Ontology Reuse and Synthesis for Modelling and Simulation

A thesis submitted for the degree of Doctor of Philosophy

by

Nurul Izrin Md. Saleh

Department of Computer Science

Brunel University London

May 2019

Abstract

The proliferation and ubiquity of Semantic Web technologies have transformed the way computer society reshapes its technology through knowledge integration, knowledge reuse and knowledge sharing. Ontology, one of the Semantic Web components, is a way to represent domain knowledge into a human-understandable and machine-readable format. Ontology in simulation has been seen as a conceptual model of a system in an explicit and unambiguous manner, where it can be applied to better capture the modeler's perspective of the domain. Regarding an ontology for simulation modeling, by reusing ontologies, it helps to reduce time and effort in attaining the domain knowledge, and at the same time assist in domain understanding. For a semantically-richer simulation ontology, it is useful to engage with real data and existing ontologies. This research contributes a rigorous method that extracts domain knowledge, synthesizes processes performed within the domain, and builds a minimal and viable ontology for simulation modeling, known as a Minimal Viable Simulation Ontology (MVSimO). The research method initially applies ontology selection techniques in Ontology Reuse Framework (ORF) to obtain suitable existing ontologies for reuse. ORF incorporates a module extraction technique during the domain conceptualization phase, where the modules will represent domain knowledge as sub-ontologies. Formal Concept Analysis is later applied to the real-world data to reveal the process details of the domain. Finally, the development of MVSimO is completed by the derivation of event semantic of the processes. The effectiveness of ontology selection and synthesizing methods, is reviewed by evaluating the selected ontology knowledge extracted, and the detailed ontological model of MVSimO. The evaluation of, MVSimO is performed to determine its agreement to the established simulation model of the domain. The evaluation results are encouraging, providing concrete outcomes of the new technique of ontology reuse and new development to the research area.

Acknowledgements

It is my pleasure to acknowledge my deepest gratitude to those who have helped along the way and influenced the formation of my understanding.

First and foremost, I owe my deepest appreciation and profound gratitude to my supervisor, Dr. David Bell for his advice, mentorship, unconditional support and encouragement throughout this research experience. Our discussions always gave me the confidence of being on the right path.

I also owe special thanks to Dr Alan Serrano-Rico, my second supervisor, for all his help, enthusiastic support and valuable input during the course of my study. I appreciate all his contributions of time, advice and guidance as well as his trust on me and the numerous opportunities he generously provided me with.

Also, I would like to thank all members of academic staff in the Department of Computer Science for their support and feedback. I owe a very special thank you to my friends and PhD colleagues in the Department, for their valuable help and continuous support during the years of my study.

Lastly, my deepest appreciation belongs to my husband, my mum, my dad, my brother and my sisters for their unconditional love, endless encouragement and unbelievable support in all of my pursuits in my life as well as for inspiring me to always follow my dreams. Without you, I wouldn't be able to reach where I am today. Thank you for this life opportunity.

Publications

The work in this thesis has led to the following publication:

- Saleh, N. & Bell, D., "Ontology derived Conceptual Modelling for Simulation", work under review with 10th Simulation Workshop (SW20).

The work of Agent-based Simulation during the early stage of the research process has led to the following publications:

- Bell, D., Kashefi, A., Saleh, N., Turchi, T., & Young, T. (2016, April). A data-driven agent based simulation platform for early health economics device evaluation. In Proceedings of the Agent-Directed Simulation Symposium (p. 5). Society for Computer Simulation International.
- Mgbemena, C., Bell, D., & Saleh, N. (2016, September). A Data-driven Methodology for Agent Based Exploration of Customer Retention. In 2016 IEEE/ACM 20th International Symposium on Distributed Simulation and Real Time Applications (DS-RT) (pp. 108-111). IEEE.

Contents

List of Figures

List of Tables

Acronyms

ABS Agent-based Simulation

A&E Admission and Emergency

ConExp Concept Explorer

DeMO Ontology for Discrete-Event Modeling and Simulation

DES Discrete Event Simulation

DEN Normalized Data Element Name

DSR Design Science Research

EOC Event Ontic Commitment

ESDF Event-Semantic Derivation Framework

FCA Formal Concept Analysis

FDiMe FCA-Relation Discovery Method

MinDO Minimal Domain Ontology

MVSimO Minimal Viable Simulation Ontology

NHS National Health Service

NLP Natural Language Processing

OQuaRE Ontology Quality Evaluation Framework

ORF Ontology Reuse Framework

OWL Web Ontology Language

RDF Resource Description Framework

SD System Dynamic

SPARQL SPARQL Protocol and RDF Query Language

SQuaRE Software product Quality Requirements and Evaluation

STP Space-Time-Process

W3C World Wide Web Consortium

Chapter 1

Introduction

1.1 Background of the Problem

1.1.1 Ontology and Semantic Web

Semantic Web architecture has been vastly characterized in the academic literature with the potential to better accommodate knowledge integration and knowledge sharing among its community. Ontology, as a part of the Semantic Web component, is a "*formal and explicit specification of a shared conceptualization*" within the domain (Gruber, 1993). With the emerging ontology-based application in fields such as data semantic management (Tao et al., 2017; Daraio et al., 2016), context modeling for manufacturing industry (Giustozzi et al., 2018), building information modeling for cost estimation (Lee et al., 2014), and simulation modeling (Traoré et al., 2018), ontologies offer an efficient knowledge representation scheme in making better-informed decisions (Sernadela et al., 2015a). An empirical study on conceptual modeling by Verdonck et al. (2018), indicates that it is significant to have sufficient knowledge prior to the development to provide more relevant application. The study shows that modelers trained with ontology knowledge performed better in modeling the advance aspect of a domain by 66.75% of total scores on content interpretation and content sophistication, compared to only 53.7% from the modelers who are not trained.

Recent developments in ontology-based applications have heightened the need for building an effective ontology. Nonetheless, the development of the new ontological model can be tedious and costly (Lonsdale et al., 2010; Grau et al., 2007; SIMPERL, 2010). Apart from that, there is increasing concern over the inadequate amount of semantic terminology in developing an ontology (Feilmayr and Wöß, 2016; Balasubramaniam, 2015; Chandrasegaran et al., 2013). The introduction of semantic content during the design stage of development leads to the construction of a well-defined model, and helps in building an effective solution (Moreira et al., 2016). The research presented in this thesis is the development of MVSimO that can be beneficial for modeling a simulation for a complex and heterogeneous application domain, such as A&E departments.

The domain conceptualization phase in building MVSmO provides the abstract view of the underlying domain. To achieve a conceptually comprehensible context of the domain, an efficient knowledge representation model is required. The decision regarding on what to model, the level of abstraction, and the perspective view of the model provide positive impacts on the overall process (Robinson, 2006). Perceivable understanding of the domain, gets more refined as it moves towards its goal to help the modeler to make better informed-decisions on what to model.

The development of MVSimO is revised at every stage to ensure the knowledge obtained by the end of the research is close enough to have a metamodel of A&E departments. The MVSimO definition is adopted from the definition of Minimal Viable Product by Ries (2009): "*the version of a new product which allows a team to collect the maximum amount of validated learning about customers with the least effort*". Regarding ontology for simulation modeling, the learning outcomes from artefacts collected uphold the execution of the research, despite the fact no expert opinions being gathered. Domain knowledge from existing ontology and generic pathways is used to perceive the domain understanding of A&E departments. The research is conducted by reusing the domain knowledge, thus saving the time and effort spent in requirement gathering and ontology design process. During the evaluation of MVSimO, an existing simulation model developed by experts is used for validation to gain an insight into the proposed work as opposed to the existing model.

1.1.2 Simulation in A&E Departments

The emergency department, widely known as accident and emergency or A&E department [1] in healthcare services plays a major role to save people's lives, and more importantly to reduce death and disease rate in public (Aringhieri et al., 2017). A&E departments refer to the department or sub-system that provides medical treatment to urgent need patients, and is the most critical unit since they are one of the first unit responsible in treating life-and-death situation (Gul and Guneri, 2015). Figure 1.1 shows the process journey in A&E departments. The figure is presented in a project "*A Better A&E*" by PearsonLloyd shows the different stages from patient's check in, assessment by staff, receiving treatment from medical staff, and the outcomes of the process.

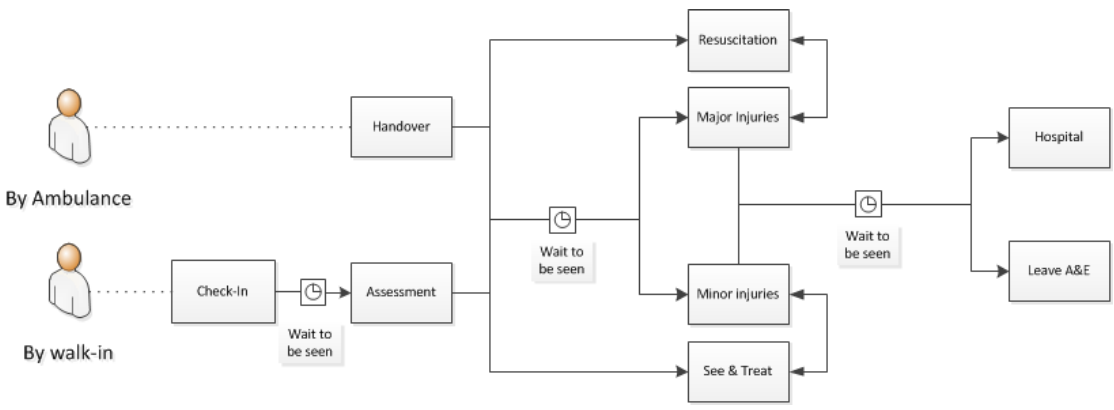


Figure 1.1: A&E Process Flow

In NHS 'Five Year Forward View' [2], the rising demand in A&E departments has caused the failure in meeting the 4-hour target despite additional funding. Stakeholders face the challenge to balance two withstanding goals: to fulfil the medical demand of an increasing population and to keep within the allocated budget (Mielczarek and Uzialko-Mydlikowska, 2012).

[1] https://www.uhs.nhs.uk/OurServices/Emergencymedicine/EmergencyDepartment.aspx
[2] https://www.england.nhs.uk/five-year-forward-view/next-steps-on-the-nhs-five-year-forward-view/urgent-and-emergency-care

The stakeholders hold the responsibility in making decisions to ensure that the services are, and to continually be accessible 24 hours a day, 365 days a year. The situation is intensified by multiple factors which include staff shortages, doctors' turnover, high bed occupancy and lack of resources in the department (Blunt et al., 2015). To investigate more about this, simulation modeling has been used by decision-makers to understand the behaviour of the complex system of A&E departments, and to provide support for the healthcare system management in an efficient use of resources for healthcare delivery (Traoré et al., 2018; Aringhieri et al., 2017; Mohiuddin et al., 2017a).

For example, to help foresee A&E departments in the real-world situation, simulation provides a model to conduct a preliminary test or trial changes for a safe and efficient care deliverable implementation (Günal and Pidd, 2009; Aringhieri et al., 2017). Simulation attributes that resemble real-world situations help decision-makers to predict the output of a proposed solution. Real-world phenomena modeled using Discrete-event simulation (Lebcir et al., 2017), System Dynamic (Pidd, 2014) and Agent-based simulation (Chahal et al., 2013) provide guidance for stakeholders in implementing new solutions. Recent surveys indicate that computer simulation facilitates the A&E model in finding optimal solution to improve patient flow and allocation strategy, in the effort to solve overcrowding (Eatock et al., 2011; Vannieuwenborg et al., 2015; Gul and Guneri, 2015; Baboolal et al., 2012; Coughlan et al., 2011).

Several simulation approaches have been proposed to facilitate various applications in A&E departments, ranging from operation and system re-design (Baboolal et al., 2012; Eatock et al., 2011), decision making in application implementation (Maull and Smart, 2009), contagious disease modeling (Luppa et al., 2011), to other feasibility study on the A&E performances in critical situations (Mielczarek and Uzialko-Mydlikowska, 2012). Though there is an increasing interest in the usage of simulation modeling techniques and tools to alleviate problem in healthcare, there are still limitations to existing models, such as the lack of research with logical component in providing more complex simulation models (Mohiuddin et al., 2017a), analysis reports in simulation are not detailed out comprehensively (Hay et al., 2006; Maull and Smart, 2009), and difficulties in deducing experts knowledge a for decision-making process (Grigore et al., 2016). Moreover, the lack of evidence shows that the proposed simulation models are not being used despite the research effort (Mohiuddin et al., 2017a)

1.1.3 Simulation and the Role of Ontology

Simulation modeling for A&E departments is required to cater for aspects such as the representing behaviours of each stakeholder and their interactions (Escudero-Marin and Pidd, 2011; Eatock et al., 2011; Ghanes et al., 2014), and analysing its properties (Isern and Moreno, 2016; Baboolal et al., 2012) to ensure the usefulness of the model in a complex and heterogeneous system. Pertaining to this idea, an ontology is generally assumed to play a role in setting up a common ground to provide knowledge sharing among subject domain (Huang, 2016) and also describe, standardize and represent an object or instance in the domain (Park et al., 2011; Schulz et al., 2012; Grolinger et al., 2012; Huang, 2016). This is because the foundation of ontology itself is a formal specification of conceptualization (Gruber, 1993). The idea can attentively define a domain using classes, properties, relationships and instances.

A number of research have been conducted to support the use of ontology with other methods to come up with a more innovative contribution to the fields of science, engineering, technology and mathematics (STEM) (Tolk and Miller, 2011). To conduct a study in the healthcare domain is not an easy task, as the data in healthcare is more complicated due to the rapid growth in technology and human population (Stroetmann and Aisenbrey, 2012). Even though most of healthcare data is publicly available, it has not been used in a way that could contribute to the healthcare community strategically (Shafee Kalid, 2017). In addition, regardless of the amount of the research effort in simulation modeling in the domain, there is not enough evidence to show that the proposed models are being used (Mohiuddin et al., 2017a). Taken together, this suggests that the aggregation of ontology-data analysis in simulation and a well-presented knowledge scheme during the design phase may help to help understand the domain and its emerging issues better, thus making the proposed model more usable and beneficial.

1.2 Aim and Objectives

1.2.1 Aims

To design an effective domain knowledge extraction method that supports the design and development of the minimal and viable ontology for simulation. The ontology minimizes the effort required to build a semantically-rich simulation model for a complex healthcare domain, A&E departments and more importantly, successful in producing the desired simulation model. To achieve this, the thesis has five objectives as listed below:

1.2.2 Objectives

- Critically assess the existing simulation modeling methods, ontology reuse framework, data analysis techniques and process mining methods to provide an understanding of the state-of-the-art.
- Design and develop a methodological Ontology Reuse Framework (ORF) with a simulation and modeling perspective, which is able to select suitable ontologies for ontology reuse.
- Produce a Space-Time-Process (STP) map from a set of generic pathways as a means to better represent the domain
- Analyse generic A&E processes and extract process elements from semi-structured data
- Derive event-semantics from the process elements extracted to further develop a Minimal Viable Simulation Ontology (MVSimO) model, and validate the research outcome with a goal-free evaluation method

1.3 Research Methodology

Design Science Research (DSR) has been adopted as a method for this research. The objective of DSR is to produce a more effective and efficient solution for a domain problem (Hevner et al., 2004). To conduct a research based on the DSR methodology, the set of methods and techniques of the activities of research should be suitable to contribute to the understanding of the knowledge and needs to be in the interest of research community (Vaishnavi and Kuechler, 2004). DSR consists of a process of getting an effective solution while adapting to changes and components, which results in products called artefacts (Hevner et al., 2004). The process involves in the design science research activities are; build, evaluate, theorize, and justify (March and Smith, 1995), and the result from the activities, which are known as artefacts, are a form of materials or properties structured into a visible manner (Orlikowski and Iacono, 2001). The artefacts of a research can be classified into four categories (March and Smith, 1995);

- **Construct:** An interpretation to define problem and solution
- **Model:** A construct to represent a problem in real-world and its solution space
- **Method:** Are used to define processes and provide guidance on how to solve a problem
- **Instantiation:** The implementation of constructs, models and methods for artefacts evaluation

The implementation of DSR allows for a search process to find effective solutions following specific requirements, in order to obtain the result for the research community. To ensure the effectiveness of the solution, DSR provides a brief and distinct mechanism of conceptual framework and standard, in understanding, developing and assessing the research (Hevner et al., 2004). The activities in DSR search for the effective solution in iteratively or incrementally until a solution that satisfies the requirement is found.

The research aims to find an effective solution to the prevailing issue; hence the adoption of DSR as a research methodology enables the research activities to iterate until it achieves its objectives (Vaishnavi and Kuechler, 2004). Through the iteration process involved in DSR, the phases applied in accomplishing the research purpose is as follows:

- **Problem Study:** This phase involves reviewing of existing literature on ontology-based simulation, ontology reuse in an ontology engineering process, domain process extraction and concept exploration in data analysis, and establishing the gap in the analytical approach of data analysis in healthcare domain
- **Recommendation:** This phase is the process of introducing the preliminary idea on conceptual modeling in ontology selection design, knowledge reuse and process and events selection for the ontology-based simulation model
- **Development:** This phase involves the development of design artefact to select existing ontology for reuse, to extract the process elements constitute in pathways and to derive an ontological model from the process-event description as research artefacts
- **Evaluation:** The evaluation process in the research involves the process of assessing the derived ontology from reusing existing ontologies and similarity evaluation by comparing the proposed ontological model with an existing simulation model
- **Conclusion:** This is the final phase of the DSR cycle for this research. The way this research helps in modeling a simulation of A&E pathways by reusing existing ontology, exploring the semi-structured data to extract process elements, and develop MVSimO in healthcare is discussed in this phase.

The iterative approach is adopted in this research, in which the research processes involve designing, building and evaluating iteratively (a little more is added each time) until the objectives are achieved. The iterative processes are the steps taken to design a construct, model, method or instantiation piece by piece, followed by building of the construct, model, method or instantiations piece by piece using suitable tools and techniques, and finally evaluating it using appropriate approaches. The processes involved in this research are as below:

- **Iteration 1**

  Core framework development for ontology reuse based on domain's specification and the evaluation of the derived ontology with identified evaluation metrics.

- **Iteration 2**

  Continuation of the work done in Iteration 1 by the creation of a Space-Time-Process (STP) map in extension of generic pathways and the analysis of semi-structured data using Formal Concept Analysis (FCA). The details of data analysis using FCA are explained. This iteration contributes an FCA-Relation Discovery Methodology (FDiMe) model and a set of process relation using first-order rules.

- **Iteration 3**

  The complete Minimal Viable Simulation Ontology (MVSimO) is developed using the semantic derivation processes. The ontological model of MVSimO and existing simulation as an exemplar model is used for similarity assessment.

1.4 Thesis Overview

In meeting the objectives of the research, the thesis is structured as follows:

Chapter 2: This chapter involves a systematic review of literature in an ontology-based simulation, emphasizing on healthcare domain, a domain where the study will be evaluated, giving a general background on problems in ontology reuse and simulation modeling and steps taken to overcome the problems. Ontology reuse framework, ontology-based simulation challenges and domain conceptualization are also discussed in this chapter. This chapter proceeds by discussing the methods to extract the ontology knowledge analytically and how conceptual models help in simulation modeling.

Chapter 3: This chapter discusses the selection of Design Science Research as the research method for conducting the research effectively. It elaborates on why DSR has been selected as a method to conduct this research, how to apply DSR in the research, the processes involved in DSR for the research, the artefacts, the framework, the theory and the evaluation conducted.

Chapter 4: This chapter presents the first iteration of DSR, the first task of developing the Ontology Reuse Framework (ORF), as well as the steps involved in the framework, namely Domain conceptualization, Ontology Discovery, Ontology Selection, Ontology Merging and Integration, and Ontology Evaluation. All five phases are explained in detail. This chapter also presents how the final candidate ontology, the Minimal Domain Ontology (MinDO), is obtained through the framework, in accordance with the simulation modeling for healthcare domain. The output of the iteration is presented as a set of DSR artefacts. An evaluation of the products is then performed. This chapter concludes with the presentation of the learning outcomes and discussion of future improvement.

Chapter 5: This chapter presents the implementation of the second DSR iteration. Here, the existing data of the A&E departments are refined and analysed in order to extract the process properties. It outlines the FCA-Relation Discovery Methodology (FDiMe), which is used to guide the steps in obtaining the process properties using the FCA approach. In between the steps, a new set of 3-dimensional map is derived to illustrate the processes in the department to depict the events in space and time dimension blocks. The map is known as the Space-Time-Process (STP) map. This chapter also presents the elements extracted from the processes, which are transformed into process relation to support the development of MVSimO.

Chapter 6: The last research iteration is executed here to improve and validate the generality of the framework and model. The full development of MVSimO is made by adding the event semantics, applying the process relations produced in Iteration 2 and combining the ontology elements of MinDO from Iteration 1. The simulation model is evaluated against an existing expert model, to measure its effectiveness. Appropriate metrics are used to measure the model's precision.

Chapter 7: The thesis is concluded with the presentation of the contributions and key findings of this research. Limitations that were discovered from applying the DSR to solve the proposed problem are also explained. An evaluation of the DSR process is performed in its ability to meet the research aim and objectives, highlighting the research limitations. Lastly, relevant conclusions will be drawn against the degree to which the proposed approach meets its objectives, while an explanation of the research limitations suggesting future improvements is presented.

Figure 1.2 shows an overview on how the research is conducted from defining and analysing the issue, formulating the problem and setting the objectives, defining the research methodology, conceptual validation (Chapter 4), data analysis and knowledge extraction (Chapter 5), operational validation (Chapter 6), and to the conclusion.

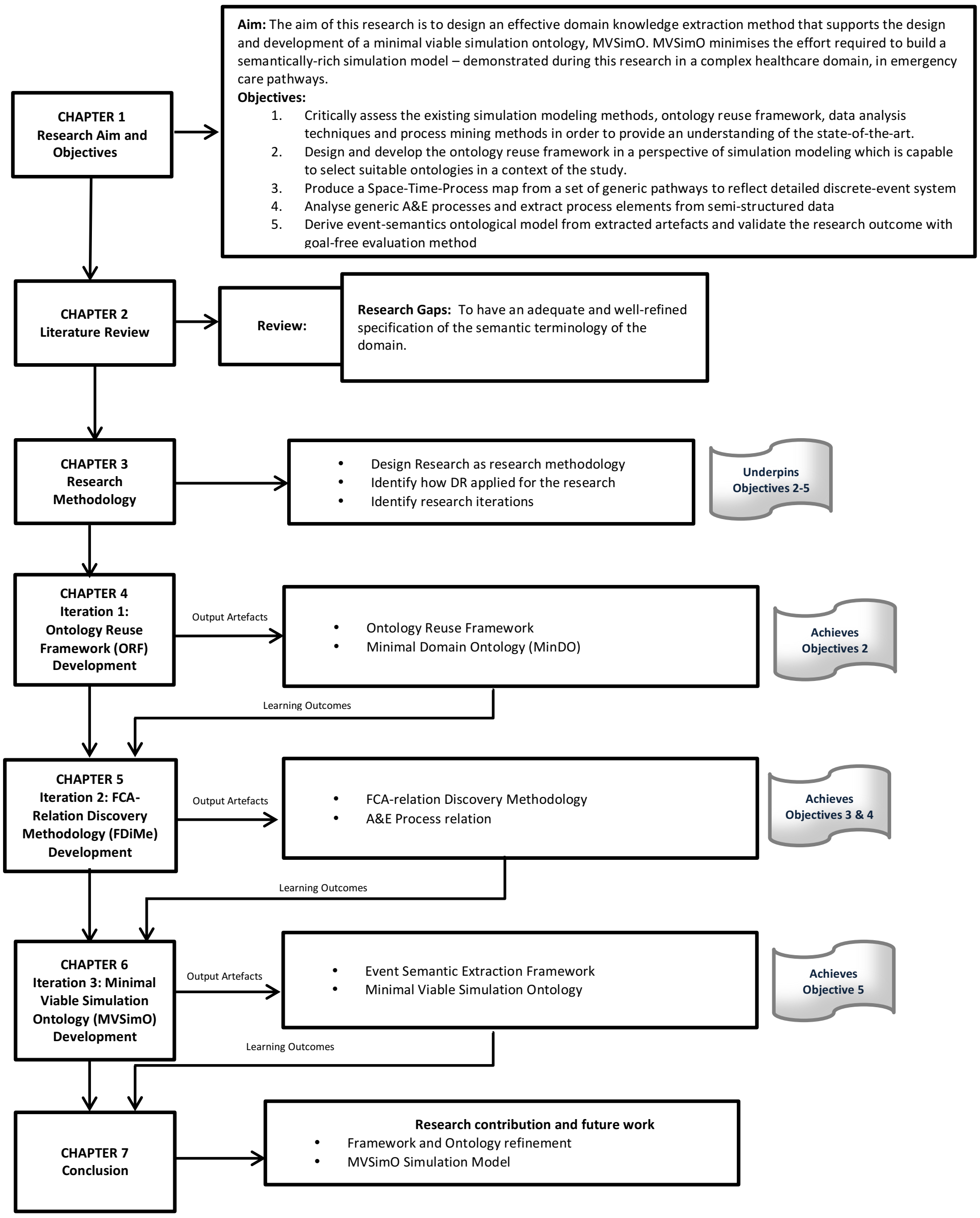


Figure 1.2: Thesis Overview

Chapter 2

Literature Review

2.1 Introduction

Recent studies in accomplishing knowledge sharing and reuse in diverse domain application can be achieved by using ontology as knowledge representation, which provides an effective way of defining a domain with a set of concepts, characteristics and relationships. The collaboration of ontology with an application such as simulation model, allows interoperability across platforms, organizations and operating systems. In a rapidly changing requirements environment, reaching the full potential of ontology-based simulation model could be challenging. Literature has shown that by adding ontology into simulation modeling, a semantically richer and more expressive model can be developed. This chapter reviews relevant literature on achieving ontology-based simulation model, discussing the challenges in incorporating ontology semantic into requirement specification, and suggestions on how to improve the design process. Ontology reuse frameworks offer a good starting point for existing ontology selection, followed by a formal concept for semantic-process enhancement. Mapping of the ontology to the A&E pathways is discussed and analysed for the purpose of developing an effective approach for a refined modeling formality.

This chapter is structured as follows. Section 2.2 describes a general review of Semantic

Web and Ontology, introduces the need for ontology reuse and semantic knowledge representation in developing an ontology-based simulation model. Section 2.3 presents a broad overview of frameworks and tools for ontology reuse, and the conceptual modeling for knowledge representation. Section 2.4 discusses simulation models in healthcare and the challenges of simulation modeling. Section 2.5 describes the need for adding semantics into simulation modeling as a way for advancing the simulation model limitation and review existing literature to identify the most important approaches in the field. Section 2.6 classifies existing semantic extraction approaches concerning the techniques applied, and the disciplines from which these techniques are borrowed. Section 2.7 introduces the application of ontology-based simulation, detailing current work in the area, highlighting issues and challenges, as well as suggesting improvements.

2.2 Achieving Ontology-based Simulation

In an extension of the classic "*Web of documents*", Semantic Web technology consists of a stack to support a "Web of data", where information is given a well-defined meaning, encouraging interaction between human and computers (Berners-Lee et al., 2001). Semantic Web technologies extend the vision of W3C to have a Web of linked data. Semantic Web technologies permit the creation of data stores on the Web, build vocabularies, and write rules for data manipulation. Linking of data in the Semantic Web is empowered by technologies such as RDF, SPARQL, OWL, and SKOS [1]. Through the increasing need for shared semantics and a web of data, this is being achieved by the adoption of common conceptualization known as an ontology (Shadbolt et al., 2006). An ontology can be defined as an "explicit, machine-readable specification of a shared conceptualization" (Studer et al., 1998).

Over the past decade, most research in Semantic Web has placed an emphasis on the use of ontology to understand systems across ranges of scale and distribution in various applications. With the growing number of complex and heterogeneous applications and high user

[1] https://www.w3.org/standards/semanticweb/

demands, the current research has shifted from data-information level to human-semantic-level interaction. Hence, the ontology-based applications such as simulation modeling (Traoré et al., 2018), semantic search and question answering system (Collarana et al., 2016), process mining (Sarno and Sinaga, 2016; Okoye et al., 2018), data semantic management (Tao et al., 2017), and internet-of-thing (IOT) (Abinaya et al., 2015) are taking over. In general, the number of published research on the use of ontology is evident in the Semantic Web community, providing a standard requirement for data and information integration.

Ontology-based application is an architectural approach with the potential to better accommodate deep level of interoperability, provided by publicly available data on the internet. In the field of healthcare, ontologies are used as a knowledge representation of a domain, as well the relations amongst attributes of the domain. (Okhmatovskaia et al., 2012) introduced SimPHO, an ontology for simulation modeling of population health. SimPHO is an explicit machine-readable specification of domain knowledge integrating both aspects of taxonomy and vocabulary in the form of logical axioms. Silver et al. (2007) developed an ontology-based simulation model that supports links between a domain ontology and a simulation ontology. An ontology for discrete-event modeling and simulation (DeMO) by Silver et al. (2011), provides taxonomies for a discrete-event simulation model that captures the essential features of the real world system. Recent study by Traoré et al. (2018) discusses on O4HCs, an Ontology for Healthcare System Modeling and Simulation, which is a domain analysis model based on System Entity Structure (SES), proposed by Zeigler (1984). O4HCs aims at facilitating communications amongst modelers, ensuring they share common semantics.

A number of approaches that work in conjunction with either an ontology or a simulation modeling to support an ontology-based application, have also been identified in the literature. The literature for each of the approaches (as below) will be discussed in detail in later sections:

- **Conceptual Modeling:** In the early stage of ontology reuse, the process of domain conceptualization to find which ontology to be selected to best represent the domain are performed based on the ideas of conceptual modeling and simulation modeling. A conceptual model observes all aspects of the simulation model development by abstracting a real-world to a simulation model. A well-designed conceptual model enhances the possibility of a more structured simulation model (Robinson, 2008b).

- **Process Mining** In the implementation of an ontology-based representation system, process mining to model the processes in a real-world system helps in structuring the ontology. Essentially, an ontology refers to the process and the representation of the domain model (Thomas and Fellmann M.A., 2009).
- **Formal Concept Analysis (FCA)** provides an illustrative analysis of datasets, resulting in more manageable and readable concepts to describe the data. This also contributes to a more structured knowledge representation for simulation modeling (Andrews and Orphanides, 2010; Wray and Eklund, 2014).

Moving towards an ontology-based simulation modeling can be conceptualized as a semantic layer being added on to a simulation modeling. It intends to give a machine-readable and well-defined semantic representation of a model (Berners-Lee et al., 2001; La-Ongsri and Roddick, 2015; Traoré et al., 2018). It is also agreeable by Heath et al. (2009) that a conceptual model is an important feature in the simulation development process where it validates system theories and assumptions before the concept is being transferred into a machine-readable model. Even though a proposed semantic representation is used to depict the model, they are not enough to provide a full assumption of the process. The problem lies in connecting the real-world interpretation of the process and the machine-readable representation (ontology) (Thomas and Fellmann M.A., 2009). The formalized concept of the model provides a better way of representing concepts and relations than just in a taxonomy order (Obitko et al., 2004).

2.2.1 Ontology

Literature clearly defines that ontologies form an important component of Semantic Web (Balasubramaniam, 2015; Garcia, 2009; Ford, 2011). Ontologies can be classified by different criteria, and the most prevalent criteria are generality and level of detail (Nicola Guarino, 1998; Nicola Guarino and Staab, 2009). The first type of ontology based on level of generality is general or foundational ontologies. This type of ontology formalize notions that go beyond particular domains (Van Heijst et al., 1997; Mizoguchi et al., 1995). Foundational ontology plays a critical role in ontology development by giving developers a guideline of how to view the target domain, and providing the basic concepts for the development of any domain-specific ontology (Guizzardi and Wagner, 2004). Examples of foundational ontologies are: COmmon Semantic MOdel (COSMO)[2], Descriptive Ontology for Linguistic and Cognitive Engineering (DOLCE)[3], Suggested Upper Merged Ontology (SUMO)[4], Basic Foundational Ontology (BFO)[5], and Yet Another More Advanced Top Ontology (YAMATO) (Mizoguchi, 2010).

The second type of ontology is the domain-specific and the task-specific ontology that describes concepts which belong to specific domains. Domain ontologies and task ontologies outline the vocabulary related to a generic domain such as the Social Web (Peñalver-Martinez et al., 2014), System Engineering (Tolk et al., 2014; Batres, 2017), Neonatal (Farinelli et al., 2016), Medically Related Social Entities (Hicks et al., 2016a) and Biomedical Data (Sernadela et al., 2015b) or a generic task for example Clinical Decision Support System (Bau et al., 2014; Wilk et al., 2013), Data Semantic Management (Tao et al., 2017), Conceptual Modeling (La-Ongsri and Roddick, 2015; Verdonck et al., 2015, 2018) and Simulation Modeling (Guizzardi et al., 2015; Benjamin and Graul, 2006; Traoré et al., 2018). Typically, a domain ontology is a model of a domain-specific definitions (Fernández-López et al., 2013; Corcho et al., 2004).

The third type is the application ontology, where ontologies are used to represent a conceptualization of a particular domain for a specific task. Task ontologies for an ontology-based application comprise of domain ontology or task ontology, to achieve the interoperable applications that provides semantic interoperability and knowledge exchange (Benjamin and Graul, 2006). It can be derived from a foundational ontology (Khan and Keet, 2012) and used

[2]http://micra.com/COSMO/COSMO.owl
[3]http://www.loa-cnr.it/ontologies/DOLCE-Lite.owl
[4]http://www.ontologyportal.org/translations/SUMO.owl
[5]http://www.ifomis.org/bfo

as domain knowledge by playing a role on semantic domain which gives meanings to requirements statements (Kaiya and Saeki, 2006).

With an increasing amount of available ontology on the Semantic Web, however, incorporating multiple ontologies from multiple sources to support ontology reuse and sharing requires an infrastructure to make them available for the community. Ontology infrastructures such as search engines, registries and repositories are systems that gather ontologies from different sources, and act as a middleware to support ontologies finding and exploring for ontology reuse (Baclawski and Schneider, 2009; d'Aquin and Noy, 2012). Ontology repositories used in previous studies such as Swoogle (Ding et al., 2004), Falcons (Qu and Cheng, 2011), Watson (D'Aquin et al., 2007), OntoSearch (Jiang and Tan, 2006), OntoKhoj (SIMPERL, 2010; Qu and Cheng, 2011), OntoSelect (Buitelaar et al., 2004), SQORE (Anutariya et al., 2007), AktiveRank (Duque-Ramos et al., 2013; Cantador et al., 2007), BioPortal (Whetzel et al., 2011; Baclawski and Schneider, 2009; Amith et al., 2018), and Ontobee (Xiang et al., 2011; Hicks et al., 2016b) are listed in literature. For ontologies to perform their role is facilitating interoperability for the ontology community, the repositories must be accessible and return a desirable ontology result.

2.2.2 Simulation Modeling

Since the introduction of computers, computer simulation has become popular for its ability to mimic a real-world system to create a better understanding of a problem. The emerging of Agent-based Simulation (ABS) model gaining its popularity in the 1990s has made a transition from the more traditional approach of Discrete-event Simulation (DES) model Heath et al. (2009). ABS has been adopted to model complex systems comprising agents for exploring emergent behaviour of a system Anagnostou et al. (2013); Baker (2015). ABS models incorporate independent parameters to represents individual behaviour of the system components (Liu et al., 2017). As opposed to ABS, DES is used to simulate a complex and dynamic system, in which the variable state changes continuously (Varga, 2001). Moreover, DES models allow for different resource allocations for model observations (Cimellaro et al., 2017; Komashie and Mousavi, 2005; Akbari Haghighinejad et al., 2016). A broader perspective of simulation modeling has been adopted by Pidd (2014) who observed how System Dynamic (SD) model investigates system structure and feedback control for grasping system behaviour. Table 2.1 summarizes the types of simulation models and the tools used.

The rapid growth of simulation software technology comes with big opportunities, which include more revolutionary applications in fields such as operational research, computer science, process modeling and healthcare. In the context of this thesis, this sections discusses simulation modeling in Healthcare domain and A&E departments, and also issues in the department in brief. Due to the growing number of patients and the increasing demand of effective healthcare services, A&E departments face a significant pressure to provide the best for the community (Manley et al., 2016; Dakin et al., 2015). One of the major issue faces by the departments is overcrowding (Saghafian et al., 2015; Baboolal et al., 2012). Overcrowding is defined as limited healthcare resources to cope with the increasing demand from patients (Di Somma et al., 2015).

Simulation Type	**Characteristic**	**Tool**
Discrete-Event	Models series of events at discrete time intervals, focusing on operational characteristics	ARENA, Programming Language, SIMUL8, MedModel, ExtendSim
System Dynamic	Models interrelations between different units in a complex system	VENSIM, Ithink/Stella, DYNAMO, Programming Language
Agent-Based	Perform some kind of behaviour in a shared environment	NetLogo, REDsim, Repast symphony

Table 2.1: Types of Simulation Model

A number of simulation studies have been carried out in order to understand how simulation model helps to solve overcrowding in A&E departments. Integrated simulation and data development analysis by Aminuddin and Ismail (2016) is a study on resource allocation in determining the A&E performance, while Akbari Haghighinejad et al. (2016) on the other end focuses on queuing theory and simulation modeling to reduce the waiting time in A&E departments. A survey by Mohiuddin et al. (2017b) is an attempt to investigate the contribution of various computer simulation for the analysis of patient flow in A&E departments.

Apart from finding a solution to overcome overcrowding problem, a number of research have been carried out to cover the topic of providing good healthcare service, particularly in A&E departments. The topics range from point-of-care testing (POCT) to reduce the cost in emergency department (Wilk et al., 2013) using early health economic device evaluation (David Bell, 2016), Big Data analysis to increase patients' privacy risk (Li et al., 2015), mathematical modeling to increase interaction between patients and their environment (Lee et al., 2011), to simulation in helping decision-making processes (Eatock et al., 2011; Vannieuwenborg et al., 2015; Gul and Guneri, 2015). For stakeholders, the decision-making process is the key in defining which solution will help in overcoming the issues in A&E departments (Marshall et al., 2016). Overall, these studies highlight the need for an efficient simulation to help in decision-making process.

2.2.3 Ontology and Simulation Model

A framework by Uschold and Jasper (1999) has been adopted by Charles Turnitsa (2006) to define four types of ontology usage in simulation modeling:

- Ontology-based search: For the discovery and selection of domain knowledge or component
- Neutral authoring: For data exchange between systems
- Ontology as specification: For specifying the definition and process of the domain, to guide decision-making
- Common access to information: Extending the purpose of neutral authoring, allowing knowledge of the data to be transmitted

In the first and second usage, ontologies play a role in search, select and representing significant knowledge for knowledge exchange (Beck et al., 2010; Leidig et al., 2011). Ontologies in the third and fourth category will result in artefacts that are primarily intended to benefit the systems' users; in the specification for development (Saghafian et al., 2015) and provide a meaning for knowledge based systems respectively (Tolk et al., 2015). In modeling a healthcare system, a survey by Isern and Moreno (2016) reports that use of ontological representation of medical staff and patients and the dependencies amongst them are the typical use of intelligent agents in Healthcare organizations-centred applications. The agent-based simulation models normally employ shared ontologies among entities in the system. For discrete-event simulation model, DeMO is introduced by Silver et al. (2011) comprises of four first-level subclasses, based on DES principal: state-oriented model, event-oriented model, activity-oriented model, and process-oriented model.

On a different note, Durak and Ören (2016)develops an ontology for simulation systems engineering based on IEEE standard of Distributed Simulation Engineering and Execution Process (DSEEP), to create a common shared conceptualization of products, information exchange, data stores, roles, steps, activities and tasks of the process. To support simulation component reuse, Trajectory Simulation ONTology (TSONT) by Durak et al. (2011), has been developed as a conceptual model for trajectory simulations. The trajectory simulation model is constructed by conforming to TSONT. Another study by Grolinger et al. (2012) introduces application-oriented simulation packages to facilitate comparison among simulation models, querying, making inferences and reusing of existing simulation models, where the simulation models are represented as instances of the ontologies.

2.3 Why Ontology Reuse?

Constructing ontology from scratch is a tedious and costly task (Lonsdale et al., 2010). With a significant number of ontologies that have been built during the last decades (Pinto and Martins, 2000), a large amount of ontologies are available on the web. An ontology provides a knowledge-sharing infrastructure that supports the representation and sharing of domain knowledge (Park et al., 2011; Pinto and Martins, 2001; Trokanas and Cecelja, 2016; Kamdar et al., 2017). This feature of ontologies allows researchers to develop cost-effective and high quality ontologies instead of building new ones. However, finding an ontology that satisfies all of the user's needs according to pre-defined requirement is not possible (Park et al., 2011). This is due to the fact that each application is unique to the domain that utilizes the ontology, the purpose of the ontology, and also the objectives of the application that the ontology is used for.

It is agreed that the divide and conquer strategy to build an ontology should use the same approach as Software Engineering Borst (1997). A study by Borst (1997) stated that domain knowledge should be divided into "small, manageable pieces with strong internal coherence but relatively loose coupling". The approach of separating domain knowledge into small manageable ontologies leads to the idea of ontology reuse. This mechanism build an ontology by combining parts of other ontologies where "assembling, extending, specializing and adapting" the candidate ontologies into one final ontology Pinto (1999b). Simperl (2009) in the study about reusing ontologies on the Semantic Web, has listed reasons why a relatively large number of ontologies with the same or related domains of interest, are not being shared or reused, though they are accessible. This is because the ontology is being developed independently, or the ontologies are too general, thus demanding a modification before it can be reused for the intended purpose. These situations give an indication that more work could be done to promote ontology reuse and sharing.

2.3.1 Ontology Reuse as Part of Ontology Engineering

The decision to reuse fragments of ontologies from existing ontological sources, is one of the activities performed in an ontology engineering process SIMPERL (2010). As pointed out by Lonsdale et al. (2010) "*Ontology reuse involves building a new ontology through maximizing the adoption of pre-used ontologies, or ontology components*". In view of all that have been mentioned so far, this section emphasises that ontology reuse is an integral part of ontology engineering. Of all the steps of ontology engineering, in the context of this research, four processes are highlighted (Ontology Discovery, Ontology Selection, Ontology Merging and

Integration, and Ontology Evaluation).

For the development of an ontology-based simulation, finding a suitable ontology to achieve the simulation objective based on pre-defined requirements (SIMPERL, 2010) is similar to the process of conceptual modeling, where the abstraction process involves some level of simplification of the real-world system (Zeigler et al., 2000). The reason is because, pre-defined requirements necessitate the simplification of the system with minimal, easy to evaluate requirement, and capable of making a feasible component of the model. Hence, suggesting that the process of discovering potential reuse candidate ontologies corresponding to the domain conceptualization will impact the development of the simulation model. Though many ontologies are available to be selected as candidate ontologies, which ontology should be chosen? Malone et al. (2016) outlined ten rules for selecting a bio-ontology which are applicable to other ontologies, because the rules covered common domain and are not only for biologists and bioinformaticians. In addition to finding candidates ontologies, selecting an ontology requires the user to understand the domain requirements.

With the growing utilization of ontologies, the problem of overlapping knowledge is unavoidable (Stumme and Maedche, 2001). To overcome the problem, Stumme and Maedche (2001) suggested a method of FCA-MERGE to merge two or more source ontologies. The method extracts unique and non-repeatable instances of source ontologies given a domain-specific text documents. It then applies the mathematical method of FCA, to produce a lattice of concept and the produced result is explored and transformed into merged ontology by experts. The process of ontology merging involves the building of an ontology in one area of the domain by reusing two or more different ontologies from the same domain. Whereas ontology integration is the process of building an ontology in a domain by reusing one or more ontologies from different domains (Pinto, 1999a).

Ontology evaluation is a step to assess the functional adequacy and to guide developers or researchers in making informed decisions on which ontology to be reused under given criteria (Cantador et al., 2007). A study by Duque-Ramos et al. (2013) has outlined an ontology evaluation characteristics known as an Ontology Quality Evaluation Framework (OQuaRE).

For the ontology quality metrics, OQuaRe has reused existing metrics developed by ontology engineering community, with the following notion:

$C_1;C_2;\ldots;C_n$: Classes of the ontology.
$R_{C1};R_{C2};\ldots;R_{Ck}$: Relationships of the class C_i.
$P_{C1};P_{C2};\ldots;P_{Cz}$: Properties of the class C_i.
$I_{C1};I_{C2};\ldots;I_{Cm}$: Individuals of the class C_i.
$Sup_{C1};Sup_{C2};\ldots;Sup_{Cm}$: Direct superclasses of a given class C_i.
Thing: Root class of the ontology.

Ontology Evaluation Metrics

- Tangledness (TMOnto): Mean number of classes with more than one direct parent.
  $TMOnto=\Sigma|R_{Ci}| / \Sigma|C_i|-\Sigma|C(DP)_i|$; where C_i is the i-th class in the ontology and $C(DP)_i$ is the i-th class in the ontology with more than one direct parent

- Attribute Richness (AROnto): Mean number of attributes per class.
  $AROnto=\Sigma|Att_{Ci}| / \Sigma|C_i|$

- Relationships per class (INROnto): Mean number of relationships per class.
  $INROnto=\Sigma|R_{Ci}| / \Sigma|C_i$

- Number of Children (NOCOnto): Mean number of direct subclasses. It is the number of relationships divided by the number of classes minus the relationships of Thing.
  $NOCOnto=\Sigma|R_{Ci}| / (\Sigma|Ci| -|R_{Thing}|)$

- Number of properties (NOMOnto): Number of properties per class.
  $NOMOnto=\Sigma|P_{Ci}| / \Sigma|C_i$

2.3.2 Ontology Reuse Framework

Knowledge representation structures for an efficient and effective reuse of ontological knowledge require methodologies and tools to perform the ontology reuse processes. The ontology reuse framework or methodology maximises the adoption of existing ontologies or ontology components, in a way to abstain from creating new ontology, which is a tedious and costly task (Pinto, 1999b). Gómez-Pérez and Rojas-Amaya (1999) discussed about re-engineering work of ontology reuse to capture the conceptual model of the implemented source ontologies to transform the more detailed ontology. The re-engineering method proposed by Gómez-Pérez and Rojas-Amaya consists of three steps: *Reverse engineering*, *Restructuring* and *Forward engineering*. With respect to the complexity of the source ontology and the need for automatic means, the study was restricted to several hundreds of concepts. Nevertheless, the proposed method was conducted manually and has taken 18 months to complete with the small size of ontologies.

On a different perspective, a study by Lonsdale et al. (2010) presents a generic architecture for automated ontology reuse to fit the scope of natural language web pages. The reuse process takes two inputs; natural language (NL) documents and source ontologies. The NL documents provide a description of targeted ontology for the automated ontology reuse system to identify all the required domain concepts. Then the relations, concept and constraint of new ontology are automatically gathered from the source ontology. The system reuses source ontologies to create small domain ontology within the scope description of NL documents. Though the results of the study were encouraging, the relationships and constraints reuse were not discussed in detailed. It has been suggested that the adaptive reuse of sub-ontologies into modeling point of view creates a new perspective to ontology matching (Stecher et al., 2008). The adaptive ontology reuse approach suggesting similar existing sub-ontologies, and provides an alternative for the user to integrate them with the draft ontology. The adaptive algorithm focuses on ontology discovery, ontology matching and automatically propose an alignment to the user.

Park et al. (2011) conducted a study on ontology selection and ranking model comprised of selection standards and metrics grounded by semantic matching capabilities. The study aimed to overcome problems of finding a suitable ontology that meets users' requirements and complementing semantic matching capability to adapt the ontology selection and ranking model. The model selected ontologies from ontology repositories (OntoSelect, OntoSearch2, Swoogle, OntoKhoj, AktiveRank) based on keywords, and rank the selected ontologies according to pre-defined semantic similarity measures. The semantic similarity is measured using concept matching, relation matching and taxonomy matching. The measurement factors such as the relevancy of search keyword to the concepts defined in the ontology, and the usefulness of the search keyword to obtain the ontology, are used for the study.

A descriptive research by Kamdar et al. (2017) was conducted to study the current extent of term reuse and overlap among biomedical ontologies stored in BioPortal repository. The study estimated the term reuse and term overlap, extracted reuse patterns from BioPortal ontologies; and extracted reuse patterns from time-stamped BioPortal Import Plugin logs using method motivated by text mining, graph theory and unsupervised learning. The result from the study has shown that most ontologies reused less than 5% of the terms from a small set of popular ontology, and have not used terms from the unpopular ontologies. It has been showed that ontology developers used incorrect representation to reuse despite the intention to reuse. In contrast to the studies by Gómez-Pérez and Rojas-Amaya (1999) and Lonsdale et al. (2010), this study focused on biomedical ontologies in BioPortal, and the result was based on Bioportal Import Plugin Logs.

In understanding the applicability of ontologies from process's point of view, SIMPERL (2010) provided a set of guidelines for ontology reuse, by investigating how the objective of the ontology has impacted the development of a particular application. Prior to the ontology reuse process, a domain requirement analysis was conducted to form the basis requirement for the potential ontology reuse candidates. The requirement provides requirement specification that entails important features of the final ontology. In her study, SIMPERL (2010) adopted the ontology engineering *Conceptualization, Implementation and Evaluation* paradigm as part of the reuse process. The first step, which is identified as *Ontology Discovery*, is where a list of candidate ontologies are obtained with the help from Semantic Web-specific search engine. The second step is the *Ontology Evaluation and Selection*. During this stage, an in-depth evaluation of the candidate ontologies was conducted, to select the reusable ontologies. The final stage is the *Ontology Integration*.

Providing a criteria of the participant and context of eReqruitment solution provider, the author conducted the study that led to the adoption of domain and ontology-specific approach. The method was evaluated through a case study research, and compared with existing ontology reuse approaches, using the method of *goal-free* evaluation, the method of evaluating different solutions of the same problem against a set of pre-defined criteria (House, 2010; Scriven, 1991). SIMPERL (2010) has concluded that the ontology reuse process is a process in which the final new ontology comprised of input generated from a set of existing ontology knowledge. Regardless of its context-oriented focus, the study demanded experts' involvement in the system analysis and design to execute this research. The studies presented thus far provide evidence that the conceptualization of the domain does not start as early as the design phase and if so, it was not discussed in detail. Table 2.2 shows the Ontology Reuse Frameworks discussed in this chapter.

Framework	Method	Domain
Gómez-Pérez and Rojas-Amaya (1999)	Reverse Engineering Restructuring Forward Engineering	Formal Ontologies: DOLCE-Lite, SUMO-OWL, OpenCyc, BFO, Sowa's
Lonsdale et al. (2010)	Natural Language and source ontologies for Automated Extraction of new ontology	Ontologies of Web page content
Park et al. (2011)	Ontology Selection Ontology Ranking	Ontologies from repositories: OntoSelect, OntoSearch2, Swoogle, OntoKhoj, AktiveRank
Kamdar et al. (2017)	Term Reuse Term Overlap	Ontologies from BioPortal repository
Simperl (2010)	Conceptualization Implementation Evaluation	eReqruitment Solution Provider

Table 2.2: Ontology Reuse Framework

2.4 Ontology-based Simulation Challenge

In the context of this thesis, the discussion on the challenges faced by ontology-based simulation modeling focuses on the issues raised by previous studies in the aspect of 1) simulation modeling in healthcare domain and 2) semantic enrichment of the simulation model. For simulation modeling in healthcare domain, the literature review was conducted by extending the preliminary work on the impact of simulation model to the delivery of A&E departments by Mohiuddin et al. (2017a). The study shows that the selective use of the available data makes the analysis and decision to define the objects and processes involved in the model unclear. In addition, conceptual models for domain requirements are lacked of semantics of the terminology of the underlying model Verdonck et al. (2018). By its very nature, getting the suitable ontologies for ontology-based simulation is an expensive and time consuming task, therefore an effective ontology-based simulation requires an efficient knowledge representation scheme, where knowledge is captured from earlier design decisions (Robinson, 2013; Chandrasegaran et al., 2013).

In view of all that have been mentioned so far, the main challenge is minimising the effort in developing a simulation ontology, which will have its validated learnings about the domain are collected throughout the research iterations. Thereby, enabling domain conceptualization in the ontology selection phase to offer a good starting point for clarification of its real-world semantics and simplifies the model, leading to a higher overall quality, and for a wider use of the simulation Guizzardi and Wagner (2010); Wu et al. (2016). Applying knowledge extraction technique to a set of semi-structured A&E data to complement the selected ontologies allows the user to add more semantic meaning to the existing ontological structure, for the purpose of facilitating the process of simulation ontology development(Wray and Eklund, 2014; Bau et al., 2014).

Semantic meaning need to be added at different levels; therefore, to obtain the ontology that faithfully represent the domain knowledge, it is important to integrate the domain understanding iteratively. The literature highlights the importance of domain conceptualization for a faster ontology development. Adopting prominent formalism in data analysis task for knowledge discovery adds more meanings to the knowledge Kuznetsov and Poelmans (2013a). Deriving semantic content from syntactic domain descriptions and representing such semantics in ontological models provides process-oriented aspects to the ontological model to reach its full potential Bell et al. (2007).

2.5 Domain Conceptualization

One of the most difficult parts in modeling a simulation is understanding the real system and determining the abstraction of a model from a real-world (Robinson, 2013). In fact all simulation models are an abstraction of real-world model (Zeigler et al., 2000). Conceptual modeling is a process to determine the model limitation on what to model and what not to model (Robinson, 2006). In developing an ontology-based application, a conceptual model is represented by a domain ontology where it gives semantic meaning to the requirements (Kaiya and Saeki, 2006). Ontology as a fundamental of domain-specific conceptualization (Guizzardi, 2005), have benefited the ontologist, the simulation modeler and researcher through the formalization of model semantics, the ability to query and infer, as well as in the reuse of the developed model (Grolinger et al., 2012; Kontopoulos et al., 2013).

A simulation model with a clearly-defined semantic of a real-world produced is of a higher quality model that it is clearly understood to work with other systems, easy to maintain, and has the capacity of adaptive evolution (Guizzardi and Wagner, 2010). The domain knowledge is represented by domain ontology, where it gives semantic meaning to the requirements (Kaiya and Saeki, 2006). Robinson (2008a) concluded that a well-designed conceptual model enhances the success of a simulation study. Zeigler et al. (2000) outlined the elements in a conceptual model for simulation modeling, namely the 'experimental frame', 'base model' and 'lumped model'. The experimental frame is a set of specific input–output behaviours. The base model is a hypothetical explanation of the real system, producing all possible input–output behaviours based on system requirements. In the lumped model, the components of a model are grouped and simplified. The aim is to generate a model that follows the requirements and reproduces the input–output behaviours with feasible solution.

To date, various method have been developed and introduced for conceptual modeling. The Unified Foundational Ontology (UFO) developed by Guizzardi et al. (2015) is a foundational ontology with an aim for advanced conceptual modeling. UFO is grounded by the results associated by OntoUML (Guizzardi, 2005), an ontologically well-founded version of UML 2.0. In a study of an ontology-driven conceptual modeling (ODCM) by Verdonck et al. (2015), a conceptual model is defined from formal ontology, cognitive science and philosophical logic to develop engineering artefacts. Both studies (Guizzardi et al., 2015; Verdonck et al., 2015) are performed on an ontology-based conceptual model which required prior knowledge of ontology to build new artefacts (UFO and ODCM). A study by (Guizzardi, 2007) presents the relations between domain conceptualization, abstraction, modeling language and model. The relations is visualized in Figure 2.1.

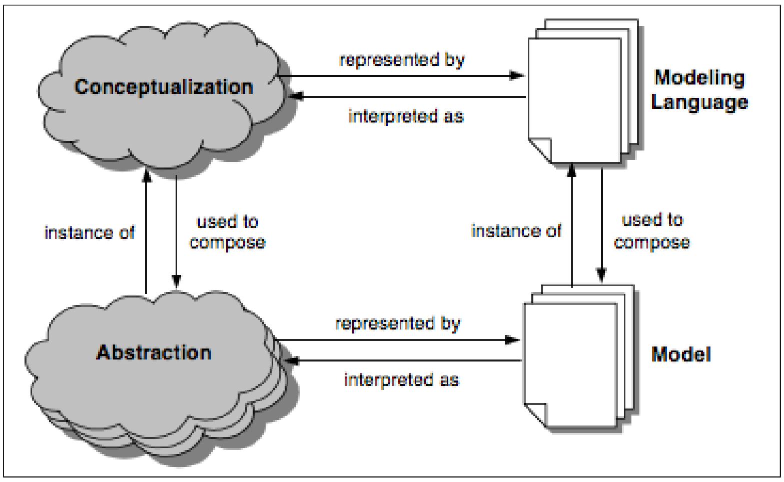


Figure 2.1: Relations between Conceptualization, Abstraction, Modeling Language and Model (Guizzardi, 2007)

Referring to Figure 2.1, the studies reviewed so far begins with *Modeling Language* (e.g ontology) interpreted by *Conceptualization* (conceptual model) used to compose the *Abstraction* of the domain *Model*. However, all the previously mentioned studies have not captured the abstraction of *Model* (domain) as an instance of *Conceptualization* that will be represented by the *Modeling Language*. For the purpose of ontology reuse and simulation modeling, an early made conceptual model with adequate specification of semantic terminology of the domain may lead to the consistent interpretation and efficient use of the knowledge (Verdonck et al., 2018).

2.5.1 Process Mining

Apart from ontology reuse and simulation modeling, this chapter aims to reveal the domain conceptualizaton process to work cooperatively with ontology reuse to achieve the objectives of this research. This is based on the ontology design idea presented by (Noy et al., 2001) that a simulation modeler makes design decisions based on the functional property of the domain to be modelled, while an ontology engineer makes design decisions based on the structural properties of the ontology class (Noy et al., 2001). The decision of structural property of the domain is discovered during the ontology development process, whereas the functional property of the domain can be discovered prior to the selection of existing ontology, in the course of understanding the domain. In understanding the domain and managing the available knowledge, accepting the fact that knowledge is often transformed from real-world system processes is inevitable (Labrousse and Bernard, 2008). Therefore, process mining is carried out to gain an insight into what happens in A&E departments.

The study by Arcelli and Christina (2007) defines that the capability of a simulation model can be improved over time to meet the rapidly changing requirements by improving the perception of the domain. Overall, there seems to be some evidence to indicate that information gained from process mining can be useful to understand the domain, and it is important to understand the activities involved for a well-presented simulation ontology. Afterall, series of domain processes organize the relationships among objects in the domain (Goodale et al., 2012). Recently, process mining has become an approach to discover significant information from the event data logs and the derived process models (Okoye et al., 2018). In a domain like healthcare, as pointed out by Poelmans et al. (2013) a process mining "is a bottom-up approach that tries to gain an understanding of the as-in process realities that are existing at the operational work floor". Günal and Pidd (2009) suggested that the principal inputs to the simulation model are requirement data, details of emergency department resources and process and its process times. Together these studies provide valuable insights into the introduction of process mining as a basis for domain conceptualization and abstraction.

2.6 Domain Knowledge Acquisition and Processing

Extracting, structuring and organizing domain knowledge is one of the key activities in simulation modeling. Apart from the existing ontology knowledge, information is obtained from data sources, to be converted to domain knowledge. In a recent study in evaluating the resilience of hospitals to disaster using DES Cimellaro et al. (2017), analysis input data of interviews with medical staffs were used to develop an emergency plan, while the data provided by hospital database was used to determine patient arrival distribution. A study by Khanna et al. (2016) is performed to identify optimal inpatient discharge time target by resconstructing the A&E pathways using patients' transfer data. Data from routine information systems provide the baseline for the arrival patterns in a study by Mould et al. (2013). The simulation built helps to identify the missing data as an evidence for evidence-based management. The data obtained not only assisted in the simulation model development, but also helped in understanding the domain through data analysis processes (Günal and Pidd, 2009).

2.6.1 Formal Concept Analysis

Data sources used as inputs for simulation models ranged from primary data, secondary data and experts' opinions (Eatock et al., 2011; Jahangirian et al., 2014). The functional and operational elements of the domain can be obtained by understanding the dependencies among the attributes in the data. Formal Concept Analysis (FCA) analyzes the dependencies between a set of objects and a set of attributes (Davey and Priestley, 2002). Not only describing the dependencies between attributes, the hierarchical property of concept lattices in FCA also makes a good foundation in defining the structural property of the applied domain (Kuznetsov and Poelmans, 2013b). FCA was first introduced as a mathematical perception for concept formalization and conceptual thinking by Wille (1982). In the field of software engineering, FCA is used to organize the concept-knowledge representation for software process appraisal and reasoning purpose (Roongsangjan et al., 2017). In the knowledge discovery and processing application, FCA organizes features and software product configuration in product development (Carbonnel et al., 2016). FCA also helps in the matrix evaluation in a software project, as a supporting tool for processing large, complex and dynamical data supplemented with additional knowledge (Kester, 2016). In information retrieval, FCA is combined with fuzzy ontology framework for conceptual clustering, data analysis and knowledge representation (Balasubramaniam, 2015).

The input data in FCA is represented in a matrix form known as the *formal context* (refer example in Table 2.3). *Formal context* comprises rows indicating the objects of the domain and columns indicating the attributes of the domain. From the table, an entry containing x

indicates that the corresponding object *coffee* has the corresponding class (attribute). For example, if the objects are coffee name as *Caffé Latte* and the attribute is from class *Milk Foam*, then *x* indicates that the particular coffee has milk foam. A cross-table with an empty entry indicates that a particular object does not have a particular attribute. For example, *Cappuccino* does not have *whipped cream*.

	Coffee	**Milk Foam**	**Stream Milk**	**Whipped Cream**	**Chocolate Syrup**
Espresso	x				
Cafe' Latte	x	x	x		
Capuccino	x	x	x		
Mocha	x		x	x	x
Macchiato	x	x			

Table 2.3: Example of Formal Context

Derivation operators in a formal context link object and attribute subsets and are used to define *formal concept* (Kuznetsov et al., 2007) (see Table 2.4). FCA produces results in two set of output data. The first set is a hierarchical relationship of all the established concepts of the domain known as the *concept lattice* (refer Figure 2.2). The second set of the output is the list of all the interdependecies among the attributes in the *formal concept*. The followings are the formal definitions of FCA:

- **Definition 2.1** Formal context

  A formal context is a triplet *(X, Y, I)* where *X* is a set of objects and *Y* is a set of attributes and *I* is a binary relation between *X* and *Y*, i.e., $I \subseteq X \times Y$. $(x,y) \in I$ indicates that the object *x* has attribute *y*.

- **Definition 2.2** Intent and Extent

  Let (X, Y, I) be a context, X' $\subseteq X$ and Y' $\subseteq Y$, the function Intent maps a set of objects to the set of attributes, whereas the function Extent maps a set of attributes to the set of objects:

Intent (X') = y ∈ $Y' \mid \forall y \in Y'$, (x,y) ∈ R

Extent (Y') = x ∈ $X' \mid \forall x \in X', (x, y) \in R$

For X' ⊆ X, Intent (X') is the set of attributes owned by all objects of X', and Extent(Y') is the set of all objects that own the attributes Y'. The two functions form a Galois connection and formal concepts.

- **Definition 2.3** Formal Concept

  A Formal Concept C in a context is a pair (X', Y') that satisfies Y' = Intent (X') and X' = Extent(Y')

  i.e., C is a Formal Concept ⇔ for X' ∈ $C and Y' \in C$, Extent(Intent(X')) = X',

  and symmetrically, Intent(Extent(Y')) = Y'.

- **Definition 2.4** Implications

  An implication A ⇒ B holds in (X,Y,I) if and only if B ⊆ A", which is equivalent to A' ⊆ B'. It then automatically holds in the set of all concept intents

I	**Coffee**	**Milk Foam**	**Stream Milk**	**Whipped Cream**	**Chocolate Syrup**
Espresso	x				
Cafe' Latte	x	x	x		
Capuccino	x	x	x		
Mocha	x		x	x	x
Macchiato	x	x			

Table 2.4: Example of Formal Concept

The corresponding formal context (X, Y, I) contains the following formal concepts:

C0 = (Espresso, Caffé Latte, Cappuccino, Mocha, Macchiato, Coffee),

C1 = (Caffé Latte, Cappuccino, Coffee, Milk Foam, Steamed Milk),

C2 = (Mocha, Macchiato, Coffee, Steamed Milk),

C3 = (Mocha, Coffee, Steamed Milk, Whipped Cream, Chocolate Syrup),

C4 = (Caffé Latte, Cappuccino, Mocha, Macchiato, Coffee, Steamed Milk).

The construction of line diagram or concept lattice can be done manually or by using software like Concept Explorer or ConExp (Yevtushenko, 2000), QuDa (Grigoriev and Yevtushenko, 2004), Conexp-clj [6], Tockit (Framework for Conceptual Knowledge Processing) [7], ToscanaJ (Becker et al., 2002), OpenFCA (Borza et al., 2010), and FCART (Neznanov and Kuznetsov, 2013).

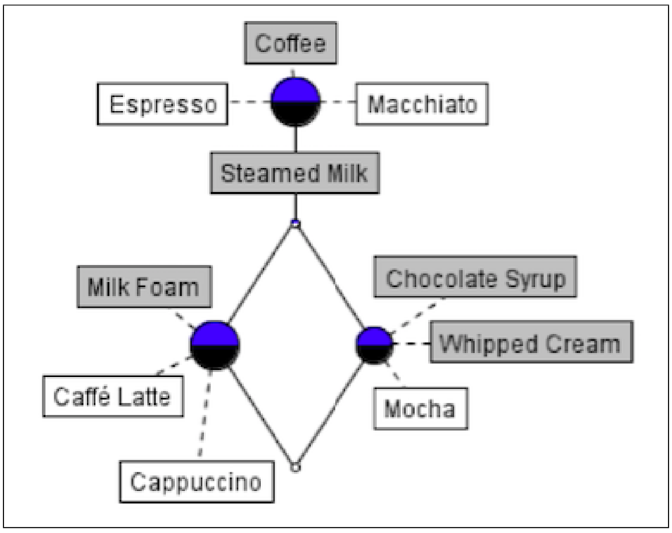


Figure 2.2: Example of Concept Latice

Figure 2.2 shows a concept lattice for coffee example built by Concept Explorer. Attribute interdependencies or attribute implications are derived from concept lattice. Concept Explorer software use Duquenne–Guigues to generate the basis of formal context, named implication basis. An implication asserts a certain relationship between two attribute sets, called premise and conclusion by the definition of "*an implication is valid in the data set if every object that has all attributes from the premise of the implication also has all attributes from its conclusion*" (Kuznetsov and Obiedkov, 2008).

2.6.2 Concept Exploration

The concept exploration task is conducted to integrate and organize the ontology knowledge of healthcare topic into simulation modeling concepts. Formal concepts are the conceptual

[6] http://daniel.kxpq.de/math/conexp-clj/
[7] http://www.tockit.org

representations of relationships between ontology classes. The conceptual result of formal concept is concept lattices, that provides hierarchically organized domain knowledge based on the class and properties of the ontology. The design of ontologies and the aim of the result of FCS supplement each other while aiming at concept modeling (Cimiano et al., 2004). In providing the formalism of a concept, concept lattice provides clear interpretability and describes the data, mutual relationships, similarities, as well as inconsistencies better than other classification, clustering and probability techniques (Cigarran et al., 2016). In conjunction with the Semantic Web initiatives, La-Ongsri and Roddick (2015) have conducted a research on a richer conceptual model by incorporating ontology-based semantics into conceptual modeling. With the growing number of web services, the use FCA formalism to leverage and organize the structured topic correlation has been studied by Aznag et al. (2014). While Yao et al. (2009), on the other hand have conducted a study on reducing the dimensionality of input data before applying the FCA for association rules mining.

2.7 Summary

Literature has illustrated the importance of a sound understanding of the domain, indicating the realization of the benefits of an existing ontology, and its capability of building a new ontology for a feasible simulation model. Understanding a variety of domain data sources and analysing the role of ontology in the Semantic Web have provided a deeper understanding of the need to apply conceptual modelling in order to advance the ontology up-take. The literature review has shown the different classification of ontology reuse techniques and approaches, and identified their applicability to complex and heterogeneous domain. It is highly acknowledged in the literature that ontology development is a costly and time-consuming process, requiring the services of highly qualified experts both in ontology engineering and the domain of interest. Widespread adoption of ontology development can be very difficult to achieve. Ontology reuse can assist in this direction by introducing domain conceptualization through conceptual modelling to extract the important elements of the domain, to ensure the success of the simulation model built from the ontology.

With the growing number of ontologies and data resources available, literature has also shown that an efficient knowledge representation scheme in ontology helps designers to organize resources and make better-informed decisions to model a simulation (Chandrasegaran et al., 2013; Kettouch et al., 2015; Verdonck et al., 2018). Since most of the research of ontology for simulation are carried out by developing a foundational ontology, there has been less work completed on the mixing techniques and developing a simulation ontology by iteratively and analytically extract the domain knowledge from existing sources. Consequently, combining conceptual modelling technique, process mining approach and Formal Concept Analysis (FCA) that deals with structuring different format of domain data to machine-readable ontology format to build a simulation model, remains an open research area.

Chapter 3

Design Science Research Methodology

3.1 Introduction

A set of systematic activities or rules agreed by the research community need to be applied to the research activity that is considered suitable for the extraction and evaluation of the knowledge gained. In a field of diverse discipline such as the Information System, there are a number of research methodologies to accommodate for the requirements of the research problems. These methodologies have their own characteristics in conducting a research. As for this research, a method known as the Design Science Research has been selected to achieve its aim. This chapter investigates and presents Design Science Research in a detailed context, techniques, phases and structure which entails the execution of the research. Design Science Research employs a set of techniques to implement research in Information Science. The notion of Design Science Research includes applying knowledge in producing artefacts, as well as analysing the use and potential of the designed artefact. Adopting Design Science Research as a valid methodology for Computer Science research, provides the justification for the selection of Design Science as the framework to conduct this research.

This chapter introduces the background of Design Science Research and the justification of encompassing design in research method in Section 3.2. Section 3.3 highlights Design Science Research as a general methodology for information system research. It provides a state-of-the-art of information system research frameworks and outlining the important strategy in those frameworks. Section 3.4 presents Design Science Research evaluation criteria associated with Design Science Research artefacts and evaluation methods. In Section 3.5, outlining the design plan of this thesis and explains how Design Science Research is applied in this study. Section 3.6 presents the research evaluation of artefacts produced throughout this research. In Section 3.7, illustrates the Design Science Research iterations involve in this research. Finally, in Section 3.8 presents the overall summary for this chapter.

3.2 Design Science Research Background

Information System design is defined as "the purposeful organization of resources to accomplish a goal" (Hevner et al., 2004). The field of Information Systems(IS) is a multidisciplinary field extends of varies disciplines such as computer science, management science, engineering and others (Baskerville and Myers, 2002). IS research conducted using a variety of research approaches, techniques, methods, methodologies and paradigms. Hevner et al. (2004) described research as producing new and innovative artefacts to research problem. Edelson (2002) and Winter (2008) perceive Design Science Research as a productive perspective of research development thus can be incorporated into a broader situation that leads to design science. Simon (1996) in his paper "The Sciences of The Artificial" differentiate the concepts of behavioral science and design science, in yielding the meaning science of the artificial. Simon introduced that artefacts lie between the inner and outer activities in the search of finding a solution that achieved the fundamental objectives with satisfactory design rather than an optimal solution. Here, the design is the learning process of producing an explicitly applicable artefact to a problem.

March and Smith (1995) introduced the idea of Design Science Research as two-dimensional frameworks that incorporates design and natural research activities with Design Science Research outputs, as in Figure 3.1

Research Activities

Research Outputs	Build	Evaluate	Theorize	Justify
Constructs				
Model				
Methods				
Instantiation				

Figure 3.1: A Research Framework (March & Smith 1995)

Adapting March and Smith's framework, the research outputs classification can help in identifying an appropriate procedure to build, evaluate, theorize and justify the research. The artefacts of research can be classified into four categories;

- **Constructs:** Constructs are sets of concepts or terms that form a specific knowledge within a domain to define problem and solution (Hevner et al., 2004).

- **Models:** Models use constructs to describe a real-world situation of the design problem and its solution space (Hevner et al., 2004); models can be used to define associations between constructs (March and Smith, 1995).

- **Methods:** Methods are a set of steps that define the solution space. They provide guidance on how to solve problems using the constructs and the models. Methods can be thought of as methodological tools that are created by design science and applied by natural scientists (March and Smith, 1995). Methods are used to define processes and provide guidance on how to solve problem

- **Instantiation:** Instantiations are the implementation of constructs, models and methods for artefacts evaluation. They prove the feasibility and effectiveness of the models, methods and constructs allowing for an actual evaluation (March and Smith, 1995). Instantiation plays an important role in enabling researchers to learn about the working artefact in a real-world scenario. As highlighted by Newell and Simon (1975), the significance of instantiations is in providing a better understanding of the problem domain and consequently to offer better solutions.

As mentioned by Takeda et al. (1990) and Owen (1998) knowledge is generated by process of knowledge manipulation and accumulation that iterates through knowledge building activities. Design means "to invent and bring into being" (Webster's Dictionary and Thesaurus, 1992) and design is considered as a process of creating a plan or convention for the construction of an object. Vaishnavi and Kuechler (2004) presented the steps involved in a design process. Design can be employed as research that generates knowledge. A number of research attempted to link theories and design to justify Design as a research approach leading to theories (Brown, 1992; Kelly and Lesh, 2012), while others attempted to put emphasis on the learning aspect of Design Science Research, and identify the types of learning that can evolve when a researcher immerses in the design process as demonstrated by Edelson (2002).

A general methodology comprises of five phases of design as proposed by Vaishnavi and Kuechler (2004) and adopted from Takeda et al. (1990) inspire the iterative process that takes learning as a key attribute of general Design Science Research methodology. The five phases are; problem *awareness*, *suggestion*, *development*, *evaluation* and *conclusion*. The initial phase, the problem *awareness*, is the phase of problem recognition through various sources, including the current development of the issue. It followed by a *suggestion* that input a proposal and produces a tentative design. The third step, *development* of artefacts extends the tentative design that motivates learning and improvement of phase 1, the problem awareness. The design is the main contribution, and not the development of the artefact.The *Evaluation* phase is the vital stage to determine artefacts performance and the progress measured during this phase. The last step is the *conclusion*, which emphasizes on the result of the Design Science Research, which contributes knowledge to the solution capacity, or contributes to the next cycle. System development (artefact construction) is considered as a research methodology that can lead to an improved, and a more effective design when applied with other research methodologies, while at the same time making a thorough contribution to knowledge (Nunamaker Jr et al., 1990).

Conforming with utility and truth as two important aims of Design Science Research and behavioural science respectively, Design Science Research is proposed by March and Smith (1995) as well as Hevner et al. (2004) as a research framework, where IT research can occur by integrating two complementary disciplines. The first of these is behavioural science, where the research is more focused on theorizing and justifying, and the second is design science research, where the research is more focused on the building and evaluating process.

3.3 Design as an IS Research methodology

The aim of Design Science Research framework is to provide the Information System community with a Design Science Research Methodology (Hevner et al., 2004; March and Smith, 1995). To achieve the aim, a process of an iterative design cycle is employed as a problem-solving process throughout the activities in the research. A well-grounded research is attained along with the design, building and evaluation process of the designated artefact. Information Science (IS) research is no different from other research as defined by Blake(1978), where research is a "systematic, intensive study directed toward fuller scientific knowledge of the subject studied". Baskerville and Myers (2002) described Information System discipline as a reference discipline due to the successful development of its own research perspective and research tradition. Information System research is considered a multi-related disciplinary field, made up of social and natural sciences management and engineering field, circled by inter-related methods of research, which promotes persistent knowledge evolution to satisfy the multi-nature of information system field (Nunamaker Jr et al., 1990).

Research in Information Technology is commonly categorized into two types; the first type is the descriptive research, where aimed in understanding the nature of IT, and the second type is the prescriptive research, where knowledge-using activities are aimed at improving IT performance (March and Smith, 1995). For both descriptive and prescriptive, research plays a role in bringing together community, organizations and technology, therefore IS consolidates IT research. Simon (1996) made a clear distinction between natural science and science of the artificial (design science), where the first is concerned with naturally occurring phenomenon, while the second relates to artificial human-made artefacts. With this distinction being made clear, it has led the IS community to realize and justify the need for design as a research discipline that combines the two Hevner et al. (2004); Nunamaker Jr et al. (1990); March and Smith (1995).

Design Science Research as an Information Systems valid research methodology, is composed by incorporating two inter-related disciplines, design and behavioural science, in a way that a researcher engages in designing an artefact for the design science aspect; and the implication of truthfulness and usefulness of research on community and organization, hence the behavioural science aspect (Hevner et al., 2004). In design science research, the truthfulness and usefulness are important and gained through an implicit cycle between design science and behavioural science, where truthfulness is provided by the theories and the usefulness is provided by the artefacts (Hevner et al., 2004). The design cycle is executed in an iterative incremental process that can be initiated by simple conceptualization providing the necessary learning that feeds into consequent iterations, where the final iteration results is an improved product that satisfies the problem requirements and constraints.

Nunamaker Jr et al. (1990) presented an earlier Design Science Research framework that connects aspects of design and design science. In the framework, Nunamaker Jr et al. (1990) assigned system development with a central role in the research life cycle, again showing an integrated approach that includes design science as a core component in an Information Systems methodological research framework. The process for conducting the research is left for the researcher to infer. On the contrary, Hevner et al. (2004) propose a descriptive Design Science framework that incorporates both natural science and design science as shown in Figure 3.2 Research rigour can be achieved by effectively applying knowledge (theories) from the knowledge base to develop and build an IS artefact, while relevance can be accomplished by assessing whether the artefact satisfies business needs. The justify-evaluate process is used to assess the applicability of the artefact in the appropriate environment.

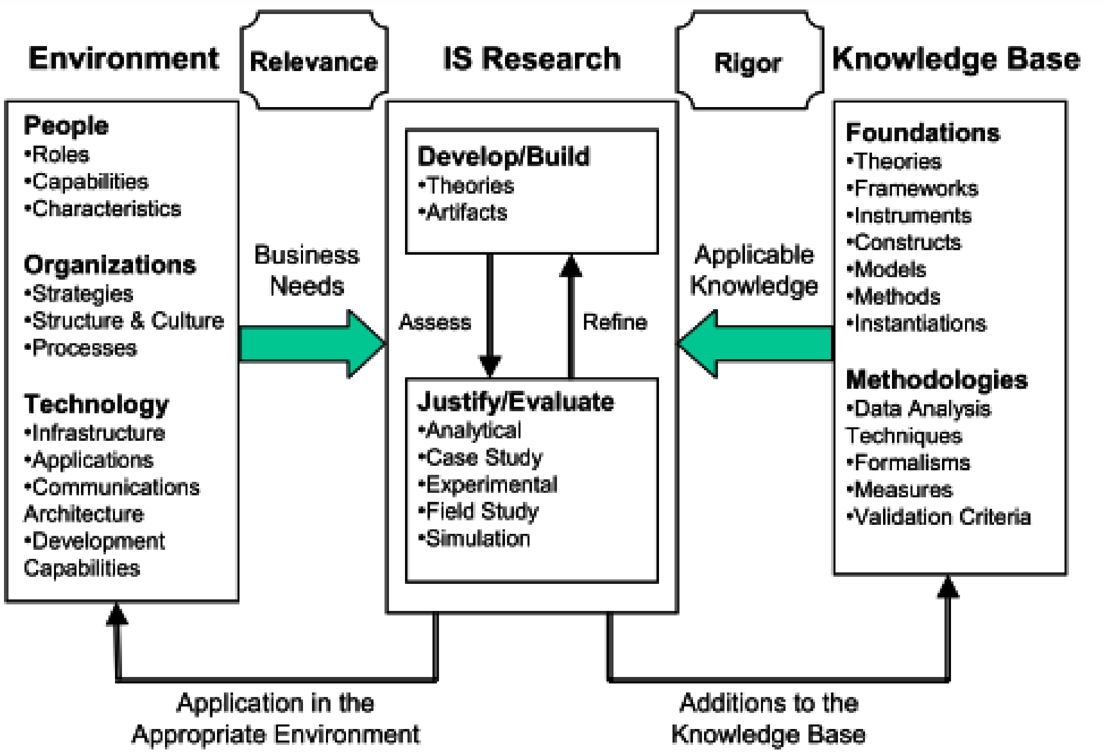


Figure 3.2: Information Systems Research Framework Hevner et al. (2004)

In (Hevner et al., 2004) a concise IS research framework is presented and used to induce Design Science Research methodological guidelines that can be followed to identify, execute and evaluate an IS research. Build and evaluate are considered to be an iterative process through which both method and product are carefully assessed by the researcher and used to assess and refine the developed product. This evaluation process typically applies measures from the knowledge base to assess the utility, efficacy and quality of the designed artefact. (Hevner et al., 2004) proposed a set of evaluation methods that can be used to evaluate the designed artefact, as discussed in the Design Science Research evaluation.

3.4 Design Science Research Evaluation

Evaluating the research artefacts is important in Design Science Research as there is a need to determine artefacts' performance and progress, measured according to designated met-

Artefact Type	Evaluation Criteria
Constructs	Completeness, simplicity, elegance, understandability and ease of use.
Model	Fidelity with real world phenomena, completeness, level of detail, robustness and internal consistency.
Method	Operationality (ability of others to efficiently use the method), efficiency, generality and ease of use.
Instantiations	Efficiency, effectiveness and impact on an environment and its users.

Table 3.1: Summarized Evaluation Criteria with Artefact Types (Hevner et al., 2004

rics (March and Smith, 1995). Problem assessment made in the problem space when the artefacts are built to perform a specific task demonstrates its usefulness, therefore validates the research. Apart from that, evaluation and validation play a role in iterative research (design science) where the knowledge generated can contribute to the next iteration cycle. Hence, developing appropriate evaluation metrics to assess the performance of the research artefacts for proving the evaluation criteria is critical (March and Smith, 1995).

Evaluation criteria proposed by March and Smith (1995) known as quality attributes, which is based on artefacts type is summarized in Table 3.1. Once the evaluation metrics and criteria are identified, an empirical work is applied March and Smith (1995), where an evaluation method is chosen appropriately. Hevner et al. (2004) emphasized that the selection of the evaluation method should be carefully considered, and when it matches with the suitable artefact, the evaluation metric are typically withdrawn from the knowledge base. An inclusive set of evaluation methodologies is summarized in Table 3.2, adopted from Hevner et al. (2004). The classifications represent the most common evaluation methods, from which a suitable method/s can be applied based on the type of artefact and the evaluation metrics used.

Design Science Research Evaluation Method Types and their Description	
1. Observational	**Case Study**: Study artefact in depth in business environment.
	Field Study: Monitor use of artefact in multiple projects.
2. Analytical	**Static Analysis**: Examine the structure of artefact for static qualities(e.g., complexity).
	Architecture Analysis: Study fit of artefact into technical IS architecture.
	Optimization: Demonstrate the inherent optimal properties of artefact, or provide optimality bounds on artefact behaviour.
	Dynamic Analysis: Study artefact in use for dynamic qualities(e.g., performance).
3. Experimental	**Controlled Experiment**: Study artefact in a controlled environment for qualities (e.g., usability).
	Simulation: Execute artefact with artificial data.
4. Testing	**Functional** (Black Box) Testing: Execute artefact interfaces to discover failures and identify defects.
	Structural (White Box) Testing: Perform coverage testing of some metric (e.g., execution paths) in artefact implementation.
5. Descriptive	**Informed Argument**: Use information from the knowledge base(e.g., relevant research) to build a convincing argument for the artefacts' utility.
	Scenarios: Construct detailed scenarios around the artefact to demonstrate its utility.

Table 3.2: Design Evaluation Methods (Hevner et al., 2004)

3.5 Applying Design Science Research

The contribution of this research is the development of ontology reuse framework to select existing ontology, the comprehensive model of A&E departments and a framework for deriving semantic event element from A&E processes and an ontological model resulting from instantiating the frameworks. To meet the research aim, Design Science Research is adopted from Vaishnavi and Kuechler (2004), as an overall research methodology. March and Smith (1995) research products classification is also adopted to illustrate the research output. Research products are identified in the form of constructs, models, methods and instantiations. The Design Science Research methodology employed for developing the research artefacts is an iterative design cycle (build and evaluate). In design science, build is concerned with the development of the artefact, and evaluation is concerned with the development of an assessment method or metric to assess the quality and effectiveness of the artefact in its context (March and Smith, 1995). The main design artefact is the ontology selection framework, an iterative process involving the five design process steps; problem awareness, suggestion, development, evaluation and conclusion, as illustrated in Figure 3.3.

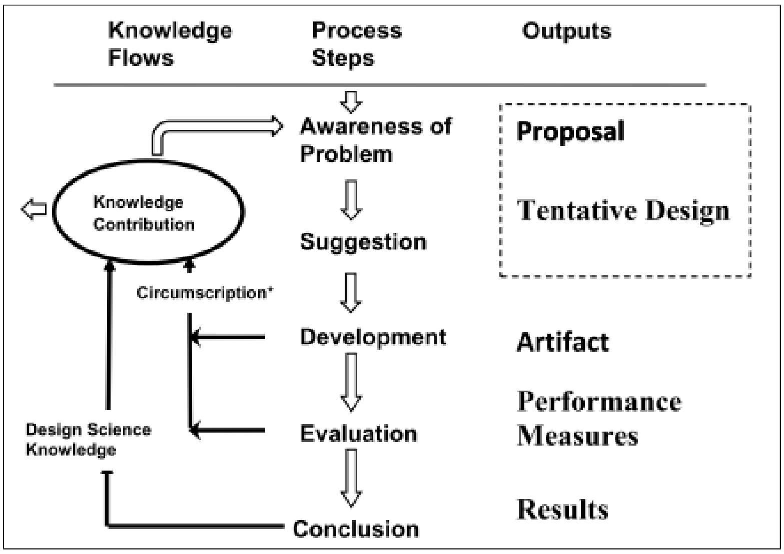


Figure 3.3: Steps of Design Science Research Hevner et al. (2004)

- **Awareness:** This phase involves reviewing existing literature of available ontology and simulation model, ontology reuse approaches, as well as domain knowledge acquisiton and processing methods, while establishing the gap in having a well-designed specification of the semantic terminology of the domain. The problem awareness was achieved in Chapter 2.
- **Suggestion:** This phase is the process of introducing the preliminary idea of how a well-defined conceptual model can help to select candidate ontologies for ontology reuse, as well as the application of Formal Concept Analysis(FCA) and event-ontic

commitment to have an explicit detail of domain knowledge. This steps began in Iteration 1, which is the development of an appropriate ontology reuse process, and making a new suggestion for process relation and event semantic in consequent iterations. As new knowledge was gained during the development and evaluation of the developed method, new suggestions from the build and evaluate cycles are used to initiate subsequent iterations.

- **Development:** The development phase was carried out as a Minimal Viable Simulation Ontology (MVSimO). The development consists of the process of reusing existing ontology, creating detailed domain mapping and extracting domain process elements. MVSimO is aimed to provide an ontological model for the creation of DES model.

- **Evaluation:** The evaluation process was carried out through an evaluation strategy that measures the effectiveness of the research, based on the significant performance improvement of the developed techniques over existing techniques, in finding the efficient way for simulation modeling in a complex and heterogeneous domain. An evaluation of the 1) Minimal Domain Ontology (MinDO), 2) MVSimO was carried out using Design Science Research evaluation criteria to examine the efficiency and generality of the techniques. In Iteration 1, to assess the efficiency and effectiveness of Ontology Reuse Framework (ORF) in A&E domain has resulted in the evaluation of the instantiation of the framework, MinDO using the ontology evaluation metrics. Finally, in Iteration 3, the final research artefact, MVSimO, was validated by comparing it with existing DES model by Bell et al. (2007).

- **Conclusion:** This is the final phase of DSR cycle for this research, where the research output was summarized, and the results of the evaluation were identified, and recommendation for future improvement was made towards improving ontology sharing and reuse for simulation modeling. Limitation of the research and suggestions for future work are also presented in this section.

3.6 Research Evaluation

Two common evaluation metrics for Design Science Research are novelty and effectiveness (Edelson, 2002). The novelty of this work lies in the development of a new ontology reuse framework model, designed to extract ontological knowledge from candidate ontologies artefacts, FCA-relation derivation method and event semantic extraction framework, and hence resulting in development of a simulation ontology. In evaluating the novelty and effectiveness of the research, DSR artefacts will need to be formally evaluated to determine whether progress has been made in the ontology reuse process and MVSimO development processes.

The effectiveness of the framework model and the derivation method is in detailing domain semantic, as well as reducing the cost and time of the simulation modeling process. When the research objective is to achieve intelligent behaviour, instantiations are used to illustrate the effectiveness and provide proof of the proposed method. It is the means through which deficiencies and improvements are identified (March and Smith, 1995). Determining whether progress is made in the ontology reuse process and in MVSimO development requires applying the appropriate metrics and similarity assessment with an established model. Due to the fact that the ontological model is the final artefact, the evaluation method can be classified according to the existing simulation model. The research artefacts can be primarily classified into two main types: (1) quality of the selected ontologies, based on specific standards, and (2) comparing the ontological model with an existing model which can be either manual evaluation by domain experts or by using the method of goal-free evaluation against an existing simulation model.

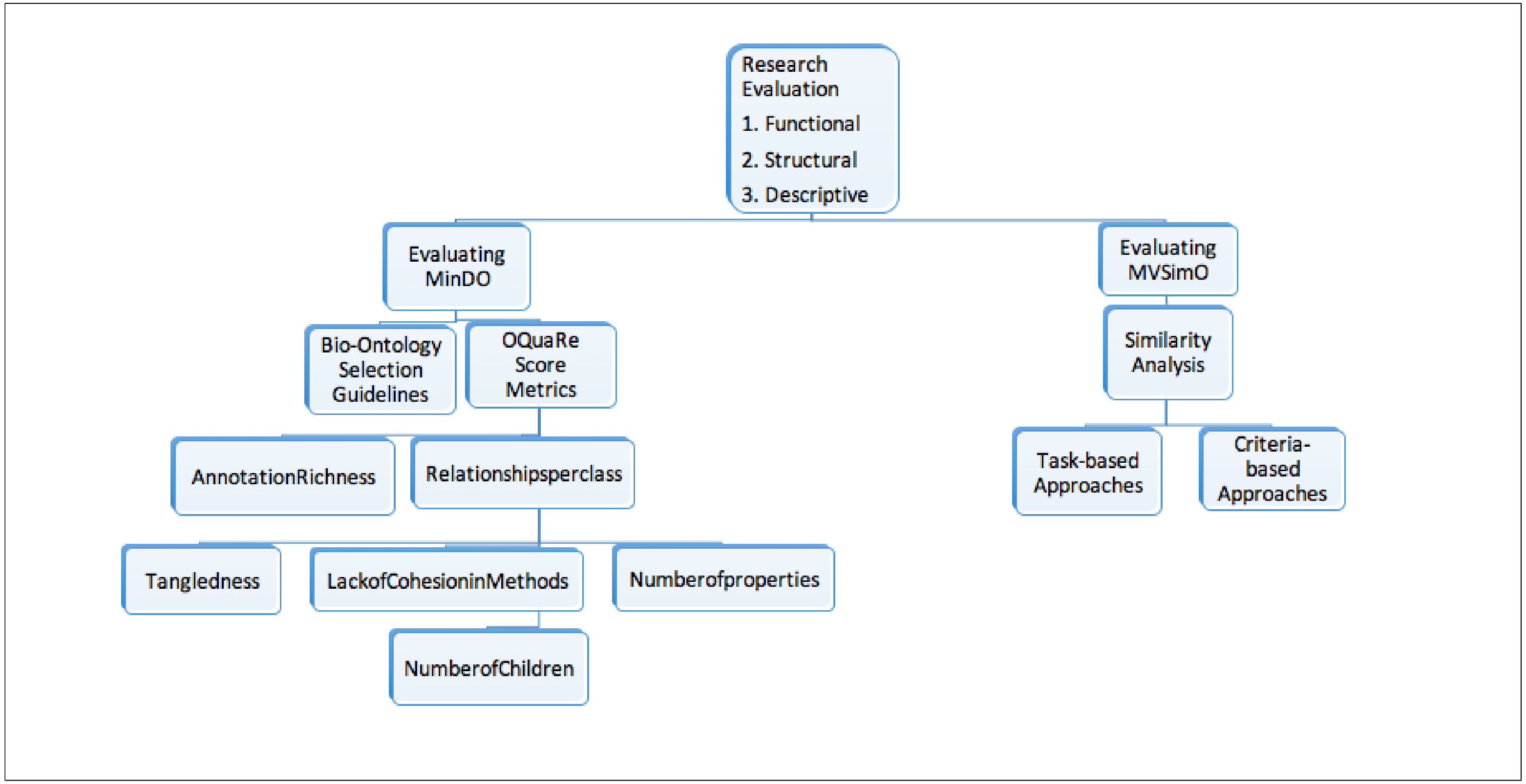


Figure 3.4: Taxonomy of Research Evaluation Approaches

Evaluation approaches can be further subcategorized according to the measure used and what they intend to measure in terms of the functional, structural and descriptive aspects as summarised in Figure 3.4. The figure illustrates how the evaluation is categorized into two main groups: first, evaluation of MinDO in Iteration 1, and second, evaluation of MVSimO in Iteration 3. The evaluation of MinDO is conducted twice; first is using Bio-Ontology Selection Guidelines by Malone et al. (2016) in Ontology Discovery phase of ORF to generally select the ontologies to ensure that the ontology is about a specific domain of knowledge and has the appropriate amount of knowledge to cover the modules. Then, in Ontology Evaluation phase of ORF, only MinDO is evaluated rigorously based on OQuaRe Quality Model by Duque-Ramos et al. (2011). For the purpose of this research, the criteria of functional, structural and descriptive are being measured in obtaining structural and functional properties of the model to assist the development of an ontology-based simulation, by which the metrics used for the evaluation strive to cover these criteria. The metrices are *Annotation-Richness*, *Relationshipperclass*, *Tangledness*, *LackofCohesioninMethods*, *Numberofproperties* and *NumberofChildren*.

The final artefact of the research, MVSimO, is evaluated using similarity analysis using *Task-based Approaches* and *Criteria-based Approaches* of goal-free evaluation. These approaches are selected ideally expecting the artefacts are appropriate and adequate to model an-ontology based simulation. Tasks and criteria of both model, MVSimO and Cumberland model (Bell et al., 2017) are assessed with the purpose of comparing different solutions of the same problem (Scriven, 1991).

Chapter 4

Iteration 1

4.1 Introduction

This iteration addresses the ontology selection task as the main component for an effective ontology reuse methodology (Stecher et al., 2008). The aim of an ontology selection process is to identify existing ontology with the most similar concept in helping developers and researchers to refine their decisions on which ontology to be reused (Minyaoui and Gargouri, 2012). Reusing ontologies saves time, effort and also the cost of ontology engineering (Lonsdale et al., 2010). Notably, and in contrast to the existing approach in the ontology reuse field, this chapter highlights the activity of introducing the real-world system component during the domain conceptualization phase, to provide more feasibility for domain-oriented ontology reuse. This iteration includes the development of an ontology reuse framework as the artefact and evaluation of the framework with quality assessment of the resulting ontology.

The rest of the chapter is organised as follows. To begin with, Section 4.2 discusses the way Design Science Research is applied for this iteration. Research artefacts are identified along with the iteration plan and research products as a design component of DSR methodology. Section 4.3 introduces the building stage of the DSR component through the presentation of a framework for domain ontology reuse. A framework consisting of the process of domain conceptualization, ontology discovery, ontology merging and integration, as well as ontology evaluation, was developed at the end of the section. Section 4.4 concludes the framework by discussing the framework evaluation. The iteration feedback and the learning outcome are presented in Section 4.6. The chapter is then summarized in Section 4.7.

4.2 Design Science Research and Output Artefacts

This iteration applies Design Science Research as a chunk of iterative processes where learning of the problem space is achieved through artefact development and evaluation. A method can be seen as a set of steps that can be followed to accomplish a particular task (March and Smith, 1995). Design science research, as the other side of the IS research cycle, creates and evaluates IT artefacts intended to solve identified organizational problems (March and Storey, 2008). From the steps, an artefact of ontology reuse framework is proposed, as a guideline in reusing existing ontologies. Provided with no prior knowledge of the available ontologies for reuse, the framework has been of assistance in the selection of the candidate ontologies by performing domain conceptualization beforehand. As illustrated in Table 4-1, an iterative cycle of artefact building, development and evaluation is employed and adopted, based on the general methodology of Design Science Research by Vaishnavi and Kuechler.

As discussed in chapter 2, a number of data sources for simulation modeling in healthcare have been identified. The applicability of the structured data, which is usually collected from hospital department and online patient's management system, make it more common to be used in researches because less effort is required to determine their tailoring ability to represent the domain (Günal and Pidd, 2009; Weber et al., 2012; Gul and Guneri, 2015). On the other hand, the unstructured data collected from experts' opinions also form part of the data resources in the simulation modeling (Komashie and Mousavi, 2005; Virtue et al., 2011). In most studies in simulation modelling in healthcare, semi-structured data are not treated as a resource in much detail (Mohiuddin et al., 2017a). This iteration aims is to systematically manipulate the semi-structured data available in existing ontology, together with the conceptual representation of the domain to build a generic concept for an A&E simulation modeling. A generic simulation model of A&E department can be built provided there is a sufficient representation of the domain (Günal and Pidd, 2009).

The concept of Design Science Research is closely associated with the idea of Minimum Viable Product (MVP). The term MVP was first coined by Frank Robinson in 2001, and was continuously evolving, later defined by (Ries, 2009) as "a version of a new product, which allows a team to collect the maximum amount of validated learning about customers with the least effort". Eric Ries proposed the idea of a combined business-driven hypothesis experimentation with iterative product release to reduce the product development process. Taking this idea into the ontology development for simulation modeling, this iteration plays a role to answer the question of "How to select suitable ontology for A&E departments?". Each iteration corroborates to the research aim, and keep iterating while constantly achieving the objectives. The final simulation ontology known as Minimal Viable Simulation-Ontology (MVSimO) created for the A&E simulation metamodel.

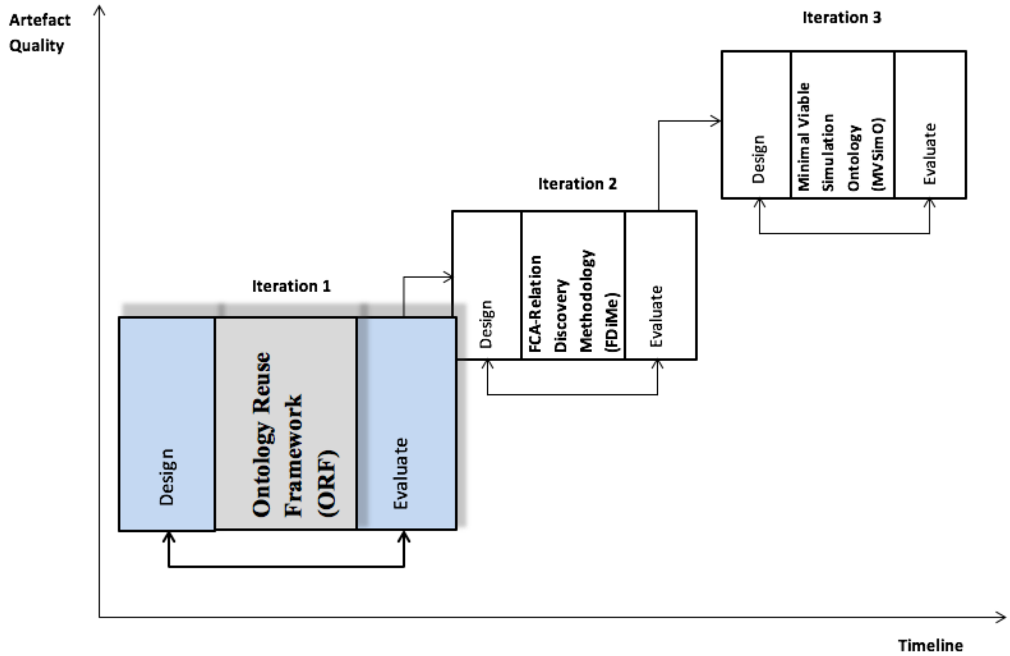


Figure 4.1: Research Iteration for Chapter 4

In line with one of the characteristics of ontology engineering; ontology reuse for knowledge sharing, the knowledge can be gained in a cost-effective way without having to develop the ontology from the beginning. This chapter provides an initial understanding of the ontology selection for knowledge reuse, and the development environment for Design Science Research cycle for domain conceptualization and ontology reuse to evolve. The fact that ontology provides a knowledge-sharing infrastructure that allows for knowledge sharing across domains, researchers and developers are able to gain benefits by sharing a common ontology for heterogeneous systems or resources to operate. However, with the vast development of available ontologies to be shared, as well as the complexity and heterogeneity of healthcare domain, it is difficult to find the exact ontology that fits the purpose. Ontology knowledge extracted from this iteration facilitates the process of simulation modeling by customary representations of the domain, and provides more refinement for a semantically-richer model.

The approach for obtaining suitable ontologies for the domain focuses on the domain knowledge and a set of standard or guidelines for the application, which in this case is a simulation modeling. In the context of this research, the focus is on the process flow in A&E. This is due to the fact that the performance in A&E is one of the contributing factors in determining the aspects for an improved procedure, reduced waiting time and increased efficiency of the employees in A&E (Sklar et al., 2010; Saghafian et al., 2015; Ghanes et al., 2014). By following the A&E process flow adapted from the project "A Better A&E" by PearsonLloyd as a guideline, this framework provides its' novelty through the use of existing ontologies for a more generic knowledge with revisions, improvements and variations to fit the intended use of the application.

4.2.1 Design Science Research Artefact

The purpose of this iteration is to develop a framework - Ontology Reuse Framework (ORF) that models an ontology selection process for ontology reuse, based on conceptual modeling terminology, and evaluation of the framework by assessing the quality of the final ontology, Minimal Domain Ontology (MinDO). The framework captures domain abstraction by incorporating techniques of an ontology-driven conceptual model to provide real-world components for simulation modeling. As illustrated in Figure 4.1, the manual process is applied in each step of the input artefact, resulting in an output that will be used as input for the next iteration. The framework in this iteration works in two folds: First, a process for selecting candidate ontologies to be reused using the domain abstraction technique. Second, implementing the evaluation techniques to assess the final ontology for reuse using the OQuare Quality Model.

The quality of ontologies is assessed using ontology evaluation method named OQuaRE (Duque-Ramos et al., 2011). OQuaRE is an evaluation method based on the standard ISO/IEC 25000:2005 for Software product Quality Requirements and Evaluation known as SQuaRE. OQuaRE elements cover the criteria for ontology evaluation: evaluation support, evaluation process and metrics. OQuaRE quality model reuses and adapts the following characteristics from software implementation to ontology implementation. Evaluation of this iteration is aimed at assessing the structural, functional adequacy, compatibility and operability of MinDO. More on ontology evaluation criteria will be discussed later in Section 4.3.4

Steps	Method	Input Artefact	Output Artefact
1.Define domain's requirement	Domain Problem Identification	Literature Review	Conceptualization of the Domain
2.Decide on a suitable Ontology Repository and search terms	Popular Search Repository	Literature Review	Repositories and Search Terms
3.Discover candidate ontology	Ontology Discovery Process	Repositories, Search Terms	Candidate Ontologies
4.Identify ontology evaluation characteristics	Ontology Evaluation Strategy	Literature Review	Ontology Evaluation Characteristics
5.Extract ontology classes and properties	Protégé and Eclipse SPARQL Query	Candidate Ontologies	Selected Ontologies
6.Combine, modify assemble and restructure ontologies	Ontology Classes Discovery	Ontology Classes, Instances and Properties	Final Ontology
7.Decide on a suitable framework evaluation method	Framework Evaluation Strategy	Framework Evaluation Characteristics	Evaluated Framework

Table 4.1: Iteration 1 - Input Output Model

4.3 Artefact Building and Development

The building stage involves problem awareness and suggestion. This iteration indicates the identification of the initial steps for the process and explaining what involves in each step. It begins with the review and analysis of existing ontology reuse methodology, assessment of ontology repositories, review of the suitable ontology evaluation method, and suggestion on the appropriate technique for framework evaluation. Review of the state-of-the-art existing standards, methods and techniques has provided a deeper understanding of the limitation of the current approach, the gaps and improvements to fill the gap. This eventually led to the identification of relevant techniques for ontology reuse in the context of simulation modeling in healthcare domain.

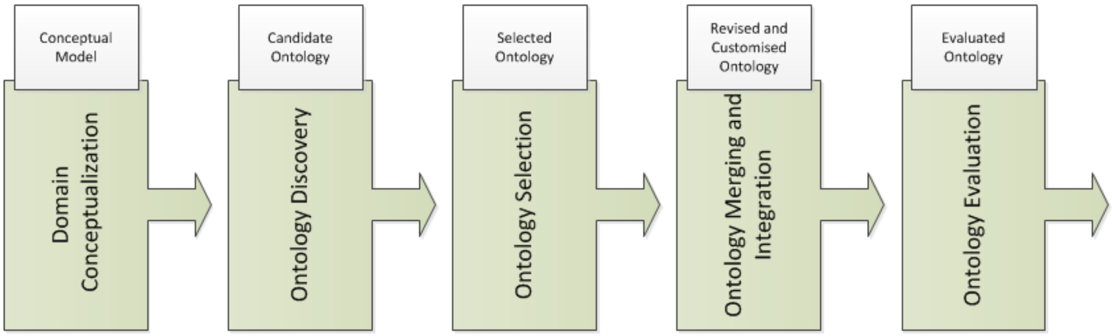


Figure 4.2: Ontology Reuse Framework

4.3.1 Domain Conceptualization

In understanding the domain, the process flow from the project "A Better A&E" by PearsonLloyd, as depicted in Figure 4.3 is adopted, as a representative model of the A&E. Stages from patient's check-in, assessment by staff, receiving of treatment from the medical staff and the outcome of the process, mimic the general process. In the UK, the A&E process flows varies from one hospital to another, therefore it is agreeable to employ a world view in general (Hay et al., 2006). In particular, understanding of the process of the A&E helps to identify key scenarios that lead to the most common problem in the UK's emergency department, namely overcrowding (Mohiuddin et al., 2017a).

As depicted in Figure 4.2, the first step of the ORF is domain conceptualization. At the domain conceptualization stage, the concept of the domain is grasped to assist in the ontology discovery process. These are outlined by the idea of a conceptual model by Stachowiak (1973). According to Stachowiak (1973), the conceptual model is composed of three main attributes: (1) a mapping feature, or a representation of a real-world system; (2) a reduction feature, a minimised version or part of the original system, and (3) the pragmatic side of the model which is designated to have a final purpose. The conceptual model helps to increase

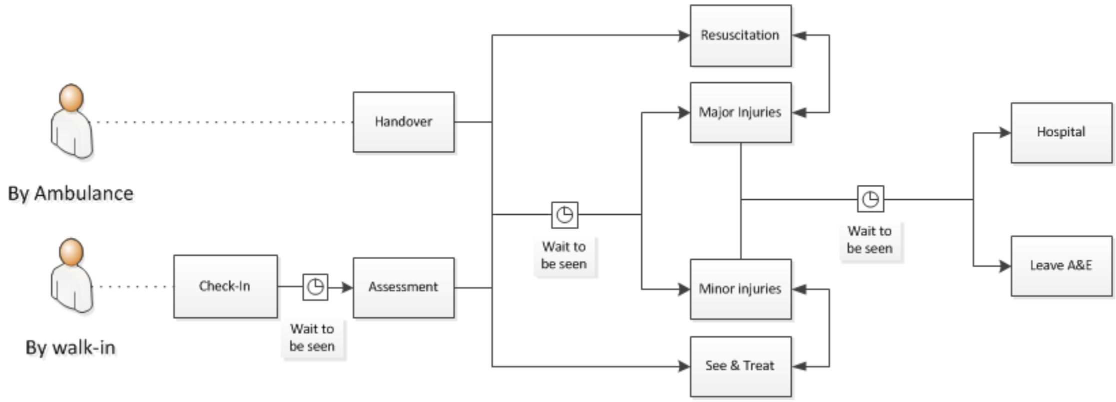


Figure 4.3: The A&E Process Flow

the understanding of the domain and functions an abstraction of the model from the real-world (Robinson, 2013). This criteria fits into the application context of the research, which is not to cover the whole NHS, but the critical part of the healthcare system, namely the A&E process flow. Based on Stachowiak (1973), the A&E department is conceptualized as below:

1. Mapping: To map the process in the A&E department, the A&E process flow (Figure 4.3 is divided into modules. For the setting of overcrowding in the A&E, the modules comprise of healthcare and hospital setting. From here, "healthcare" and "hospital" are inferred as modules.

2. Reduction: The A&E department is a large system involving several resources and heterogeneous patient types within a complex and well-organized process (Ghanes et al., 2014). Building and understanding the whole A&E system is a difficult and tedious task (Steward et al., 2017; Mohiuddin et al., 2017a). Hence, one needs to identify the assumptions and ontological commitments that each module should comply with. The description of the conceptual model is the commitments outlined in the specification requirements of future ontology.

3. Pragmatic: The purpose of the ORF is to obtain a suitable ontology to reuse for the development of a final ontology, the Minimal Viable Simulation Ontology (MVSimO). The simulation developed using MVSimO is intended to model the patient flow through the processes in the A&E or emergency department. The following processes in the ORF

will adhere to this purpose.

From these attributes, five modules are deduced; “healthcare”, “hospital”, “emergency department”, “process” and “patient data”.

In order to determine whether the deduced modules are used to model a simulation in the A&E departments, a systematic review has been conducted based on the identification and inclusion process by Mohiuddin et al. (2017a). The review extends the work by Mohiuddin, through the inclusion of literature from the year 2016-2017. Through the inclusion of the identification and inclusion criteria as illustrated in Figure 4.4, a systematic review on literature has been conducted to produce a list of objects, processes and data involved in simulation models between the year 2016 and 2017. With keywords;(1) “computer simulation” AND “emergency department” and; (2) “computer simulation” AND “patient flow”, eight bibliographic databases (MEDLINE, EMBASE, COCHRANE, WEB OF SCIENCE, CINAHL, INSPEC, MATHSCINET and ACM DIGITAL LIBRARY) were searched to investigate if any simulation methods were used to capture patient flow within A&E departments, as recognised by the UK National Health Service hospital.

Generally, 25 studies met the inclusion criteria, showing flow process of object patient and data resources from hospital records. From the study, 16 out of 25 mentioned processes in A&E departments and the same figure (16 out of 25) mentioned about patients’ data. Throughout the review, it was found that more than half of the study included *Process* and *Patient Data* in their simulation model. Overall, all the simulations in the review are in accordance to the *Healthcare*, *Hospital* and *Emergency Department* modules, as the review were conducted for simulation models in A&E bounded by the Healthcare domain. Therefore, five modules have been selected to be integrated into the next ORF phases.

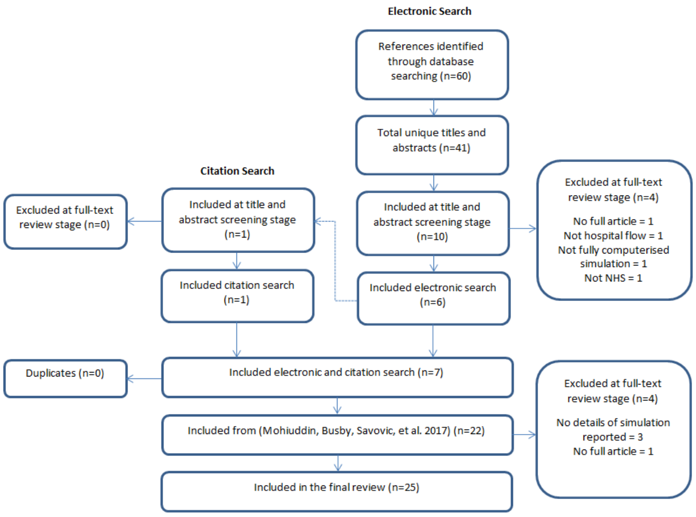


Figure 4.4: Identification and Inclusion Criteria

4.3.2 Ontology Discovery

Ontology discovery is a process to find available ontologies based on predefined concepts to represent the knowledge involved in simulation modeling. The aim of this process is to find the available candidate ontologies to be integrated into MVSimO as ontology for simulation. During the early stage of this research, an analysis on the type of simulation that best represents the domain problem in general is conducted. For agent-based simulation (ABS), codes are developed in PHP and JavaScript to parse JSON data to investigate the movement of agent or entity in A&E departments. ABS describes the system as constituent units and shows the interactions between agents Bonabeau (2002). The simulation is constructed to gain additional insight into a complex system behavior and to demonstrate the real situation for decision making. The simulation demonstrated a Sepsis patient within emergency room environment to determine patient's duration of stay based on their condition (Bell et al., 2016). The example of ABS model is shown in figure 4.5.

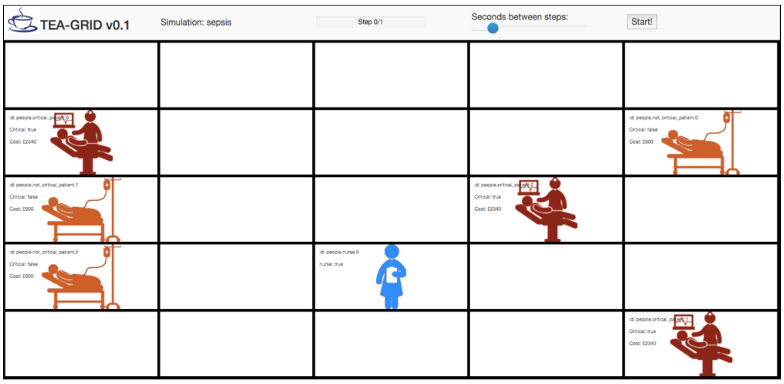


Figure 4.5: The example of ABS for Sepsis patient in Emergency Room

While the aim of this thesis is to obtain an ontology for simulation that represents the process flow of multiple departments in emergency unit rather that the behavior of one unit or department, ABS is not suitable . DES is more suitable for the reason that the underlying concepts of DES, convey the fundamental simulation paradigm (Guizzardi and Wagner, 2010),and it focuses more on queuing especially in state changes in discrete time and modeling operations and for analysis of patient flows (Bedoya-Valencia and Kirac, 2016; Cimellaro and Pique, 2016). This makes DES more appropriate in modeling a workflow and process simulation.

Why domain ontology? Benjamin et al. (2006) suggested using domain ontologies in simulation modeling process for making simulation models unambiguous and consistent. To obtain a list of existing ontologies that are domain-related, or having a similar structure with the domain of study, a comprehensive search has been carried out using the databases of ACM, IEEE, ScienceDirect and Scopus. With keywords: (1) “ontology search”; (2) “ontology discover” and; (3) “ontology reuse” from year 2011-2017, has resulted in the discovery of 40 papers and the search engines or repositories listed are Swoogle, Falcons,Watson, OntoSearch, OntoSearch2, OntoKhoj, OntoServer, OntoRank, OntoSelect, SQORE, Sindice, Semantic Search, SWSE, Dome, AktiveRank, Vocab.cc, Yars, BioPortal, Oxford and TONES. Out of these, 5 most popular ontology repositories have been ranked, based on the number of appearance in Google Scholar since 2015. These five search engines give a significance result with the web URL search using the Google scholar result. The summary of the comprehensive literature search is shown in Table 4.2 according to its popularity and Google Scholar search result.

From these repositories/search engine, only BioPortal and Ontobee are under active development, while Falcon and Swoogle were last updated about 10 years ago and Watson is unaccessible. For this reason, the search process is performed only using BioPortal and Ontobee repositories. A comprehensive search or discovery strategy supports the process of standardization by using the information that is publicly relevant. To ensure the ontologies describe the considerable amount of data for healthcare and A&E, the ontologies should be about the domain and reflect the current understanding (Malone et al., 2016).

Search Engine	URL	Google Scholar result
BioPortal	http://bioportal.bioontology.org/	588 results
Swoogle	http://swoogle.umbc.edu	159 results
Ontobee	http://www.ontobee.org/	131 results
Watson	http://watson.kmi.open.ac.uk/WatsonWUI/	51 results
Falcons	http://ws.nju.edu.cn/falcons	29 results

Table 4.2: Google Scholar result

4.3.2.1 Ontology Repositories

With an increasing amount of available ontology on the Semantic Web, however, incorporating multiple ontologies from multiple sources to support ontology reuse and sharing require an infrastructure to make the ontologies available for the community. To have current ontology knowledge, it is important to search available ontologies from search engines that are under active development.

Funded by the National Institute of Health of United States, BioPortal supports biomedical researchers in The National Centre for Biomedical Ontology (NCBO) to provide access to a library of ontologies and terminologies [1], via the NCBO web services. The web services enable multi-layered access to ontology data, capturing all terms in the ontology and extracting the metadata of the term (Whetzel et al., 2011). It provides the opportunity for the community to engage by evaluating, adding new content and submitting a review. BioPortal, a service-oriented architecture, stores more than 700 ontologies with almost 9 million classes, from various groups, not only from the biomedical organizations, but also the non-biomedical organizations. With access to multiple formats of ontologies and terminologies: Web Ontology Language (OWL), Resource Description Framework (RDF), Open Biological and Biomedical Ontologies (OBO) format, Protégé frames and Rich Release Format, the integrated web services contribute to ontology evolution.

The Ontobee [2] program was originated from a Vaccine Ontology (VO) web browser developed by The He Group based at the University of Michigan Medical School. Ontobee is a project to support the Linking Open Data (LOD) community to make RDF datasets available on the internet. The linked-data server mapped searched term to RDF through the web. The ontologies featured in Ontobee are mostly from OBO Foundry Library [3]. The browser responds to the search requests by SPARQL querying from the RDF triple store. The search results of the SPARQL queries are given in the JSON format or OWL format to produce machine-readable RDF files (Xiang et al., 2011).

4.3.2.2 Ontology Selection

The most important thing in selecting an ontology to be reused is understanding the requirements before making a decision to involve any particular ontology. Considering the integration step has started earlier at the conceptualization level, it simplifies the process of selecting available ontologies (Pinto, 1999b). The discovery process follows the guideline developed by Malone et al. (2016), to ensure that the ontology is about a specific domain

[1] http://bioportal.bioontology.org/
[2] http://www.ontobee.org/
[3] http://www.obofoundry.org/

of knowledge, or in this case an appropriate amount of knowledge to cover the modules. Assumption was made that BioPortal and Swoogle are contributed by domain experts from the Semantic Web community and are under active development. Although the guidelines are for selecting a Bio-ontology, however, its generality and the nature of healthcare domain makes the guidelines applicable to this research. (see Table 4.3)

Rule 1: The Ontology Should Be about a Specific Domain of Knowledge
Rule 2: The Ontology Should Reflect Current Understanding of Healthcare Systems
Rule 3: The Ontology Classes and Relationships Should Persist
Rule 4: Classes Should Contain Textual Definitions
Rule 5: Textual Definitions Should Be Written for Domain Experts
Rule 6: The Ontology Should Be Developed by the Community but Not Incapacitated by It
Rule 7: The Ontology Should Be under Active Development
Rule 8: Previous Versions Should Be Available
Rule 9: Open Data Requires Open Ontologies

Table 4.3: Rules for Selecting a Bio-Ontology (Malone et al., 2016)

Modules obtained from domain conceptualization : “healthcare”, “hospital”, “emergency department”, “process” and “patient data”, are used to perform ontology search using selected repositories, the search returned 31 results for BioPortal and 23 results for Ontobee (see Figure 4.6 and 4.7). From the total of 54 ontologies minus duplicate ontologies from the same keywords, only 7 ontologies fit the criteria as outlined by Malone et al. (2016). The list of 7 the selected ontologies is shown in Table 4.4

Ontology Name
1. (HEIO) Regional Healthcare System Interoperability and Information Exchange Measurement Ontology
2. (OMRSE) Ontology of Medically Related Social Entities
3. (GENEPIO) The Genomic Epidemiology Ontology
4. (OOSTT) Ontology of Organizational Structures of Trauma centres and Trauma systems
5. (TRIAGE) Nurse Triage
6. (TRANS) Nurse Transitional
7. (RNPRIO) Research Network and Patient Registry Inventory Ontology

Table 4.4: List of Selected Ontology

Ontology	Rule 1	Rule 2	Rule 3	Rule 4	Rule 5	Rule 6	Rule 7	Rule 8	Rule 9
"healthcare"									
(HCPS) Healthcare Common Procedure Coding System			√	√	√	√	√	√	√
SMASH Ontology	√	√	√	√	√	√		√	√
(CARELEX)Content Archive Resource Exchange Lexicon	√	√	√	√	√	√		√	√
(MHMO)Mental Health Management Ontology		√	√	√	√	√	√	√	√
(HEIO)Regional Healthcare System Interoperability and Information Exchange Measurement Ontology	√	√	√	√	√	√	√	√	√
(SITBAC)Situation-Based Access Control Ontology	√	√	√	√	√	√		√	√
(ITEMAS)Medical Technology Innovation in healthcare centers	√		√	√	√	√	√	√	√
"hospital"									
(OMRSE)Ontology of Medically Related Social Entities	√	√	√	√	√	√	√	√	√
(BHN)Biologie Hors Nomenclature	√		√	√	√	√		√	√
(ICECI)International Classification of External Causes of Injuries	√		√	√	√	√		√	√
(PCMO)Pediatric Consultation and Monitoring Ontology		√	√	√	√	√	√	√	√
"emergency department"									
(SSO)Syndromic Surveillance Ontology	√	√	√	√	√	√		√	√
"process" (Catgeory: Health; Human)									
(OMIT)Ontology for Genetic Susceptibility Factor	√		√	√	√	√	√	√	√
(OGSF)Ontology for MicroRNA Target	√		√	√	√	√	√	√	√
(SURGICAL) Nurse Surgical	√		√	√	√	√	√	√	√
(ADMIN) Nurse Administrator	√		√	√	√	√	√	√	√
(TRANS) Nurse Transitional	√	√	√	√	√	√	√	√	√
(TRIAGE) Nurse Triage	√	√	√	√	√	√	√	√	√
(CVDO) Cardiovascular Disease Ontology	√		√	√	√	√	√	√	√
(PSO) PatientSafetyOntology	√		√	√	√	√	√	√	√
(DCO) Dispedia Core Ontology	√		√	√	√	√		√	√
(RB) RegenBase ontology	√		√	√	√	√	√	√	√
(PTRANS) Pathogen Transmission Ontology	√		√	√	√	√	√	√	√
(PREMEDONTO) Precision Medicine Ontology	√		√	√	√	√	√	√	√
"patient data"									
(RNPRIO)Research Network and Patient Registry Inventory Ontology	√	√	√	√	√	√	√	√	√
(CPRO)Computer-Based Patient Record Ontology			√	√	√	√		√	√
(NDDO)Neurodegenerative Disease Data Ontology	√		√	√	√	√	√	√	√
(GENE-CDS)Genomic Clinical Decision Support Ontology	√		√	√	√	√		√	√
(TMO)Translational Medicine Ontology	√		√	√	√	√	√	√	√
(SOPHARM)Suggested Ontology for Pharmacogenomics	√		√	√	√	√	√	√	√
(ONSTR)Ontology for Newborn Screening Follow-up and Translational Research		√	√	√	√	√	√	√	√
(CTCAE)Common Terminology Criteria for Adverse Events		√	√	√	√	√	√	√	√

Figure 4.6: Result from BioPortal

4.3.3 Ontology Merging and Integration

The search for ontology reuse aims to provide a list of ontologies which are relevant to the domain requirements. The ontologies have provided the researcher with a representation, ontology definition, classes, data and object properties, as well as relations between the classes. Further on, the reuse process is completed with the merging and integration of the candidate ontologies into an application setting as stated in the conceptual model. Technically, it will produce a new final ontology built by importing candidate classes into a new ontology. The final ontology is called the Minimal Domain Ontology (MinDO). Top-down approach is used

Ontology	Rule 1	Rule 2	Rule 3	Rule 4	Rule 5	Rule 6	Rule 7	Rule 8	Rule 9
"healthcare"									
NCI Thesaurus OBO Edition			√	√	√	√	√	√	√
Ontology for MIRNA Target			√	√	√	√	√	√	√
Monarch Disease Ontology			√	√	√	√		√	√
Vaccine Ontology			√	√	√	√	√	√	√
"hospital"									
Ontology for Biomedical Investigations	√	√	√	√	√	√		√	√
EuPathDB ontology			√	√	√	√	√	√	√
Environment Ontology			√	√	√	√		√	√
Genomic Epidemiology Ontology	√	√	√	√	√	√	√	√	√
NCBI organismal classification			√	√	√	√		√	√
Infectious Disease Ontology			√	√	√	√	√	√	√
Influenza Ontology			√	√	√	√		√	√
Ontology of Organizational Structures of Trauma centers and Trauma systems	√	√	√	√	√	√	√	√	√
"emergency department"									
Ontology of Organizational Structures of Trauma centers and Trauma systems	√	√	√	√	√	√	√	√	√
NCI Thesaurus OBO Edition			√	√	√	√	√	√	√
(OMRSE)Ontology of Medically Related Social Entities	√	√	√	√	√	√	√	√	√
"process" (Category: shift)									
Ontology for MIRNA Target			√	√	√	√	√	√	√
(SBO) Systems Biology Ontology	√		√	√	√	√	√	√	√
(PDRO) The Prescription of Drugs Ontology	√		√	√	√	√		√	√
(OBCS) Ontology of Biological and Clinical Statistics	√		√	√	√	√		√	√
(OAE) Ontology of Adverse Events	√		√	√	√	√	√	√	√
NCI Thesaurus OBO Edition			√	√	√	√	√	√	√
Ontology of Organizational Structures of Trauma centers and Trauma systems	√	√	√	√	√	√	√	√	√
"patient data"									
NCI Thesaurus OBO Edition			√	√	√	√	√	√	√

Figure 4.7: Result from Ontobee

to create classes, by introducing the higher-level class first, then the medium-level class and the lower-level class. Information extraction is conducted using the Protégé 5.1.0 [4], in the following steps:

- Input: Selected ontologies
- Steps:

  1.1 Load URI or Ontology file (owl or RDF/XML) into Protégé.

  1.2 Analyse each class, object, property and instance of the class

  1.3 Define the relationship between the class and modules described in the conceptual model. 1.2 Get all objects, instances, classes and properties (instance if any) of the class 1.3 Load ontology objects, instances, classes and properties
- Output: Ontology classes, instances, properties and relations

The extracted classes for a manual ontology development is shown in Table 4.5. The manual development of the new ontology is performed by importing selected ontologies into Protégé. The imported ontologies displayed the direct and indirect imports for the active (final) ontology. The ontologies are loaded into Protégé by loading its RDF/XML file. The location or hyperlink where the ontology is loaded from is shown. Later the selected classes are in-

[4]https://protege.stanford.edu/

tegrated and rearranged accordingly to fit the domain and application specification requirements. Class re-definition started with the general class of the domain, 'Hospital' followed by 'Process' class and 'Role' class. After the classes were aligned, the relationships between the classes were added.

The steps performed to obtain the final ontology are as follows:

- Input: Selected ontology
- Steps:
  - 1.1 Import ontology elements from the selected ontology (classes or properties)
  - 1.2 Create a new class('Hospital' > 'Process' > 'Role')
  - 1.3 Re-arrange the class and remove class redundancy according to the process flow.
  - 1.4 Add a new property (object property, data type property) and annotation.
  - 1.5 Save as a new ontology.
- Output: Final ontology (MinDO)

Figure 4.8 shows the classes in the final ontology, MinDO, with *Hospital*, *Process* and *Role* classes. See Appendix A for the full version of MinDO.

Ontology	Class	Modules Representation
HEIO	'Electronic Health Information Type'	Hospital
OMRSE	'health care facility'	Healthcare; Hospital; Emergency Department
OOSTT	process	Process; Emergency Department
Nurse Triage	Nurse; Patient; Triage_Nurse	Process; Emergency Department
Nurse Transitional	Bed_Occupancy_Indicator; Nurses_Dependent_Role; Nurse_Independent_Role; Nurse_interdependent_Role; Referrals	Process; Emergency Department
RNPRIO	Patient_Population	Patient Data
GENEPIO	Organism datum; 'personal health datum'	Patient Data

Table 4.5: List of Extracted Classes

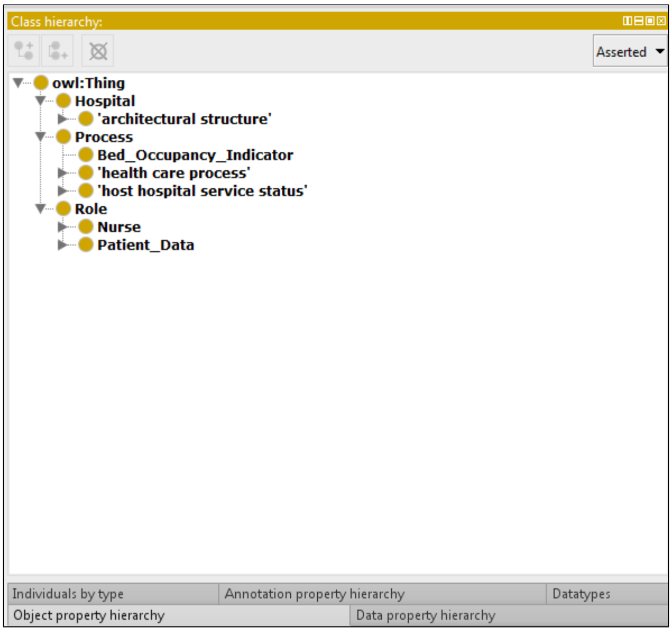


Figure 4.8: MinDO Classes

4.3.4 Ontology Evaluation

Instantiation can be viewed as an existing implementation of a framework, and is used to evaluate constructs, models and methods (March and Smith, 1995). To meet the objective of this iteration, the framework was developed and implemented following the steps to search for ontology for reuse. The evaluation of this process was made by assessing the usefulness of the framework in conceptualizing the domain, discovering the ontologies, selecting the ontology, as well as merging and integrating the ontologies, leading to the development of the final ontology. For this research, the aim is to evaluate the structural, functional adequacy, compatibility, operability and quality of the final ontology. The process measures the ontology for its consistency, tangledness, degree of accomplishment of its functional requirement, its capability to perform under stated condition, and the efforts needed for using the ontology. In order to evaluate the quality of the MinDO, a SPARQL query was formulated for subclass count, class hierarchy, direct subclass and superclass, the length of path between nodes, and the class annotation. Example in Table 4.6 shows a snippet of the SPARQL used to calculate the path between classes.

```
PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX owl: <http://www.w3.org/2002/07/owl#>
PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
PREFIX xsd: <http://www.w3.org/2001/XMLSchema#>
select ?super ?sub (count(?mid) as ?distance) {
?super rdfs:subClassOf* ?mid .
?mid rdfs:subClassOf+ ?sub .
}
group by ?super ?sub
order by ?super ?sub
```

Table 4.6: Snippet the SPARQL Query: To calculate the path between classes

Based of the deduced criteria in the domain requirement, the ontologies selected should accomplish the task as modules in a simulation modeling environment, the completeness of the final ontology is determined by having the classes represented by the reused ontologies. OQuaRe evaluation metrics are used: Tangledness (TMOnto); Annotation Richness (ANOnto); Relationships per class (INROnto); Number of Children (NOCOnto); Number of properties (NOMOnto); Lack of Cohesion in Methods (LCOMOnto). The metrics were used in association with OQuaRe evaluation subcharacteristics. After the characteristics were evaluated based on the metrics, the values are ranged from 1 to 5; where "*1 means not acceptable, 3 is minimally acceptable, and 5*" "*is exceeds the requirements*"(ISO25000, 2005). The evaluation metrics used are as follows:

- Tangledness (TMOnto) $TMOnto=\Sigma|R_{Ci}| / \Sigma|C_i|-\Sigma|C(DP)_i|$;
- Annotation Richness (ANOnto) $ANOnto=\Sigma|A_{Ci}| / \Sigma|C_i|$
- Relationships per class (INROnto) $INROnto=\Sigma|R_{Ci}| / \Sigma|C_i$
- Number of Children (NOCOnto): $NOCOnto=\Sigma|R_{Ci}| / (\Sigma|Ci| -|R_{Thing}|)$
- Number of properties (NOMOnto) $NOMOnto=\Sigma|P_{Ci}| / \Sigma|C_i$
- Lack of Cohesion in Methods (LCOMOnto) $NOMOnto=\Sigma|Length(path(|C(leaf)_i|))/m$

Ontology evaluation is performed to assess the functional adequacy and quality of the ontology for reuse that has fulfilled the structural, functional adequacy, compatibility and adequacy characteristics. The characteristics are based on the OQuare Quality Model (Duque-Ramos et al., 2011). The model adapted the SQuaRE characteristics to be applied into the ontology quality. The evaluation metrics used in this iteration are intended to satisfy the OQuare Quality Model characteristics. Based on a complete description of the OQuaRE Quality Model [5], The *Structural* is the characteristic for formal and semantic ontological properties. The capability of the ontologies to provide concrete functions, the appropriateness for its intended purpose (Stevens and Lord, 2009) is measured by its *Functional Adequacy*. The

[5] http://miuras.inf.um.es/evaluation/oquare/Contenido.html

Compatibility and *Operability* characteristics are to evaluate the ability of two or more ontologies and the effort needed for using an ontology respectively.

Table 4.7 shows the OQuaRe scores for the characteristics and metrics used in evaluating the quality of MinDO. The evaluation of a particular characteristic and the association metrics is important in understanding how the quality of MinDO is being measured. OQuaRe adopted the Compendium of Software Quality Standards and Metrics (Lincke and Löwe, 2007) for the association between metrics and the scores for each characteristic. The scores for each metrics were transformed into a 1-5 range before the score for each characteristic was obtained. The weighted average of the scores of all the quality characteristics was calculated for the overall quality of MinDo following the formula:

- $x = (w_1x_1 + w_2x_2 + \ldots + w_nx_n) / (w_1 + w_i + \ldots + w_n)$ Where, w_i are weights and x_i are values

The weighted average for the OQuaRe Score (Refer Table 4.7) is 4.25. The OQuaRe quality model does not define the quality level of the evaluated ontology, but describes the main feature of the ontology (Duque-Ramos et al., 2011).

Characteristics	Metrics	Values	Score
Structural	Tangledness (TMOnto)	7.167	2
	Annotation Richness (ANOnto)	159%	5
Functional	Annotation Richness (ANOnto)	159%	5
	Relationships per class (INROnto)	0.95%	5
Compatibility	Number of Children (NOCOnto)	1	5
	Number of properties (NOMOnto)	0.413	5
Operability	Number of Children (NOCOnto)	1	5
	Lack of Cohesion in Methods (LCOMOnto)	1.465	2

Table 4.7: OQuaRe Score

4.4 Framework Evaluation

This section concludes ORF by discussing the framework evaluation following a study by Venable et al. (2016).

4.5 Learning Outcome

The primary points of learning are:

- Modules integration may identifies ontology integration possibility (Pinto and Martins, 2001). This agrees to the hypothesis that domain fragments add domain knowledge into the final ontology to provide a structure to the final ontology (SIMPERL, 2010). The identification of modules as knowledge representative are relevant for this study, as no prior knowledge is required about the ontologies but the identification allows the organization of MVSimO into thematic-based concepts that will be understood as relevant domain knowledge.
- Usually, a decision tree is built after the application of specific objectives specified by a modeler or an ontology engineer to determine the characteristics required in the application setting (Sure et al., 2002; Benjamin and Graul, 2006). Rather than having to encode information multiple times to make decisions in different application settings, the idea of *modules* is to have revisable domain elements, which can be put into several uses such as ontology integration, evaluation and maintenance.
- It is clearly evident that there are duplications in ontology and classes of ontology selected, which can be dealt at different stages of the ontology reuse process. The **first** option would be to remove the redundant ontology with the same search keyword. Involving the manual selection at an earlier stage to resolve redundancies before the ontology was actually built can result in more accuracy. **Second** would be to remove redundant and unaffiliated classes during the merging and integration process after list of the classes, instances, properties and relations are derived. At this stage the re-

moving activity is carried out in a stricter manner before they are added to the final ontology, to eliminate most of the duplicity created in the new relations and concepts. This could potentially result in achieving a higher precision ontology.

- As illustrated in the sample taken from the merged final ontology, MinDO in Figure 4.8 *Hospital*, *Process* and *Roles* are the super-classes from the selected ontologies, the sub-classes that fall under these classes might be unrelated. At this stage, manual prunning can refine the final ontology by eliminating the duplication or undesirable classes and links concepts in different ways to provide more modeling possibilities. An automatic refinement based on the exact string matching can also be done to perform deeper ontology merging to accurately check for the existing concepts before they are added into the ontology. Hence, integrating the automatic refinement with ontology matching techniques before the ontology merging is a desirable improvement that should lead to the elimination of the majority of the redundant concepts. This should also reduce the efforts of the ontology engineer in the pruning step.

- The merging and integration of the final ontology is an iterative process. Ontology elements such as classes, properties and instances can be later refined once new ontology is selected during the ontology selection process by the domain expert. The refinement process in every part of the ontology reuse framework allows customary requirements and feedback for the reuse process to evolve. This enables an efficient knowledge representation scheme to allow for the designer to make better-informed decisions (Chandrasegaran et al., 2013).

- The evaluation of the final ontology requires manual identification of desirable and undesirable concepts beforehand. This is performed during the refinement in the ontology merging and integration phase. The total number of classes, instances, properties and relations are determined by the manual identification. The number changes accordingly, therefore, the evaluation scores might vary. Eventually, the OQuaRe quality model that is used to evaluate the quality of MinDO does not discuss the 'level' of the quality but rather, focuses on defining the characteristics of the ontology

4.6 Summary

This chapter validates the theory of this research, that reusing structured domain knowledge as modules to represent sub-ontologies may lead to the building of a new ontology. This iteration contributes to an improved ontology reuse framework. A formal definition of the output of the phases consisting of the framework is provided. Another main contribution of this chapter is the thorough ontology discovery and the selection processes to prove the applicability of the candidate ontologies, despite the inexistence of experts' opinion. The evaluation has demonstrated that there is enough domain knowledge in MinDO. The approach adopted in ORF has proven to be efficient in reusing domain ontologies and building new ontology, according to domain and application specification requirements. Overall, the method has proven to be efficient by introducing the new technique of module integration that not only guides the domain study but explicitly justifies the choice of underlying modeling decisions. The learning that emerged from this iteration highlights a number of issues and challenges that can be developed to direct future research.

Chapter 5

Iteration 2

5.1 Introduction

Classes and properties are basic elements provided through ontology languages irrespective of the underlying formalities. Data from the real system complements the semantic meaning in ontology by using its relation to represent ontology properties. In simulation modeling of A&E, the process-element related semantics from the data can disentangle the scope for interpretation connected with the use of natural language and improve simulation ability. In studies by Noy et al. (2001); Labrousse and Bernard (2008) as discussed in Chapter 2, the structural and functional decisions to model an ontology-based simulation are derived from the understanding of real-world system processes. Therefore this iteration has conducted process mining and formal concept analysis to extract the formalism of A&E processes. The aim of this iteration is to syntactically and semantically discover the attribute relation from semi-structured data without the involvement of experts' judgement (Benjamin et al., 2006; Mohiuddin et al., 2017a). This discovery process is an attempt to prevail in the definition of Minimal Viable Simulation Ontology (MVSimO) as a specification of simulation modeling conceptualization by continuing the effort done in Iteration 1 using Formal Concept Analysis (FCA).

The processes started with the introduction of the FCA as a mathematical method involving semi-structured qualitative data; followed by the formal context acquisition in a form of cross table from the domain requirement understanding; concept lattice and implication basis generation; as well as attribute exploration to select the implication basis convenient to the context. Here attributes relations are identified based on the output of the FCA and these representations are then utilized for the semantic enrichment of MVSimO. The attributes relations are instantiated using FCA-Relation Discovery Methodology (FDiMe) that is used to develop MVSimO for simulation modeling in the next iteration.

This chapter is structured as follows: Section 5.2 provides the research design and the research outputs for this iteration. Section 5.3 describes the incorporated elements that contribute to the concept analysis in the perspective of; Discrete-event Simulation (DES) in the A&E department. Section 5.4 presents the building and development of the design artefacts, FDiMe - illustrating and detailing the notion of concept exploring theory; including a formal context formation in a representation of a certain, well-defined sense and rigorous knowledge extraction. Section 5.5 discusses the attribute exploration of implication basis to represent the domain. Section 5.6 describes the validation of the research outputs. The learning outcome of this iteration is presented in section 5.7, and finally, the chapter is summarized in section 5.8.

5.2 Design Science Research and Output Artefacts

As discussed in Chapter 2, a collaboration technique of process mining and ontology development works to have a deep understanding of the domain (Verdonck et al., 2018; Wray and Eklund, 2014). In sharing a common understanding, an ontology architecture provides a connection between multiple domain data resources (Kettouch et al., 2015). The purpose of this Design Science Research iteration is to build an FCA-relation discovery methodology by incorporating the mathematical analysis funded by FCA to a set of healthcare data which later complements the ontology acquired from the previous iteration to build the MVSimO.

The methodology necessitates finding objects and attribute relations in the A&E data. This iteration extends the framework in Iteration 1 to include techniques for concept exploration by applying the FCA theory of lattices and implication basis. As mentioned in Chapter 2, FCA is used to organize knowledge and enriching it using Web Ontology Language (OWL) to transform information into reusable knowledge and also to identify conceptual structures among the dataset. This method allows for the discovery of fundamental concepts and relations in real dataset, as well as support ontology development in a way suitable for knowledge exchange, information retrieval, and in the context of this research - ontology for simulation modeling. This chapter proposes a methodology for objects-attributes relations task and semantically link the relations to ontology knowledge gained in the previous iteration in the quest for answering the question of - "How to generate detailed healthcare processes from pathways and data?". Figure 5.1 shows the research iteration for Chapter 5.

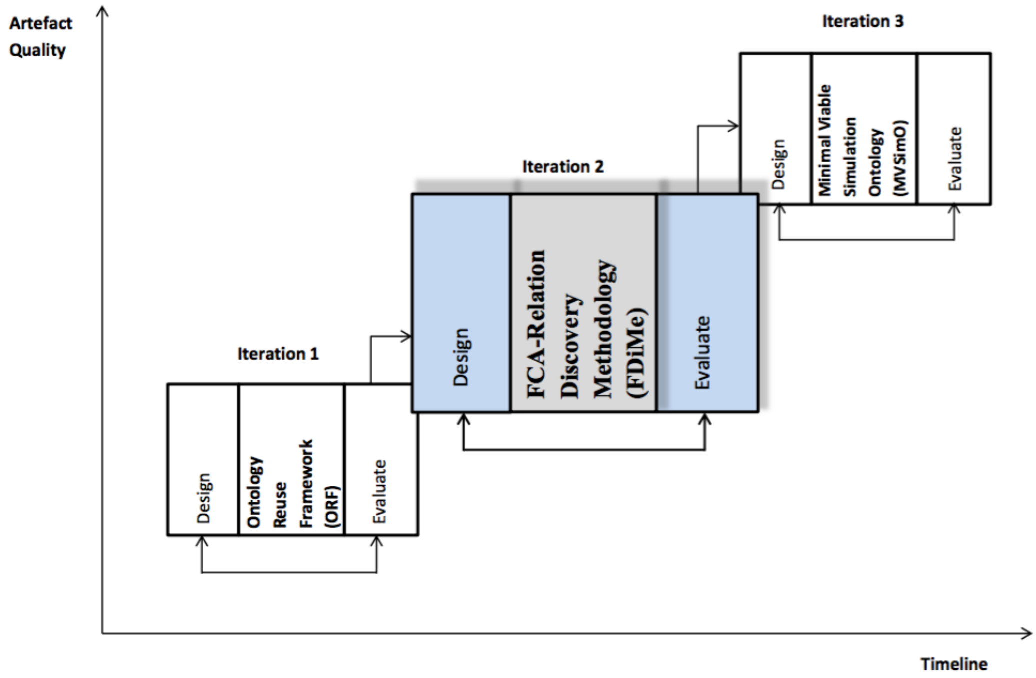


Figure 5.1: Research Iteration for Chapter 5

5.2.1 Design Research Artefacts

This chapter introduces a concept exploration technique to acknowledge object-attribute relations in semi-structured data. The approach executes the segregation of knowledge mapping into blocks. The first step started with the pre-processing of the A&E data from Hillingdon Hospital. During this step, a modeler should know clearly the real-world problem that needs to be modeled or analysed. The mapping of objects and attributes onto the cross-table, which is also known as the formal context, helps to construct the data in the way intended by the modeler, so that it conforms to the simulation requirement. Secondly, formal concepts are derived from the formal context using the FCA software. Finally, the third phase involves the disaggregation of the knowledge captured from the analysis, and correlations between the process, as well as the analysis result are shown. To achieve the objective, this iteration executes the steps in Table 5.1.

Steps	Method	Input Artefact	Output Artefact
1. Data preprocessing	Data generalization, Plain Scaling	A&E Data	Refined Data
2.Formal context mapping	Pathways mapping, Line of Best Fit Analysis	Refined data A&E pathways	Cross-table STP Map
3.Formal concept exploration	Concept Analysis using ConExp and FCART	Formal context STP Map	Formal concepts Implication basis Concept Lattice
4.FCA-Relation transformation	Process relation extraction	Implication basis	Process relation

Table 5.1: Iteration 2 - Input Output Model

5.3 Artefact Building and Development

The proposed FCA-Relation Discovery Methodology (FDiMe) is developed in a four-phase approach. In Figure 5.2 depicts the complete process of the methodology development. The contextual or logical structure of the data is generated because the data is required to be in a standard form, known as the formal context. (Ganter and Obiedkov, 2016). Each step on FDiMe is further described in the following subsections which starts with data pre-processing; formal context formulation process and cross-table development; formal concept generation using the FCA software and; the FCA-relation transformation of the set of implication basis. The application of the methodological framework using the real A&E data collected from April 2014 until March 2015 to replicate the domain is detailed and demonstrated in the following subsection.

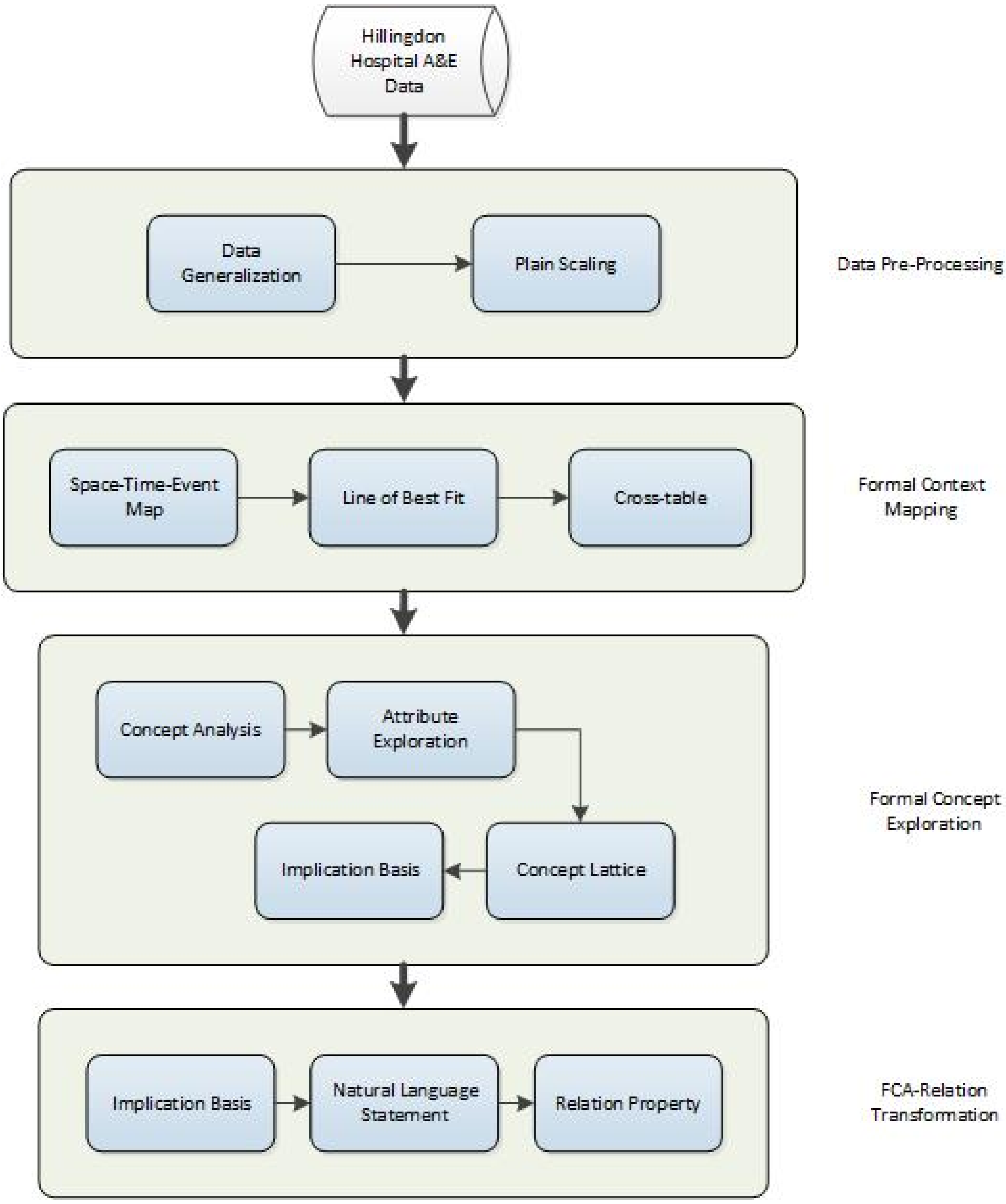


Figure 5.2: FCA-Relation Discovery Methodology (FDiMe)

5.3.1 Data Pre-Processing Phase

Before the start of FDiMe, the A&E data are pre-processed by first, normalizing the column name [1] [2] of the dataset into a more understandable name (See Appendix B for the list of column name, Normalized Data Element Name (DEN) and its description), and second, by removing the null and redundant records. The many-valued columns are then transferred into one-valued columns using the plain scaling method. The plain scaling method substitutes attributes in the original many-valued context with a set of columns representing each one of the allowed values for the attributes (Jiang and Chute, 2009). The process converts 25 many-valued columns to 98 single-valued columns. The application of the plain-scaling method enables multiple attributes to be grouped accordingly. For example, *EMAttendance-Date* is grouped into weekdays and sub-grouped into two 12-hour periods. The first 12-hour period is designated as a.m. It runs from midnight to noon. The second period, marked p.m, covers the 12 hours from noon to midnight. Table 5.2 shows examples of many-valued to single-valued transformation.

5.3.2 Formal Context Phase

The formal context of the data is illustrated using a rectangular table, a cross-table with one row for each object (represents Patient ID) and one column for each attribute, having a cross in the intersection of row x with column y if and only if object x has attribute y. The formal context phase adopted here is a process of describing data into a more understandable context according to this research. During this phase, the *Line of Best Fit* analysis is conducted to ensure that useful and manageable (Andrews and Orphanides, 2010) lattices are generated and to avoid any exceptions during the analysis by demonstrating a correlation between the number of implication and the number of objects using samples of 50, 100, 200, 300, 400 and 500 objects.

[1] http://www.datadictionary.nhs.uk
[2] https://www.nhscc.org

Process Element Extraction

The data provided for this research came in the form of comma-separated value (CSV), a and manual semantic selection is conducted to improve the representation of the data for simulation modeling purpose. For this research, the aim is to model patient flow in the A&E thus, discrete-event simulation is selected to show a discrete sequence of events in the department. The Space-Time-Process (STP) mapping process helps to derive all of the possible process element that will lead to MVSimO and simulation modeling. The process flow from the A&E pathways is used in this phase as the guideline. To obtain the STP map, the process element based on space, time and event has adopted the dataset analysis as proposed by Sider et al. (2001) and De Cesare et al. (2014), for the ontology development method.

5.3.3 Formal Concept Phase

The FCA-relation discovery technique in this methodology describes the process of generating the concept and its implication basis; and incorporating the elements that contribute to the discrete-event simulation modeling to produce MVSimO. The technique requires an understanding of the domain and the research context, as described in the following subsections.

5.3.4 *FCA-Relation Transformation Phase*

From the set of implications obtained in the earlier phase, relevant implications between the attributes of the process elements are selected and transformed into a natural language statement and used to represent the class properties in MVSimO. This phase is the assignment of the concept formalized as process relations to model ontology class properties with a logical reasoning mechanism (Xiao Hang Wang et al., 2004). Thus the semantic of the process element formulated in the natural language can be represented formally and in a machine-processable way through MVSimO. The transformation relation process is conducted using the first-order predicate - a subject, a verb and an object. For example, from the implication-transformed natural language: "*patient attended A&E by ambulance; on Friday (PM), assessed into A&E departments; admitted as lodged patient*" can be described as "*Patient, attended, by ambulance*; "*Patient, assessed, Friday (PM)*"; and "*Patient, admitted, lodged patient*". The chosen implications as relations (and their underlying descriptions) are described in more detail in Section 5.5.1

5.3.5 Methodology Validation

A domain expert is typically used to validate and modify the resulting relation property and filter out any irrelevant relations or concepts. An implementation of the next stage of MVSimO, is described in the next iteration. Further validation of MVSimO is made in the simulation modeling for A&E, which provided the required development environment for implementing the ontology. The user is then able to view the generated ontology, model a simulation and make any further changes or amendment to the rules or ontology to fit the simulation purpose.

Many-Valued	Single-Valued
EMAttendanceDate	Mon: am
	Mon: pm
	Tue: am
	Tue: pm
	Wed: am
	Wed: pm
	Thu: am
	Thu: pm
	Fri: am
	Fri: pm
	Sat: am
	Sat: pm
	Sun: am
	Sun: pm

Many-Valued	Single-Valued
Age	Age: 0-4
	Age: 5-18
	Age: 19-64
	Age: 65 and above

Many-Valued	Single-Valued
EMReferralSource	Ref: Educational establishment
	Ref: Emergency services
	Ref: General Medical Practitioner
	Ref: Healthcare provider
	Ref: Other
	Ref: Police
	Ref: Self referral

Table 5.2: Example of Multi-Valued to Single-Valued Attribute

Original Column Name	Revised Column Name
zNHSNumberPseudo	PatientID
ProviderCode	HospitalCode
ProviderName	HospitalName
zArrivalDateTime	ArrivalDate
zEMAttendanceConclusionDateTime	EMAttendanceDate
zEMDateTimeSeenforTreatment	EMDateSeen
Arrivalmonth	ArrivalMonth
YearMonth	YearMonth
FiscalYearLabel_1	FiscalYear
CoreHRG	CoreHRG
hrg_desc	HRGDescription
EMModeofArrival	EMModeofArrival
EMModeofArrivaldesc	EMModeofArrivalDescription
EMReferralSource	EMReferralSource
EMReferralSourcedesc	EMReferralSourceDescription
A&E_DepartmentDescription	DepartmentDescription
GPPracticeCode	GPPracticeCode
GPPracticeName	GPPracticeName
Age	Age
CCG	CCG
EMDiagnosisFirst	EMDiagnosisFirst
diag_desc	DiagnosisDescription
EMAttendanceDisposal	EMAttendanceDisposal
Attendance_disposal	AttendanceDisposal
PostcodeSectorofUsualAddress	PostcodeSectorofUsualAddress

Table 5.3: Revised Columns Names

5.4 Implementation of FCA-Relation Discovery Methodology

The initial data has 126,986 rows with 25 columns. For categorisation and generalisation of the attributes, the column names are revised to a more meaningful name suitable to this research (Refer Table 5.3), and is transformed into 98 single-valued columns. Columns with a null record are removed, and rows are randomly generated for the formal context in the form of cross-table . The basic structures of FCA are context and concept hierarchy. A context is comprised of a set of Patient ID as objects, a set of attributes, and a binary relation describing which objects possesses which attributes. As introduced in Chapter 2, the followings are the formal definitions of FCA:

- **Definition 2.1** Formal context

  A formal context is a triplet *(X, Y, I)* where *X* is a set of objects and *Y* is a set of attributes and *I* is a binary relation between *X* and *Y*, i.e., $I \subseteq X \times Y$. $(x,y) \in I$ indicates that object *x* has attribute *y*.

- **Definition 2.2** Intent and Extent

  Let (X, Y, I) be a context, X' $\subseteq X$ and Y' $\subseteq Y$, the function Intent maps a set of objects to the set of attributes, whereas the function Extent maps a set of attributes to the set of objects:

  Intent (X') = y $\in Y' \mid \forall y \in Y'$, (x,y) $\in R$

  Extent (Y') = x $\in X' \mid \forall x \in X', (x, y) \in R$

  For X' $\subseteq X$, Intent (X') is the set of attributes owned by all objects of X', and Extent(Y') is the set of all objects that own attributes Y'. The two functions form a Galois connection and formal concepts.

- **Definition 2.3** Formal Concept

  A Formal Concept C in a context is a pair (X', Y') that satisfies Y' = Intent (X') and X' = Extent(Y')

  i.e., C is a Formal Concept ⇔ for X' $\in C and Y' \in C$, Extent(Intent(X')) = X',

  and symmetrically, Intent(Extent(Y')) = Y'.

- **Definition 2.4** Implications

An implication A ⇒ B holds in (X,Y,I) if and only if B ⊆ A”, which is equivalent to A’ ⊆ B’. It then automatically holds in the set of all concept intents

5.4.1 Line of Best Fit Analysis

From 126,986 rows of data with 98 single-valued columns - 50, 100, 200, 300, 400 and 500 objects are selected to produce six cross-tables. Each cross-table represents the Patient ID in rows and attributes in columns. The number of implications derived from the six cross-tables are then compared to identify the line of best fit to show the correlations between the number of rows and the number of implications, as shown in table 5.4. Using FCART software, the basis of implications are calculated in 3 cycles - 2,000; 3,000; and 4,000 iterations. At this stage, no concept or implication is selected because the purpose of the analysis is to determine the best approximation number of rows to be selected to represents the data set, and to determine the correlation between the number of rows and the number of implications. The line of best fit is produced by the following steps:

- **Step1** Calculate the mean of the x-values and the mean of the y-values
  $\overline{X} = \dfrac{\sum_{i=1}^{n} x_i}{n}$;
  $\overline{Y} = \dfrac{\sum_{i=1}^{n} y_i}{n}$;
- **Step 2** The following formula gives the slope of the line of best fit
  m = $\dfrac{\sum_{i=1}^{n}(x_i - X)(y_i - Y)}{\sum_{i=1}^{n}(x_i - \overline{X})^2}$
- **Step 3** Compute the y-intercept of the line by using the formula
  b = $\overline{Y} - m\overline{X}$
- **Step 4** Use the slope m and the y-intercept b to form the equation of the line
  From the steps, table of 5.4 is produced with;
  Slope = 0.005390795 and Intercept =15.50737808

From the graph generated, as shown in 5.3, the number of rows affect the number of implications; more rows generate more implications. Considering the data is quantitative data, by margin of error at a 95% confidence level, 500 objects are selected as samples to best represent the data and the result is used throughout this research. [3].

No. of Rows	50	100	200	300	400	500
Average Lines	16.167	15.667	16.633	16.8	17.933	18.2

Table 5.4: The Average Implications per Cycle

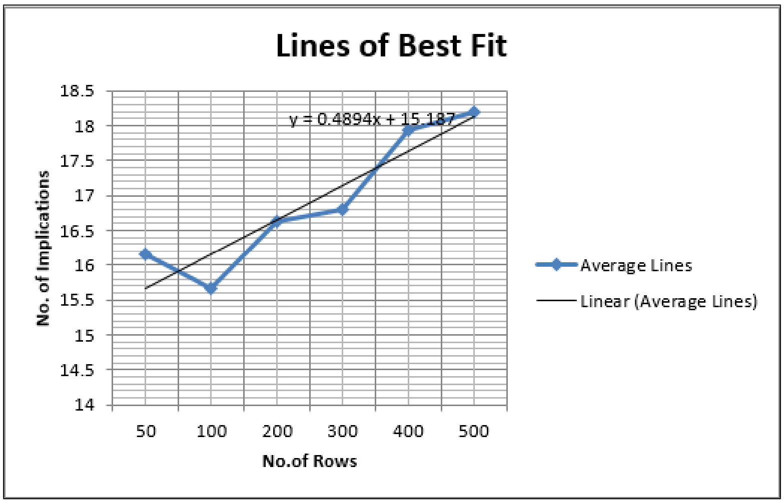


Figure 5.3: Line of Best Fit

5.4.2 Space-Time-Process Map

The graphical representation of the process elements is mapped to show the process flow in A&E department based on the generic pathways. The Space-Time-Process (STP) graphical representation with the textual description of the process activities and the roles involved in each activity is depicted in Figure 5.4. From the map, the processes are separated into blocks of activity and entity. The blocks are lined into sequence order to resemble the flow in A&E. STP map provides the first insight on how to model a simulation from a modeler's perspective. From the pathways, the processes are categorised into 4 main processes: Check-In, Assessment; Treatment; and Outcome. Each process block is labelled as process, activity and

[3]https://help.surveymonkey.com/articles/en_US/kb/How-many-respondents-do-I-need

entity. The process-event block illustrates each event that will be modeled for the discrete-event simulation.

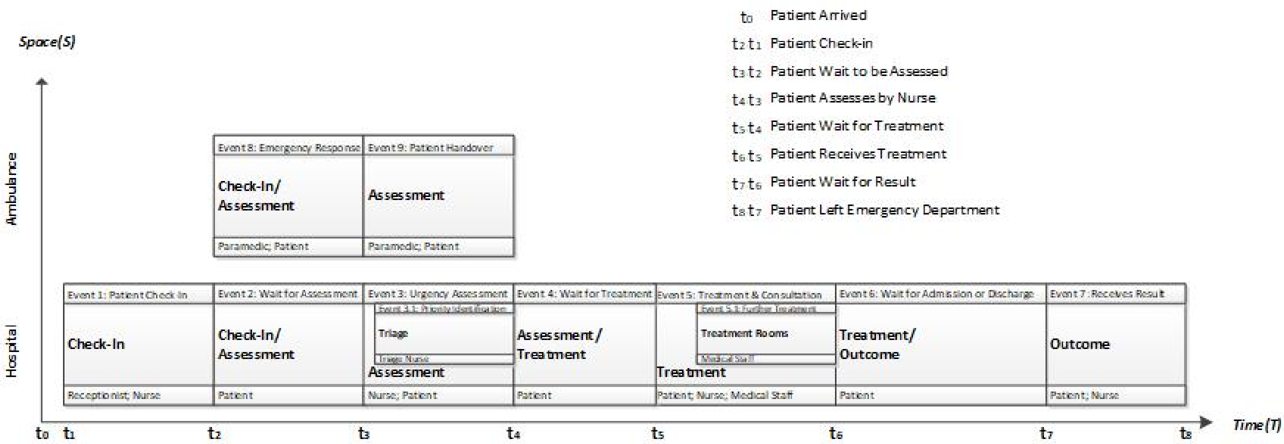


Figure 5.4: A Space-Time-Process of 'Patient Flow in A&E Departments'

The process begins with a patient check-in, either by walking to the hospital or by using an ambulance service, patient then waits to be assessed by a nurse before proceeding to the triage assessment, receives treatment by a nurse or a medical staff, and finally waits for the outcome of whether to be discharged from the hospital or admitted to the ward. The processes take place in their designated space or location in the department within a specific time. During the transition from one process to another, the patient has to wait in a queue. This suggests a delay in the process, for example between t_2 and t_3 of the process *Check-In/Assessment*, the event of *Wait for Assessment* with the entity of *Patient* occurs. In another example between t_3 and t_4, the entity of *Patient* and *Nurse* are involved in the process of *Assessment* for *Urgency Assessment* event. During this event, there is also a sub-event which occurs in a sub-location in the department. The *Triage* process is conducted by a Triage Nurse to run *Priority Identification*.

From a 3D perspective, STP determines which terms such as space, time and event, in deciding the way the process element can be extracted and modelled. Starting at this point, all steps taken are based on STP because it maps the outline process of A&E; determines object's roles and boundary, as well as the level of details; and models events for the discrete-event simulation. In MVSimO, there are various kinds of objects, properties and relations between kinds and their instances (Benjamin et al., 2006). STP acts as a bridge for the modeler and the user, and translating the real-world knowledge into a process definition. The rationale behind STP in this iteration is to identify the process elements which can then be applied in the conceptual exploration to extract suitable concepts, and their taxonomic and non-taxonomic relation.

5.5 Conceptual Exploration

In this research, FCA is used to determine all the concepts of a given context, and to describe the dependencies between the attributes. Focus is given to two algorithms; first, to produce a relevant concept and second, to construct the minimal set of implications. The conceptual exploration process emphasizes on the process in the A&E thus the minimum set of attributes are determined to define the process concepts and its hierarchy in the formal context (Yan et al., 2015). For the definition of formal concept in FCA, refer to subsection 5.4. Initially, the conceptual exploration using FCART, with 500 objects and 98 attributes has generated 6920 concepts (Refer figure Figure 5.6). FCART returns all 6920 concepts describing pairs that satisfy all objects (Patient ID) sharing all attributes (Intent) and all attributes shared by all objects (Extent). The concepts generated at this stage included all the attributes without focusing on attributes with process elements only. This is performed earlier to determine how the number of objects affect the number of implications, therefore 500 objects are selected for concept extraction.

The formal concept extraction process comprised of four steps that are applied to object and attributes from the cross table to describe the processes in A&E department. **Firstly**, a set of process elements from STP - Check-In; Assessment; Treatment; and Outcome are compared semantically to find any associated elements from the data. The elements can be process name, process entity, process date and time, and process location according to STP. Here the attribute reduction is performed to reduce the number of attributes when the number of

Patient ID remains. The attributes selected are shown in Figure 5.5. This is to obtain concepts with process-related attributes. Concept reduction by reducing the attributes makes implicit knowledge discovery easier, and also makes the representations of implication basis more process-oriented. The reduced concepts focused on process element by including the attributes of:

Age (Age)

EMAttendanceDate (Attendance Date/Time)

EMModeofArrivalDescription(Mode of Arrival)

AttendanceDisposal (Outcome - Admitted or Discharged)

DepartmentDescription (Department)

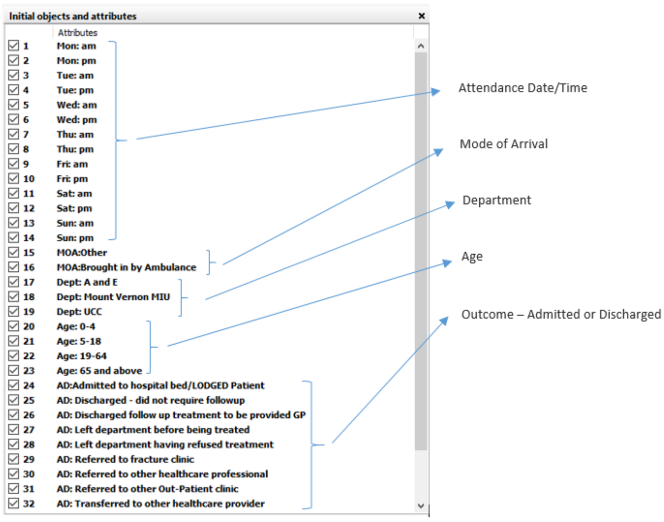


Figure 5.5: The Selected Attributes

Secondly, after selecting the process element by its column, the initial formal context is reduced giving a more accurate concept of the domain. Figure 5.7 shows the formal context of 500 objects and 32 attributes after the concept reduction generated using FCART. The concepts have been reduced to 1019 concepts. As a result, five concepts with two or more objects are selected from the reduced concepts for concept lattice development. Table 5.5 shows the list of selected formal concepts in natural language statement. After the initial concepts are reduced to get the concepts that contain process elements, a new formal context and concept lattice is generated using ConExp.

Formal concept, expressed in natural language	**No.of Object**
Fri: pm, MOA:Brought in by Ambulance, Dept: A and E, Age: 65 and above, AD:Admitted to hospital bed/LODGED Patient	3
Thu: pm, MOA:Other, Dept: Mount Vernon MIU, Age: 65 and above, AD: Discharged - did not require followup	2
Sun: pm, MOA:Other, Dept: UCC, Age: 0-4, AD: Discharged - did not require followup	2
Mon: am, MOA: Brought in by Ambulance, Dept: A and E, Age: 19-64, AD: Admitted to hospital/LODGED Patient	2
Wed: pm, MOA: Other, Dept: mount Vernon MIU, Age: 5-18, AD: Discharged - did not require followup	2

Table 5.5: Selected Formal Concepts

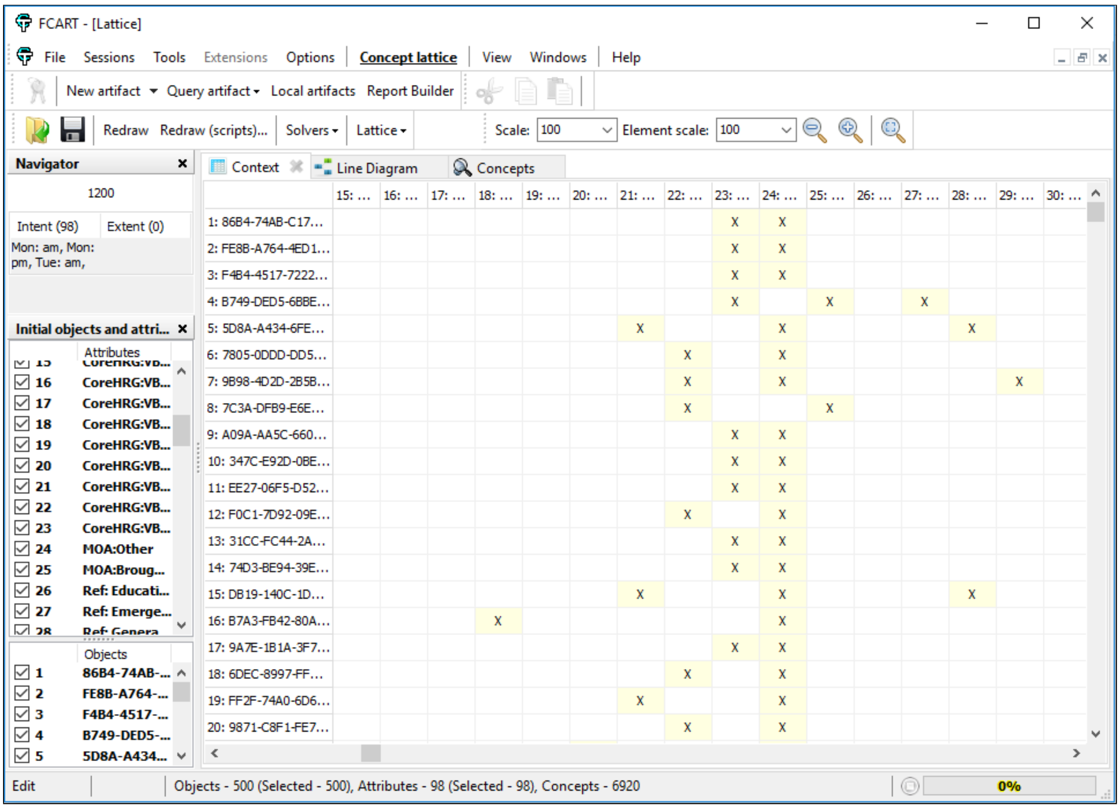


Figure 5.6: The Formal Context with 500 rows and 98 attributes

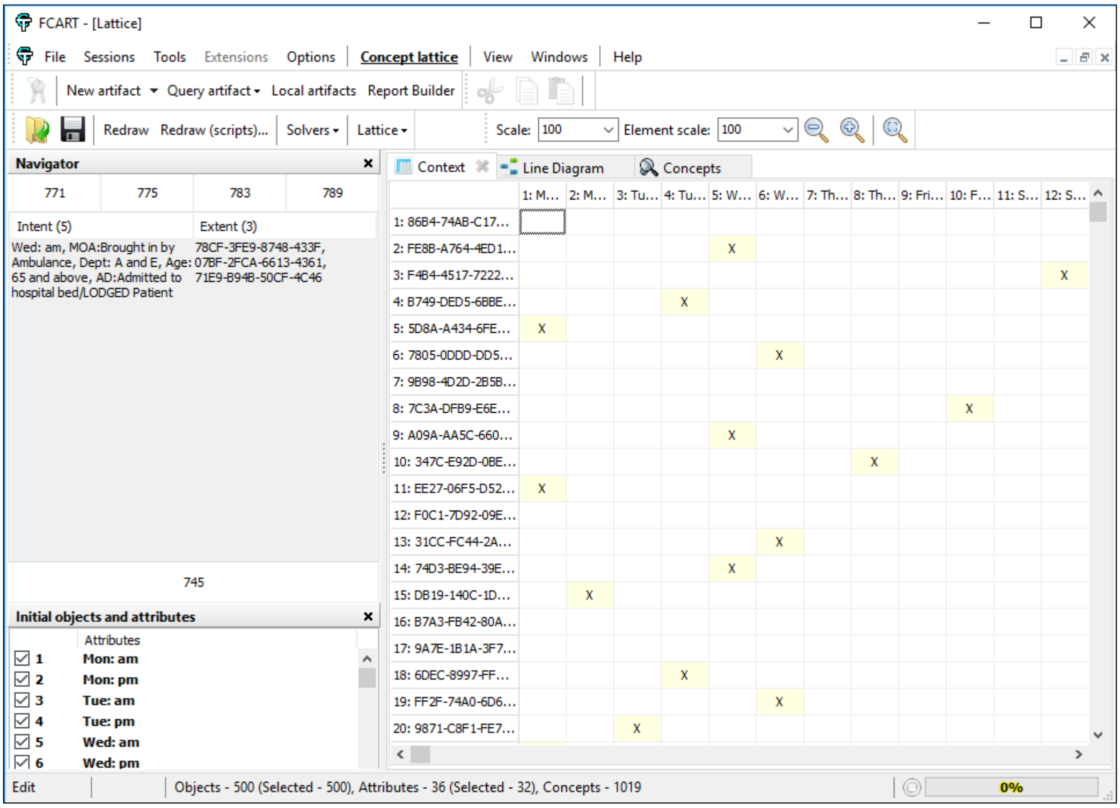


Figure 5.7: The Reduced Concepts

Thirdly, the concept lattice is developed using ConExp. Initially, FCART is used but from here, an analysis is performed using ConExp software because FCART only support line diagrams but not concept lattice, while ConExp on the other hand cannot support large data for cross-table and formal concept generation. Figure 5.8 shows the formal context for the selected attributes. The purpose of generating the formal context from ConExp is to generate a new concept and a concept lattice from the reduced attributes. A concept lattice provides a hierarchical order visualization between the discovered formal concepts as shown in Figure5.9. The concept lattice of (X,Y,I) is the set of all the formal concepts of (G,M, I), ordered by the subconcept–superconcept relation. For X' ⊆ X, Intent (X') is the set of attributes owned by all objects of X', and Extent(Y') is the set of all objects that own attributes Y'.

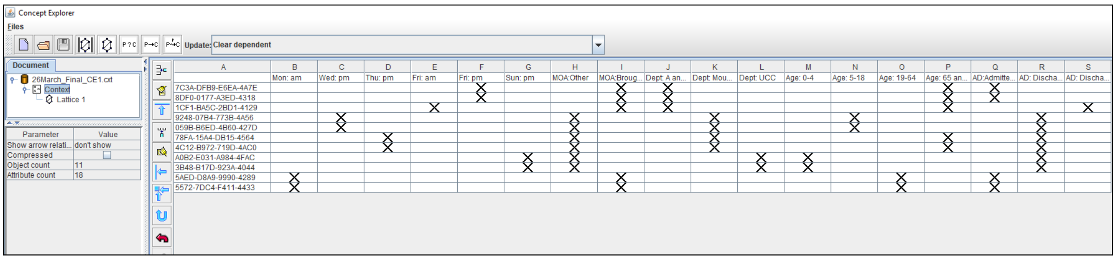


Figure 5.8: Formal Context with Reduced Attributes

In a cross-table, associating an object to the attributes created a concept hierarchy that can be visualized using the concept lattice. Based on the definition of concept intent and concept extent in subsection 5.4, a detailed of concept example of hierarchy as shown in figure 5.10 described that set A contains five objects (*Patient ID: 5AED-D8A9-9990-4289; 5572-7DC4-F411-4433; 7C3A-DFB9-E6EA-4A7E; 8DF0-0177-A3ED-4318; and 1CF1-BA5C-ED6A-4A5E*) as known as extent consists of objects that have all the attributes of the intent, and set B contains attributes (MOA: Brought in by Ambulance) that all objects in the extent have in common. Other attributes that fall under B are called its subconcept. Concept intent B and intent B will lead to the generation of the implication basis in the next section.

5.5.1 Implication Relation Extraction

From the definition of implication in subsection 5.4, an implication $A \Rightarrow B$ holds in (X,Y,I) if and only if $B \subseteq A''$, which is equivalent to $A' \subseteq B'$. It then automatically holds in the set of all concept intents. ConExp generates all the implication for the formal context. Figure 5.11 shows 19 valid implications (with one or more objects) for the formal context that implicitly described the process relation in A&E department and class properties in MVSimO. Translation of the implication to class properties using the logical reasoning method by (Xiao Hang Wang et al., 2004), led to the development of a new set of class properties. Considering attribute *Age* and *Attendance Date/Time* belonging to *Patient* class (Subject) in MVSimO, only attributes *Mode of Arrival, Admission or Discharge* and *Department* are taken as class properties to describe the process in A&E departments. Based on the logical reasoning of the first-order predicate - a subject, a verb and an object and referring to the classes in MVSimO, a subject is a class e.g *Patient*, a verb is a properties-derived implication, and an object is a class e.g *Department*.

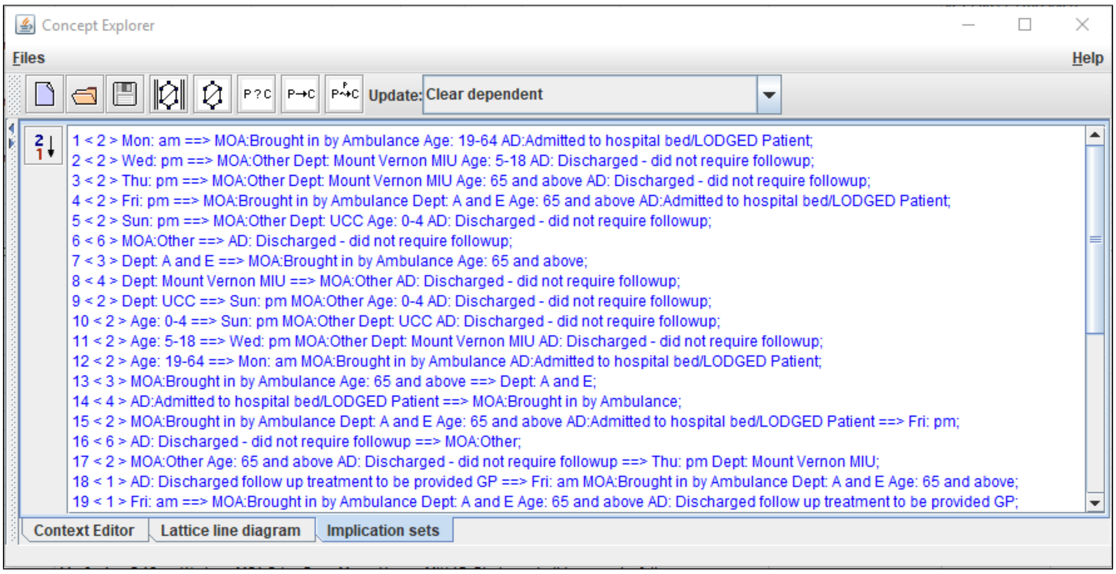


Figure 5.11: Set of Implications Basis

Implication	Natural Language
<2 >Wed: pm ==>MOA:Other Dept: Mount Vernon MIU Age: 5-18 AD: Discharged - did not require followup;	2 patients attended on Wed (pm) also checked-in by walk in; received treatment in MIU; and with discharged no followup
<2 >Thu: pm ==>MOA:Other Dept: Mount Vernon MIU Age: 65 and above AD: Discharged - did not require followup	2 patients attended on Thu (pm) also age 65 and above; checked-in by walk in; received treatment in MIU; and discharged with no followup
<2 >Fri: pm ==>MOA:Brought in by Ambulance Dept: A and E Age: 65 and above AD:Admitted to hospital bed/LODGED Patient;	2 patients attended on Fri (pm) also age 65 and above; checked-in by ambulance; received treatment in A&E; and admitted to hospital
<2 >Sun: pm ==>MOA:Other Dept: UCC Age: 0-4 AD: Discharged - did not require followup	2 patients attended on Sun (pm) also age 0-4; checked-in by walk in; received treatment in UCC; and discharged no followup
<4 >Dept: Mount Vernon MIU ==> MOA:Other AD: Discharged - did not require followup;	4 patients received treatment in MIU also arrived by walk in; and discharged with no followup

Table 5.6: Implications to Natural Language

From 19 valid implication, only 11 implications that have all process element - *Mode of Arrival; Outcome - Admitted or Discharged; and Department* are selected to represent process relation. This is to ensure this step provides as significant results as possible. Valid implication-translated to natural language statement is summarized in Table 5.6. From the table, the subject-verb-object predicate are derived. Concisely, **Subject** represented by *Patient*; **Verb** is represented by *Check-in; Treatment; Outcome* and **Object** is represented by *Department; Hospital; Ambulance*. From this, properties like *Process* is takes place at *Department*; *Check-in, Treatment, Outcome* is a *Process* and *Patient* has a *Patient ID* are derived.

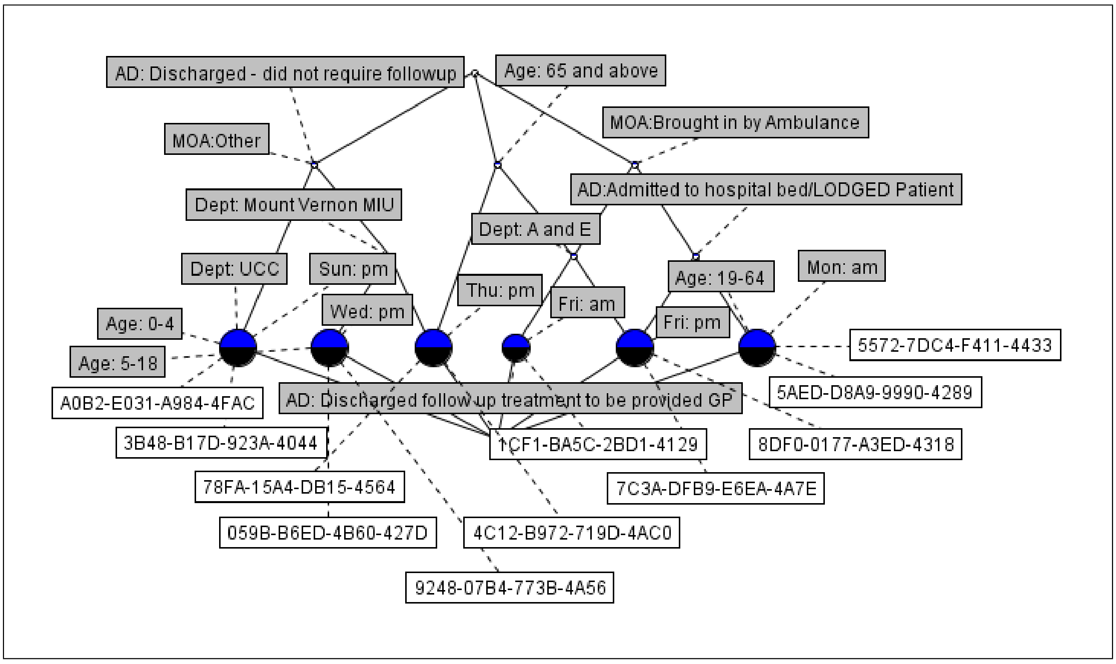


Figure 5.9: Concept Lattice

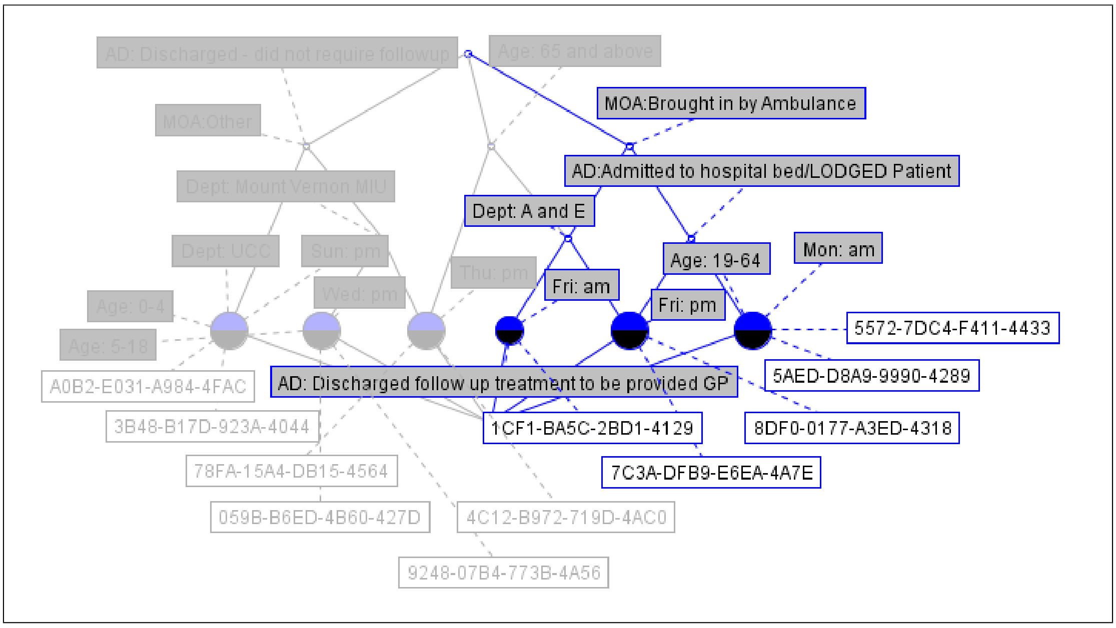


Figure 5.10: Concept Intent and Concept Extent

5.6 Concept Validation

Applying FCA to extract domain concepts and process relations has resulted in the production of a hierarchy model of concept lattice that captures the clustered relationship of the data. The lattices output represent data in an ordered manner of subconcept-superconcept relation. The translation of statements derived from generic A&E pathways (e.g "*Patient check-in by ambulance*") to ontology-class-property statement (e.g "*Process is taking place at Department*") justify the purpose of this iteration to transform process relation into ontology class property using FCA techniques. The building and evaluation of the research artefacts is an attempt to justify the research processes for meeting the objectives of the iteration (March and Smith, 1995).

The effectiveness of FCA technique in extracting domain concepts and relations are evaluated in next iteration, where MVSimO is created to develop an ontology-based simulation model. The evaluation strategy of this iteration composes of comparing the simulation modeled using MVSimO with an existing simulation model (Castellanos et al., 2017). The artefact of this iteration, the activity here is the A&E process elements-extraction using FCA that will be used as MVSimO class properties that need to be evaluated after.

5.7 Learning Outcome

The primary points of learning are:

- Single-valued column is defined based on the application objectives. As illustrated in the sample taken from Table 5.2 of the many-valued to single-valued transformation, a many-valued of `EMAttendanceDate` is transformed to single-valued of day and AM (Ante meridiem) or day and PM (Post meridiem). This is to minimise the efforts in building MVSimO, where the grouping will only reflect the relations between one attribute to another as a larger group (of AM or PM), rather than of smaller groups (of hour or minutes). As the aim of the minimal viable ontology, the research is to operate with manageable data that is capable to work successfully through which the learning can be validated (Ries, 2009).

- The transformation of implication to relation process using the first-order predicate: a subject, a verb and an object, may lead to the identification of implication-transformed natural language. Implication basis with one or more objects are selected.

- The formal context is generated twice, **first** to find the best number of Patient ID to best represent the relationship between objects, **secondly**, generated and reduced to attributes with process elements for a minimal abstraction of the domain. The minimal abstraction may lead to the identification of another type of process that exists in A&E departments. As for example in subsection 5.5, the focus is on the process after receiving the treatment (*Outcome*) - either the patient is being admitted or discharged. Another process can also be selected, e.g *CheckIn*, *Assessment* or *Treatment*. The selection of the *Outcome* process is made based on the availability of the data.

- The concepts from the concept lattices can be used in simulation modeling by providing the class properties in MVSimO. In order to create a model of object-oriented approach that complies with the system's functional requirements that is independent of implementation constraints, the requirements are organized around the objects, which bind both behaviours (processes) and states (data) based on real world scenario that the system interacts with. The artefacts produced in this phase are the formal concepts of the data and its taxonomic relations of superconcept and subconcept. Non-taxonomic relations such as *is a*; *has a* and *is takes place at* are also semantically derived from the implications basis. Apart from that, the *Space-Time-Process* map is developed earlier has an influence on the decision in the selection of the process elements throughout this iteration.

5.8 Summary

This chapter validates the theory of this research, that understanding and abstracting the domain is essential in making explicit assumptions in modeling a simulation, or in the context of this research, developing a simulation ontology for DES model, understanding and abstracting the domain is essential. This iteration contributes to a detailed mapping of A&E departments and integrates the element *Space*, *Time* and *Process* of DES model. The iteration also upholds the finding in previous iteration with the abstraction of process component to achieve minimal working research output. The approach adopted in FDiMe has proven to be efficient in detailing and generating the process relation to be used in MVSimO development. Overall, the method has proven to be efficient through the introduction of the process mining technique to support process relation extraction for MVSimO class properties. The learning that emerged from this iteration highlights number of issues and challenges that can be employed to direct future research.

Chapter 6

Iteration 3

6.1 Introduction

Providing an efficient phase-refined knowledge representation scheme that covers the semantics of the terminology, is one of the challenging tasks in making a better-informed decisions tool (Chandrasegaran et al., 2013; Verdonck et al., 2018; Robinson, 2013). The result achieved in the previous iteration from applying FDiMe framework on the A&E data followed by performing the concept analysis, has resulted in the development of a set of process relations and transformation rules that can be applied to the final ontology - the Minimal Viable Simulation Ontology (MVSimO). As mentioned earlier in this research, MVSimO is used for semantic-based simulation modeling for the A&E domain. The process relations extracted from the previous iteration contribute to the ontology development by providing properties to the classes. This iteration extends the earlier work by emphasizing the process of extracting the semantic elements of the A&E processes and MVSimO development. Deriving semantic elements from existing domain knowledge like ontologies and real dataset descriptions, provides process-oriented semantic representations of the ontological models to reach its full potential (Bell et al., 2007). Furthermore, it contributes to a detailed practical evaluation of the simulation model, addressing the different aspects of the simulation.

The rest of the chapter is organised as follows. The chapter starts with Section 6.2 that discusses how Design Science Research is applied to execute this iteration. Section 6.3 explains the first mini iteration that introduces the framework for the semantic event elements extraction that fulfils the realm of Discrete-Event Simulation (DES), called an event. The framework is structured into three series of operation; Event Interpretation; Event-Content Scoping; and Event Harmonization. Section 6.4 presents the second mini iteration of MVSimO refinement and domain expert evaluation, detailing the extraction techniques and the findings. The implementation of the framework and the evaluation measures to evaluate the framework as the second mini iteration is discussed in Section 6.5. The learning outcome of this iteration is discussed in section 6.5. Finally the chapter summary is presented in Section 6.6.

6.2 Design Science Research and Output Artefacts

The learning outcome of Chapter 4 and Chapter 5 have directed the development and implementation of MVSimO in this iteration towards proving the semantics satisfactory and the ability to provide feedback for future ontology and simulation development (Ries, 2009). In essence, it provides the theoretical grounding for the research, to illustrate how and why the approaches proposed previously, can be used for simulation modeling. In order to take this research to the next level, it is vital to validate the generality of the obtained ontologies, the Space-Time-Process map and the developed process relations by understanding how and why they are applicable for semantic-based simulation modeling. The purpose of this iteration as a whole is to achieve a minimal semantic terminology of the domain to be used for simulation modeling, and applying a rigorous evaluation to the simulation ontology developed. The proposed framework demonstrates the formalization of event element-related semantics, leading to a consistent interpretation and use of existing domain knowledge to develop a semantic-based simulation model. Figure 6.1 shows the research iteration for Chapter 6.

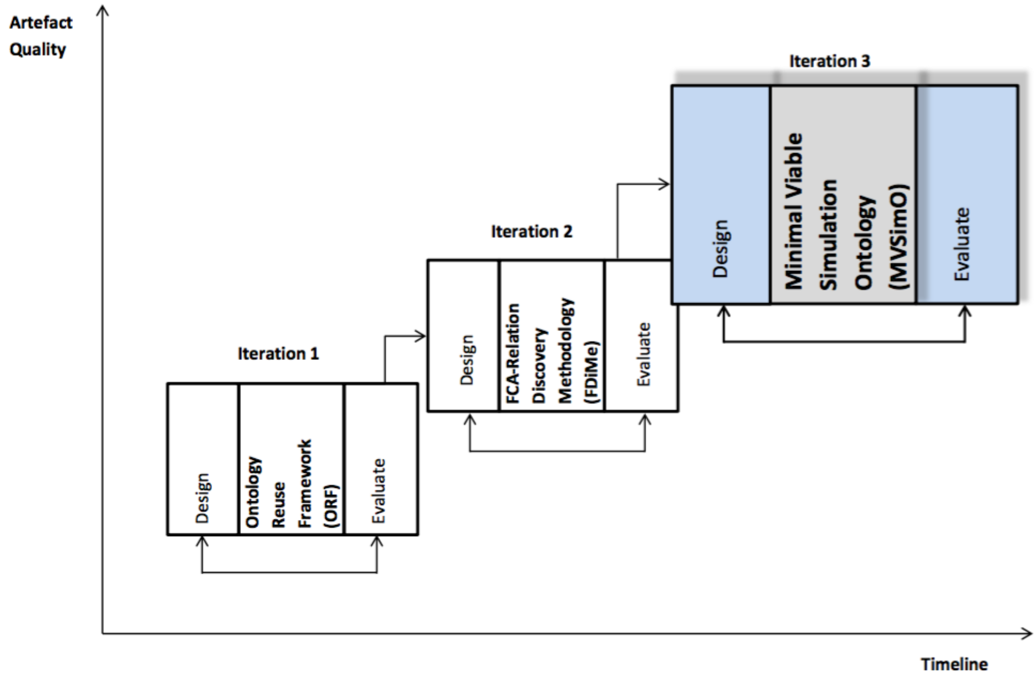


Figure 6.1: Research Iteration for Chapter 6

This iteration follows a design science research approach which first identifies the problem area and its relevance, by suggesting new refinement to the research artefacts (ontologies and process relations). Second, it builds the framework as a design artefact. These steps operate collectively as one mini iteration. Two artefacts, namely the Event Ontic Commitment (EOC) and the Event Harmonized Model were produced through the first mini iteration and finally evaluated the application of the framework through a relevant scenario as the second mini iteration. This chapter proposes a framework for deriving a semantic event element from A&E processes, which attempts to answer the query of "*How to build a semantic-based simulation model?*". The framework adopts the implementation of deriving semantic web services framework by Bell et al. (2007) in deriving the event semantics. As illustrated in Table 6.1, the iterative cycle of this iteration is performed based on the general methodology of Design Science Research by Vaishnavi and Kuechler (2004).

Steps	Method	Input Artefact	Output Artefact
1. Event Interpretation	Real-world commitment interpretation	STP map, Ontologies, Process Relations	Event Ontic Commitment (EOC)
2. Event-Content Scoping	Incorporation of classes, properties and individuals	STP map, Ontologies, Process Relation	Event-Content Representation
3. Event Harmonization	Integration and Semantic Binding	EOC	Event Harmonized Model, MVSimO
4. Framework evaluation	Framework Evaluation Strategy	Framework Evaluation Characteristics	Evaluated Framework

Table 6.1: Iteration 3 - Input Output Model

6.3 Artefact Building and Development

A framework is developed for extracting the semantic content from syntactic A&E pathways and representing such semantics in ontological models. An analysis of obtaining MVSimO is carried out iteratively begins with Iteration 1 for decisions to select existing ontologies, followed by Iteration 2 for domain attributes and process associations, and finally in this iteration, to develop and implement MVSimO. The framework is based on the principles of Content Sophistication presented by Partridge (1996) as well as Daga et al. (2005). Content Sophistication constitutes a process for improving the semantic contents of the system by providing a model of the domain that focuses on the semantics and relationships of objects existing in the real world. The framework proposed in this chapter follows the same manner and provides the basis for enacting the semantics of A&E processes.

The proposed framework addresses three objectives: (1) Semantic derivation of previously extracted process relations; (2) Representation of the derived semantic in an ontological model; and (3) An integration of process semantic with an existing domain ontological model. In essence, the framework aims to iteratively develop an ontological model within and across the domain for Semantic Web. These objectives define the scope and the process of the framework in the first mini iteration. The process, which drives the discovery and representation of semantic contents of the process, is summarized in three series of operations, Event Interpretation; Event-Content Scoping; and Event Harmonization. The evaluation of the framework is evaluated in the second mini iteration. The framework incorporates Space-Time-Process map (refer Figure 6.2) to replicate the A&E domain; and the domain knowledge from existing ontologies is supported by the process relations from Formal Concept Analysis.

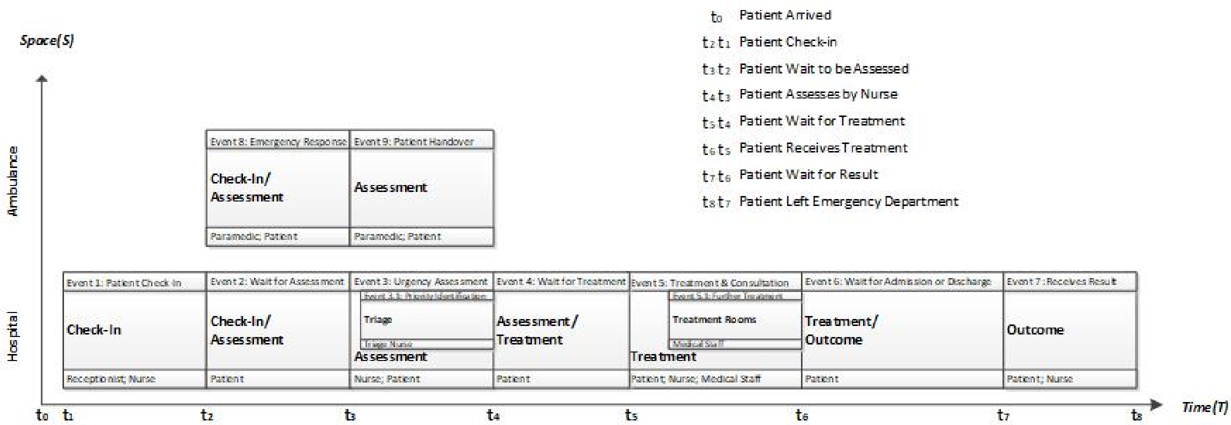


Figure 6.2: A Space-Time-Process of 'Patient Flow in A&E Departments'

6.3.1 Event Interpretation

Event Interpretation model as shown in Figure 6.3 transcribe the events experienced by a patient or a medical staff in A&E as described in Space-Time-Process(STP) map that has limited semantic underpinning. The event in STP is outlined by the process block in space and time dimensions. Events here are in the form of a process block with the location and the entity involved during the event. For example, in Hospital space of *Check-In* process, a description can be found in STP as a combination of event activity *Patient Check-In* and *Wait for Assessment* performed by the entity of *Patient* and Nurse. Interpretation of the events are performed under the capacity to have an effective DES, and its semantic description is enhanced through the integration of existing ontology artefacts obtained in Iteration 1.

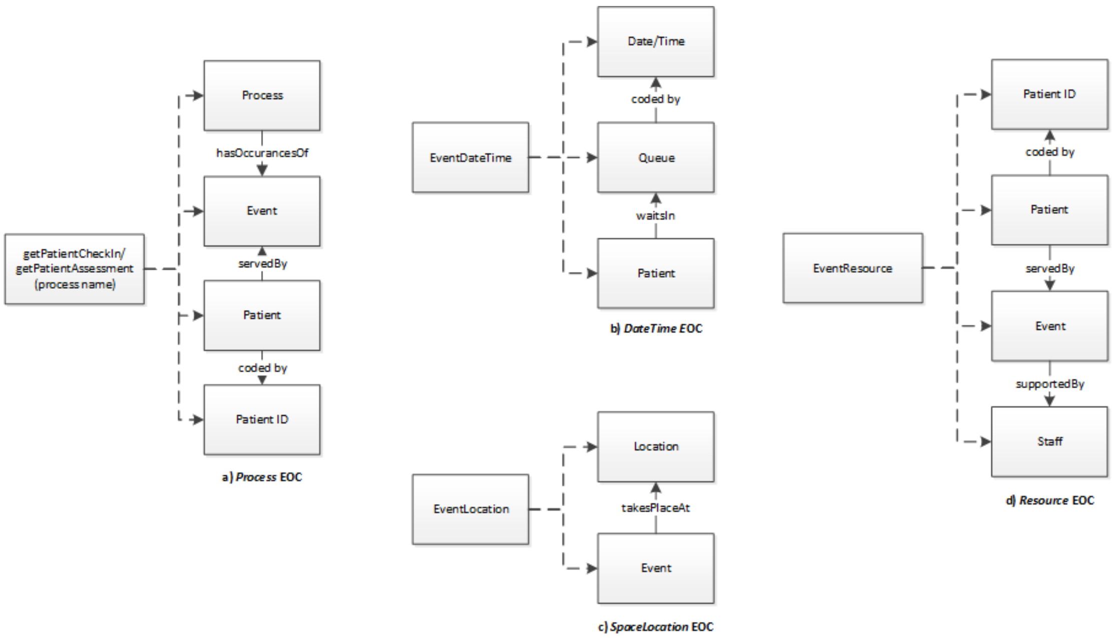


Figure 6.3: Event Ontic Commitment (EOC)

Event interpretation aims to add a perceivable semantic description based on its commitment towards the process. This means disentangling and making as explicit as possible the real objects of which their existence are recognized by. Interpretation activity breaks events into its fundamental parts, which incorporates DES elements of process, date/time, location and resource. Each part is modeled as Event Ontic Commitment (EOC) to represent its real-world commitment towards a process. The concept of the EOC models adopt the Object paradigm by Partridge (1996), which was specifically designed to represent real-world semantics in business modeling. Clarifying the ontic commitment of the process enables the extraction of explicit information that a process commits to. Furthermore, the interpretation of event into an EOC model as an object allows the object to be reused later in the application domain.

6.3.1.1 Event Ontic Commitment (EOC)

The minimalism aspect of MVSimO drives this activity to create objects focuses on process, date and time, location and resource based on the simplified version of the simulation model in Pidd (2014). The activity has extracted *Process EOC*, *DateTime EOC*, *SpaceLocation EOC* and *Resource EOC*. The process name in the EOC represents the process that is going to be built in a simulation modeling. For this research, the simulation model observes the processes of *Patient Check In* and *Patient Assessment* processes, which are renamed to *getPatientCheckIn* and *getPatientAssessment* respectively. Technically, multiple events are performed to complete the process and patient coded with unique Patient ID served by the events. For example, *Check-in*, a Patient waits in a queue for triage (Event 1); Patient being assessed by Nurse (Event 2); and Patient waits for Treatment (Event 3). Date and time commitment for the simulation is a time taken in a specific date while patient waits in a queue. *EventDateTime* for *Patient Check-In* and *Patient Assessment* processes indicate the time patients enter the queue until the time patients leave the queue. For example, from STP, the queue between the *Check-In* and *Assessment* is represented by the *Check-in/Assessment* block with queue a between t_2 to t_3.

The location of each event takes place in A&E department is denoted by *EventLocation*. From STP, the Hospital represents a bigger physical space where all the processes or events are performed, while the smaller blocks such as the *Triage* and *Treatment Rooms* are smaller physical locations where specific events are performed. The process relations extracted from FCA shows the "*Process* is takes place at *Department*" where *Department* is a location in Hospital. The resources involved in the execution of the events is denoted as *EventResource*. In EOC, the resources commitment for the events are *Patient* and *Staff*. Here, *Staff* can be a Nurse, a Medical Staff, a Paramedic or a Receptionist. The resources commit to the event by supporting it, or served by it. By inputting the process relations from the real data, event interpretation activity reaches its best, in fact, the real input greatly help in clarifying the meaning of type-level data (Bell et al., 2007).

6.3.2 Event-Content Scoping

The activity of Event-Content Scoping aims at allocating process-dedicated object from EOC models to existing ontological model obtained from Iteration 1, the Minimal Domain Ontology (MinDO). The domain requirement analysis conducted in Iteration 1 revealed an important relation and object depictions, thus providing the semantic content of the research area . The objects identified in the previous activity and represented with their relationships are scoped to MinDO. The process is conducted using a restricted set of the first-order formula (a subject, a verb and an object) with reference to Ontology Web Language (OWL), where subjects and objects represent the class; and a verb represents the class's property. For example, "Process hasOccuranceOf Event" is denoted by Class *Process*, *hasOccuranceOf* property, and class *Event*. At this stage, classes in MinDO is refined and extracted to create a new ontology, MVSimO

6.3.2.1 Incorporating Existing Ontologies into MVSimO

Incorporating MinDO enables the detailing of the relationship of the real world knowledge. The artefact can be regenerated to incorporate new ontologies to extract new classes and relationships. This step is necessary to allow for the flexibility of the framework, and to enable the ontology engineer to go back to this step to add new classes and properties and create new ontology. The process of existing ontology adoption in this activity has resulted in the creation of new *classes, properties* and *individuals.* Apart from class properties from the existing ontologies, as mentioned earlier, new properties of MVSimO are also derived from the real-world knowledge through the formal concept analysis. Individuals or instances are from the A&E data. Table 6.2 summarizes the decisions made and actions taken to incorporate MinDO into MVSimO. Each object is combined and to create the initial ontological model.

Objects	**MinDO Class**	**MVSimO Class**
Process hasOccuranceOf Event	Process	Process, new class (Event)
Patient codedBy Patient ID	Patient, Electronic Health Information Type	Patient, new class (Patient ID)
Patient waitsIn Queue	Patient	Patient, new class (Queue)
Queue codedBy Date/Time		new class (Queue), new class (Date/Time)
Event takesPlaceAt Location	Facility	new class (Event), renamed (Facility) to (Location)
Patient servedBy Event	Nurse Designation	Patient, new class (Staff) with subclass (Nurse)
Event supportedBy Staff	Nurse	new class (Event), new class (Staff) with subclass (Nurse)

Table 6.2: Event-Content Scoping

The EOC models identified in the previous activity and represented with their relationship are scoped into MinDO to illustrate the first version of MVSimO. Figure 6.4 shows the first cut of MVSimO. The initial version of MVSimO ontological model shows patient flow during the events of *Check-In* and *Assessment.* It is called the initial model because only two processes has been interpreted as the example from this research. This iteration enables further developments with a range of processes to refine the model. Classes from MinDO like *Nurse, Patient, Date/Time* and *Location* derived from MinDO proves the semantic relation from existing ontology and the EOC models. Further refinement in harmonization of the first-cut MVSimO ables to create a more detailed ontological model.

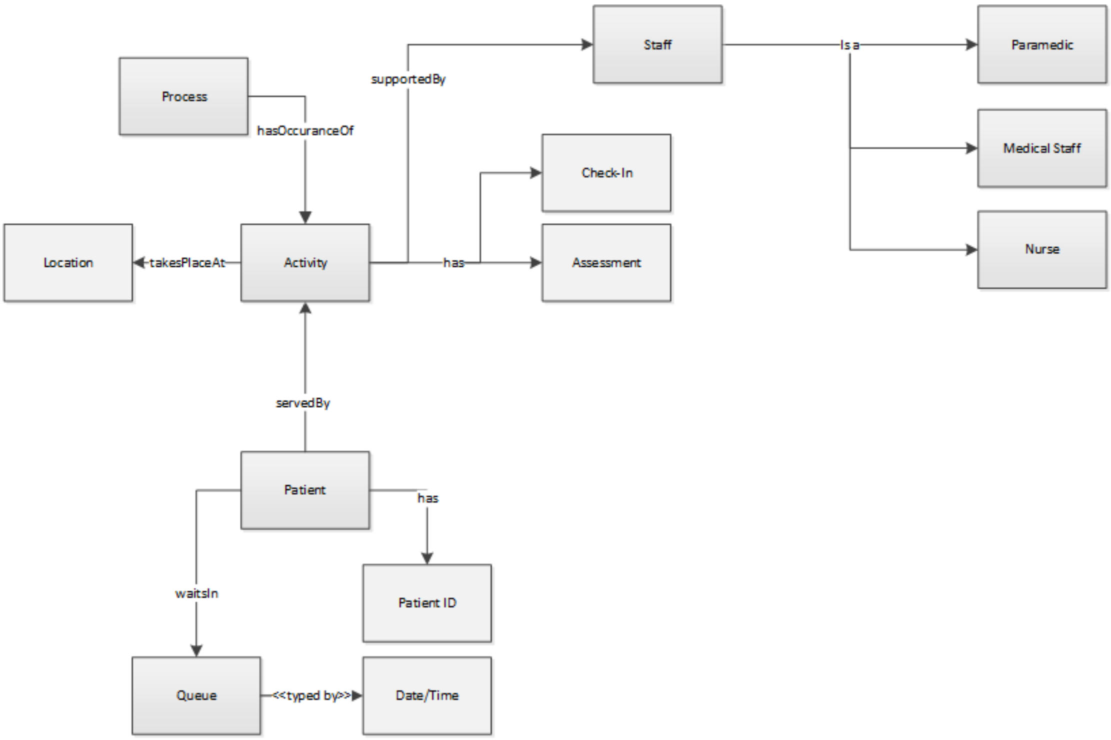


Figure 6.4: First-cut MVSimO

6.3.2.2 Event Harmonization

A preliminary analysis of domain requirement performed in Iteration 1 has led to the identification of Discrete-event Modeling Ontology (DeMO) to support this research. DeMO bridges the gap among the semantic concepts deduced in the DES world views and software tools (Silver et al., 2011). In the event harmonization activity, the first-cut domain model is combined with DeMO. To be more specific, this activity uses a process-oriented ontology subclass in DeMO named the *ProcessOrientedModel* or PIModel. Ontologically this enables an explicit mapping between real-world DES elements and the domain they serve. The harmonized model is mapped into DeMo PIModel accordingly, and may be translated later to an XML and then into a simulation model. This allows researchers, domain experts and modellers to share a common understanding of the concepts and the relationship of the domain. Ontologically, harmonized model enables an explicit mapping between a process (with its parts), and the domain it serves supported by A&E data.

Figure 6.5 presents the harmonization model derived from the previous Event Interpretation and Event-Content Scoping activities. The *getPatientCheckIn* and *getPatientAssessment* processes are defined within DeMO and their parameter are typed in relation to the respective classes. From the diagram, *Process* and *Patient* classes are derived from MinDO; , *Location*, *Event*, *Staff*, *Nurse*, *Queue*, *Patient ID* and *Date/Time* are new classes or renamed classes after the Event-Concept Scoping activity. Class *Check-In*, *Assessment*, *Treatment*, *Outcome*, *Paramedic* and *Medical Staff* are from Space-Time-Process map, and finally *Activity*, *Entity*, *Queue*, *Location*, *Resources*, and *Process* are mapped-out from DeMO.

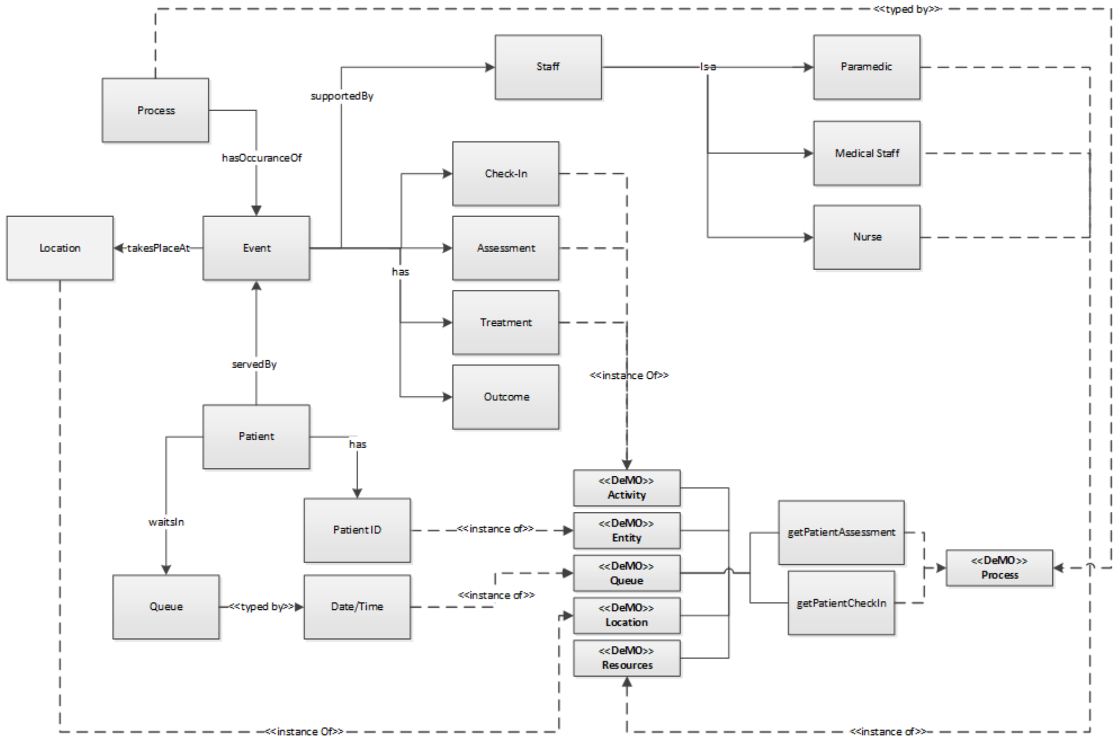


Figure 6.5: The Harmonized Model

The initial MVSimO ontological model (Appendix D - MVSimO Ontological Model) developed from EOC models shows the processes of *getPatientCheckIn* and *getPatientAssessment*. The processes are defined within DeMO and their parameter are typed in relation to the respective MinDO. The result is a combined semantic graph in which both domain objects are represented and linked together. Figure 6.6 exemplifies the fragment of mapping of *getPatientAssessment* event-domain topology provided by the subclasses of DeMO (Process, Location, Entity, Queue, and Resource).

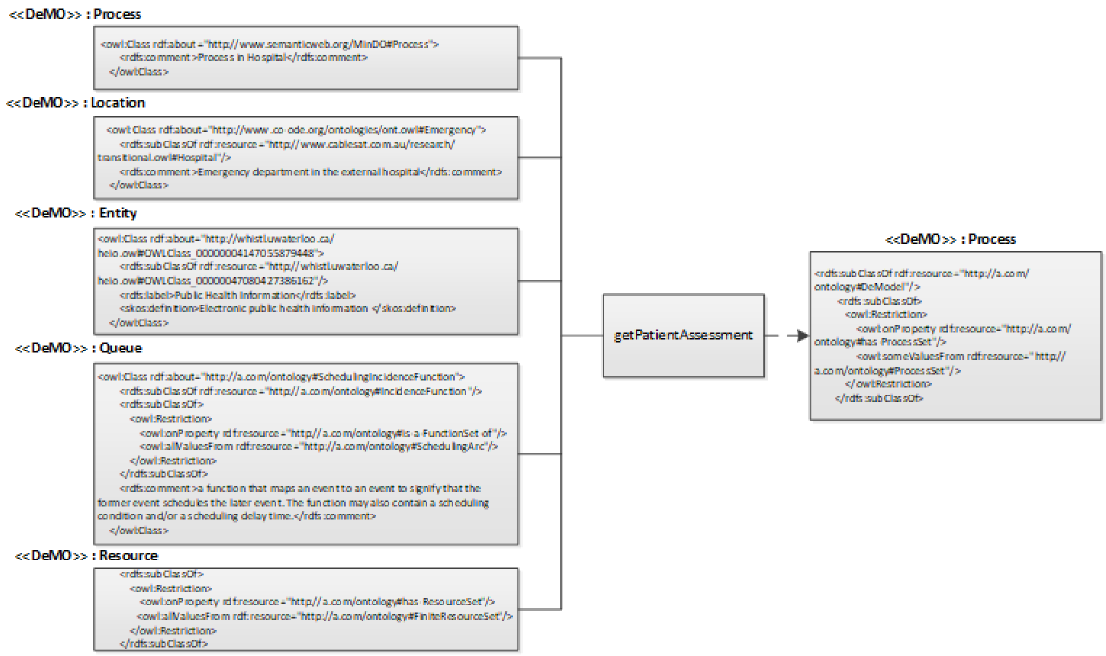


Figure 6.6: 'getPatientAssessment' Process-Ontology Topology

6.4 MVSimO Similarity Evaluation

The evaluation of MVSimO is a step to evaluate the functional adequacy and quality of the ontology provided its context and goal are given (Cantador et al., 2007). The evaluation involves assessing the quality of MVSimO, by conducting an object-by-object analysis to judge the model. The evaluation process is conducted by comparing MVSimO with an existing simulation model and measure their similarity. Existing model by Bell et al. (2017) known as Cumberland model is developed using the same set of data as this research is selected as an evaluation baseline. The hybrid model focuses on the distributions and routing patterns before and after A&E departments, tested on a number of different scenarios. Even though MVSimO model illustrates the process happening in A&E departments, both models assess the same objects to establish the process.

The model represents the possible strategic change scenario accepting data from System Dynamic model (SD) to DES model. The hybrid model focuses on the distributions and routing patterns, tested on a number of different scenarios. Table 6.3 shows the similarity assessment of the elements in MVSimO with elements in the Cumberland Model. Each element in MVSimO is compared to the elements in Bell's model to assess its representation of real-world phenomena. The assessment checks the viability of MVSimO in conforming to various characteristics or qualities that are expected of it. The process of calculating the similarity starts by manually extracting the elements in both models. Then the elements are categorized into different features as presented in Bell et al. (2017). The categories are: 1) Scope; 2) Scenario; 3) Data Group; and 4) Objects. Each category is then compared with the elements in MVSimO model. Finally, element-by-element analysis of both models is conducted to find the similarities.

Category	Cumberland Model	MVSimO
Scope	Before and after A&E	During A&E
	Discharge by hour	getPatientAssessment
	Length of stay by hour	getPatientCheckIn
	Pediatric arrival	
Scenario	Bed number	
	Care home support	
	Ambulatory services	
	Additional primary care	
Data Group	Patient Age	Age
	Arrival Method	EMModeofArrivalDescription
	Arrival Time	EMAttendanceDate
		DepartmentDescription
		AttendanceDisposal
Objects	Departments	Process
	Demographic Data	Activity
	Resources	Location
		Patient
		Resources
		Queue

Table 6.3: First-cut MVSimO Objects Assessment

In the *Scope* category, Cumberland model focuses on non-elective care and its associated services before and after A&E departments capacity. But for MVSimO, the ontological model aims at the processes within the department. Though each model has diverged into a different perspective, both enfolded the same domain structure. For the simplification of this research, no scenario has been presented because the research aims at developing the minimal-viable ontology as a foundational ontology. New scenarios and the detailed scope or processes can be added during the refinement process of a detailed simulation model. Data grouping similarity of MVSimO elements is categorised based on the analysis of its class instances. The analysis checks whether the instance serves the same purpose of the data group in Cumberland model. If it does, the column belongs to the instance is considered to be similar to the data group in Cumberland model. For example, in Cumberland model the Arrival Method matches the EMModeofArrivalDescription column name in MVSimO; with instances of *Ambulance* and *Other*. Please see Appendix C for a snippet of A&E Data Instances of MVSimO classes.

6.4.1 Similarity Analysis

The evaluation to assess the similarities is conducted by assessing the interception of both models in term of shared elements. Based on the four categories described in Table 6.3, the *Scope* and the *Scenario* categories of Cumberland model do not have any similarity with MVSimO due to the fact that each model has different aims and has different level-of-detail. In the context of assessing the similarity, only two categories are measured; the *Data Group* and the *Objects*. The similarity assessment is evaluated as depicted in Figure 6.7.

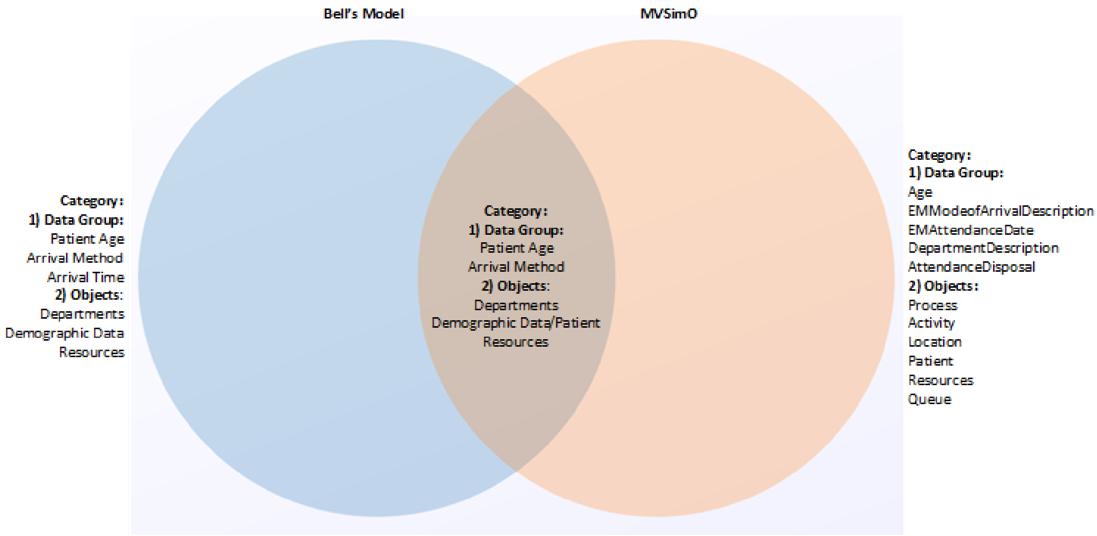


Figure 6.7: Harmonized Model

From Figure 6.7, it can be concluded that MVSimO has 83.33% similarity with Cumberland model (see below) and more elements are extracted from MVSimO.

Number of elements extracted from Cumberland Model = 6
Number of elements extracted from MVSimO = 10
Number of similarities = 5
Similar elements from MVSimO divided by the number of elements extracted from Cumberland Model = 5/6
The percentage of similarity = 5/6 × 100 = 83.33%

6.5 Learning Outcome

The primary point of learning are:

- The analysis of the structural layer of MVSimO revealed that there are few elements in Cumberland model, that MVSimO has missed, and also a few concepts in Cumberland that can be improved by MVSimO elements. Identifying new scopes and scenarios for intended model may lead to a further refinement of MVSimO, which is also known as Ontological Improvements (Sabou, 2005). The minimal viable ontology produced during this iteration is derived from ongoing research processes initiated by the Ontology Selection Framework (OSF) followed by FCA-Relation Discovery Methodology (FDiMe). The preliminary analysis of applying the previous artefacts to produce MVSimO has led to the identification of the need to extend it to adopt other processes in A&E domain.

- Clarifying the ontic commitment may lead to the identification of the simulation components for reuse. For example, in Figure 6.3 the process of *getPatientCheckIn* in *Process EOC* is represented by the commitment of *Process*, *Event* and *Patient ID*. This interpretation allows the reuse of *getPatientCheckIn* as a process component not only for the DES model within the domain, but also across the domain.

- The existing ontology from Iteration 1, the Minimal Domain Ontology (MinDO) is incorporated into the framework artefacts to enrich the semantic content of existing knowledge. It is clearly evident that the knowledge captured from earlier design decision has enabled the simulation of scenarios (Chandrasegaran et al., 2013). The union between the components of the representation and the real world determines the semantic content of the data being represented (Benjamin et al., 2006).

6.6 Summary

This chapter validates the theory of this research that efficient phase-refined knowledge representation schemes, that covers the semantics of the terminology can support data sharing and reuse across applications. This iteration contributes to an event semantic derivation framework. A formal definition of the output of the phases consisting of the framework is provided. Another main contribution of this chapter is a thorough analysis of discrete-event modeling, mapped into the Space-Time-Process map and offers real-world effects by data instantiations. The first-cut of MVSimO contains less detailed process and fewer objects compared to the existing simulation model due to the fact that its creation is without any expert's opinion, but yet produced 83.33% similarity. This proves that the framework used to derive the event semantic in developing the DES model of A&E departments has led to the creation of a competent model equivalent to the existing simulation model. The learning that emerged from this iteration highlights a number of issues and challenges that can be developed to direct future research.

Chapter 7

Research Summary

7.1 Research Summary

Domain knowledge is one of the important factors for the accomplishment of a good quality requirements specification, given adequate time and effort for acquiring the knowledge for building an efficient application for the domain, especially if ontology development is involved. With the emerging ontology-based application in fields such as data semantic management (Tao et al., 2017; Daraio et al., 2016), context modeling for manufacturing industry (Giustozzi et al., 2018), information modeling for cost estimation (Lee et al., 2014), and simulation modeling (Traoré et al., 2018), semantically annotated domain knowledge in domain ontology provides shared knowledge representation. Domain ontology also allows an opportunity for interoperability of the application that used the ontology across platforms, organizations and operating systems Lee et al. (2004); Gil et al. (2016); Daraio et al. (2016). Apart from that, ontologies offer an efficient knowledge representation scheme in making better-informed decision (Sernadela et al., 2015a). Despite the growth in ontology-based application, there is an increasing concern over the sufficient amount of semantic terminology in the design phase (Chandrasegaran et al., 2013; Verdonck et al., 2018). Therefore, the omnipresent adoption of the integration of semantic content at each stage of the ontology-based application development may lead to the construction of a well-defined model (Moreira et al., 2016).

To make ontology-based application a practical reality, one needs to have a well-defined and suitable ontology that fits the application purposes. Building an ontology from scratch, nevertheless, is expensive, time consuming and requires the intervention of experts (Kaiya and Saeki, 2006; Lonsdale et al., 2010). The significant amount of ontologies has been built during the last decade (Pinto and Martins, 2000), has resulted in a large amount of ontologies make available on the web. An ontology provides a knowledge-sharing infrastructure that supports the representation and sharing of domain knowledge (Park et al., 2011; Pinto and Martins, 2001; Trokanas and Cecelja, 2016; Kamdar et al., 2017). The notion of domain knowledge as "*small, manageable pieces with strong internal coherence but relatively loose coupling*" (Borst, 1997) leads to the idea of reusing existing ontologies to save time and effort, rather than building new ones. Consequently, this thesis sought to assist ontologists and simulation modelers in building a time efficient ontological model for simulation modeling.

This aim was achieved by analytically extracting domain knowledge and developing the minimal viable simulation-ontology to help minimise the efforts made on modeling a complex and heterogeneous A&E departments' system. The objectives as set out in Chapter 1 are summarised below:

- Objective 1

  Critically assess the existing simulation modeling methods, ontology reuse framework, data analysis techniques and process mining methods to provide an understanding of the state-of-the-art

- Objective 2

  Design and develop a methodological Ontology Reuse Framework (ORF) with a simulation and modeling perspective, and is able to select suitable ontologies for ontology reuse

- Objective 3

Produce a Space-Time-Process (STP) map from a set of generic pathways as a means to better represent the domain

- Objective 4

  Analyse generic A&E processes and extract process elements from semi-structured data

- Objective 5

  Derive an event-semantics from the process elements extracted to further develop a Minimal Viable Simulation Ontology (MVSimO) model, and validate the research outcome with a goal-free evaluation method

In achieving the aim and objectives of the work, Chapter 2 reviewed the varieties of existing ontology-based simulations, ontology reuse frameworks and applicable techniques of domain conceptualization, through the provision of an in-depth understanding of the theory and practice of currently available approaches. In the context of this research, the literature provided the basis for proving how ontology reuse and a well-defined domain conceptualization can assist in a faster simulation ontology development processes (Chandrasegaran et al., 2013; Robinson, 2013; Verdonck et al., 2018; Mohiuddin et al., 2017a). Although applying ontology to a simulation model is mainly limited to the use of application-specific built ontology as the foundation, each real-world model is unique with its own set of requirement specifications. Current research is mainly focused on 1) selecting candidate ontologies from a set of pre-defined requirements, with users' involvement for process modeling in system analysis and design (SIMPERL, 2010; Lonsdale et al., 2010); 2) building a new ontology, or reusing a whole set of existing ontologies for simulation modeling (Durak and Ören, 2016; Grolinger et al., 2012; Silver et al., 2011). Consequently, an opportunity for contribution lies in the introduction of domain modules integration in the conceptualization phase of the ontology reuse framework (Chandrasegaran et al., 2013; Verdonck et al., 2018) and the explicitly detailed domain process mapping for semantic extraction (Mohiuddin et al., 2017a).

Chapter 3 set out the means for achieving the objectives via Design Science Research. This approach provided a means by which to engage the design problem - providing the necessary learning to improve the proposed solution, while at the same time enriching the solution space with the Design Science Research output. The main Design Science Research artefact is the Minimal Viable Simulation Ontology (MVSimO). The overall research methodology is executed in incremental iterations, where each of the three iterations forms a design problem that executes the build and evaluates design activities (Vaishnavi and Kuechler, 2004). The iterations were designed in a way that; Iteration 1 developed the core framework for the ontology reuse, which include a domain conceptualization technique, Iteration 2 extended the effort performed by adding a process elements extraction method, and Iteration 3 strengthened the design artefacts by adding the ontic commitment derivation technique , scoping it around simulation model topology, and validating the artefacts through an established model similarity assessment.

Given that the literature review exhibited limited understanding and effort in the proposed research area, Design Science Research was particularly appropriate, allowing an iterative learning process to feeding the ongoing understanding of the design problem. More specifically, in the case of ontology-based simulation, domain conceptualization in the design analysis was identified as an important stage at the beginning of each cycle. Practical conceptualization methods were not yet well defined, thereby posing another learning challenge in the knowledge space. The products of Design Science Research included constructs, methods and models to facilitate the artefacts development. The build and evaluate design activities were applied in incremental iterations to build and effectively evaluate each of the Design Research products as illustrated in Table 7.1. The evaluation for the Design Science Research products was made by synthesizing the Design Science Research evaluation criteria, as the table illustrates, to create a suitable evaluation method derived from the MVSimO knowledge base as presented in Chapter 3. The evaluation demonstrated a successful application of each product in the final MVSimO method and tool.

	Build	Evaluate	Theorize	Justify
Construct	MinDO STP Map	Completeness Simplicity Ease of Use	Explain why and how constructs work by employing them to describe real case scenarios (addressed in Ch.4&5)	Prove that constructs work scientifically by applying them in models and methods (addressed in Ch.4,5&6)
Model	ORF FDiMe EDSF	Fidelity Completeness Internal Consistency	Adapting theories from the current ontology reuse and simulation model discipline (achieved by theorising in Ch.4,5&6)	Test the models on a real life example to prove them (addressed in Ch.4, 5 & 6)
Methods	ORF Process STP Mapping Process FDiMe Process EDSF Process	Operationality Efficiency Generality Ease of Use	Explain why and how methods are applied (achieved in Ch.4,5&6)	Prove the methods work formally by instantiating them using real examples (achieved Ch 6)
Instantiation	Process Relations MVSimO Application	Effectiveness Efficiency Impact on Environment	Understanding how and why application works in the domain (achieved in Ch.6)	Prove that MVSimO works by testing it DES model of A&E domains (achieved in Ch.6)

Table 7.1: Design Research Products X Activities

Chapter 4 described the first iteration, which concentrated on developing an ontology reuse framework and incorporating comprehensive domain conceptualization for a deeper understanding of the domain. An approach of ontology reuse utilizing existing ontology and promotes knowledge sharing among the Semantic web community has been adopted in this chapter. The Ontology Reuse Framework (ORF) has implemented domain conceptualization in the earlier phase of ORF development, defined a conceptual model and derived a set of modules as ontology search keywords. An initial set of constructs, models and a method was build and evaluated, meeting Objectives 1 and 2 of the study. The domain conceptualization technique formed the pre-processing stage of Ontology Discovery of ORF. The first stage laid out the foundation of the ontology reuse framework by fulfilling the domain specification requirements study. The conceptualization technique applied started by identifying the problem to be modelled and performing modules extraction based on domain requirements. The modules extracted were used as keywords to search for available ontologies using popular ontology repositories.

A thorough study was conducted to determine the most used and recent ontology repositories utilized by researchers of Semantic Web. As a result of ontology discovery, a list of candidate ontologies based on modules-derived keywords was obtained, contributing to the next phase of ORF, the Ontology Selection phase. This early form of ORF output was selected and evaluated based on the selection guidelines for the ontologies to fit the research purpose. The final ORF artefact, the Minimal Domain Ontology (MinDO) was revised and customised, and rigorously evaluated with the OQuaRe Quality Model, where the objectives of this iteration directed the next iteration towards analysing and synthesizing A&E data of Hillingdon Hospital to form another dimension of the semantic model. This iteration highlighted the need to further investigate how to extract process elements in A&E departments, and initiated the following Design Science Research iteration that allows for a process relation that leads to property definitions between ontology classes.

Chapter 5 extended the work to produce Minimal Viable Ontology (MVSimO), and addressed Objectives 2 and 3 of this study by adding Formal Concept Analysis (FCA) to the A&E dataset to derive process relations of the domain. The processes of FCA-relation derivation followed FCA-Relation Discovery Methodology (FDiMe). This second iteration contributed to another set of Design Science Research artefacts facilitating the extraction of process relations based on Space-Time-Process map derived from the generic A&E pathways. The functional aspects of MVSimO were learned through the application of FCA, where minimising implications basis is important to focus on the set of implications basis scoped by the research context. A technique known as Concept Exploration was applied by implementing natural language statement approach to the set of implication basis to identify the class relation properties. 11 implications with process elements of *Mode of Arrival*; *Outcome - Admitted or Discharged*; and *Department* were selected to represent the process relations.

An instantiation of FDiMe as process relations was made and used to prove and evaluate the method. The output of FDiMe, the process elements obtained by following STP scope were compared to the elements from generic A&E pathways to demonstrate the efficiency and effectiveness of FDiMe. It was clearly visible that the process relations derived have produced a more-detailed semantic of the domain by providing properties to the MVSimO classes. It was evident by the end of Chapter 5, that in order to justify and theorize the development of MVSimo, a further iteration was required to take the research to the next level.

Chapter 6 addressed Objectives 4 and 5, showing that the effort of developing MVSimO could be minimised using its validated learning, which were collected throughout the research iterations. The effort was continued by Event-Derived Semantic Framework (EDSF) processes to extract event semantic from the artefacts gathered from previous iterations. The EDSF has demonstrated that there was enough knowledge from which existing ontologies can be reused, complemented by domain data can effectively build a feasible simulation ontology. Overall, the framework has proven to be efficient by introducing the event semantic of domain processes. This last iteration used the learning produced by evaluating, theorizing and justifying the activities, to develop the final ontology, MVSimO for a semantically-rich and well-represented simulation model. A deeper understanding of how and why the MVSimO worked was achieved in the last iteration, by performing a thorough evaluation that enabled knowledge and learning to emerge while MVSimO was applied and allowed to be refined iteratively. Finally, the learning that emerged from the third iteration highlighted a number of issues and challenges that could be employed to direct future research.

In summary, figure 7.1 provides the clear description of the process that has been developed throughout this research, ontology reuse and synthesis. Following the typical methodology of Design Science Research (March and Smith, 1995), the blueprint categorized the tasks for each iteration as requirement analysis, analysis and design, development and evaluation to develop research artefacts. In the swimlane 'Ontology Selection', shows the process of Ontology Reuse Framework (ORF) development in Iteration 1. The steps taken started with Domain Conceptualization & Ontology Discovery, followed by Ontology Selection, Ontology Merging and Integration, and Ontology Evaluation. The framework produced an initial ontology, named a Minimal Domain Ontology (MinDO), as the main artefact for Iteration 1. MinDO complemented the development of MVSimO in the third iteration. In the 'Process Mining' swimlane, it explains the activities of extracting the processes or events performed in the department used Formal Concept Analysis (FCA) approach. Here, the research produced a Space-Time-Process (STP) map from a set of generic pathways as a means to better represent A&E departments to guide the 'journey' in extracting the processes involved in the departments. Formal Context Mapping and Formal Concept Exploration tasks were performed before the process of FCA-Relation Transformation set of process elements to be implemented as class properties of MVSimO. Further development of MVSimO is shown in 'Ontology Development' swimlane through 'Event Interpretation, Event-Content Scoping and Harmonization. The validation of MVSimO was conducted by similarity evaluation to finalize the processes in getting the minimal and viable simulation ontology.

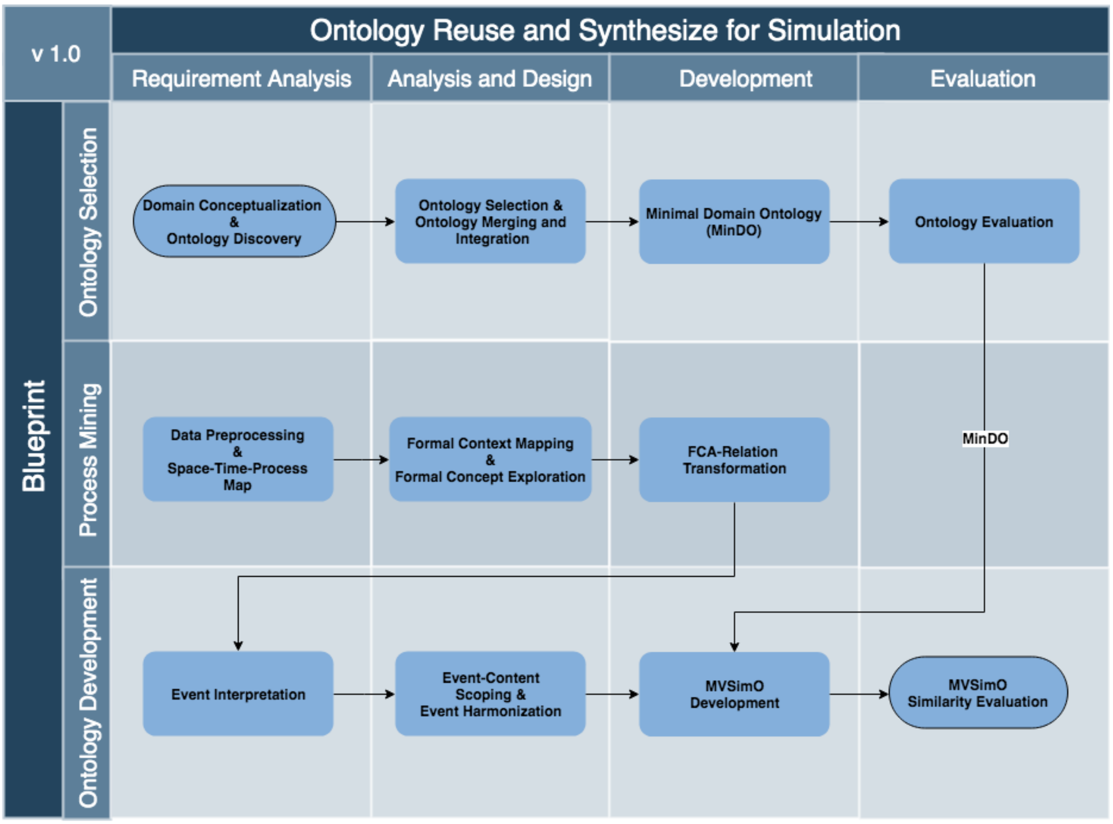


Figure 7.1: Ontology Reuse and Synthesis Blueprint

7.2 Contributions and Conclusions

Research contributions are categorized according to Design Science Research product classification (March and Smith, 1995). In overall terms, the contributions of this research are a novel ontology reuse approach that applied domain conceptualization to transform domain knowledge into modules, which played a role as sub-ontologies. FDiMe which was used to transform process relations to ontology class properties is also one of the contributions. Within the literature, a number of ontology reuse frameworks have been identified, but the studies did not specify it in a context of simulation modeling. The Reuse framework by SIMPERL (2010) provided a guidelines for ontology reuse and, was applied in eRecruitment solution provider environment. A systematic analysis by Kamdar et al. (2017) focused on reusing ontology with a collaboration of term reuse and term overlap across Biomedical Ontologies. A methodological guidelines study by Fernández-López et al. (2013) applied FCA to a set of axioms and its definition to decide which ontology to be reused. All of these approaches, however, are aimed solely at existing ontology knowledge and decisions from domain intervention to assist in decision making, thereby posing a greater challenge for the ontology engineers and simulation modelers to understand the domain within a limited time and budget.

More specifically, the main research contributions and their value are detailed below:

- **The Ontology Reuse Framework (method)** is one of the main contributions made by this research, and it can be applied in different scenarios in a reuse-oriented ontology development lifecycle. The framework emphasizes on ontology selection task as a main component for effective ontology reuse (Stecher et al., 2008). Identifying existing ontology with the most similar concepts and being able to refine the selection decision on which ontology to select require a clear definition of the domain which can be achieved through conceptual modeling (Robinson, 2006). ORF can be applied in different scenarios in a reuse-oriented ontology development lifecycle. Therefore, this approach has the potential to be integrated as the first step of a more complex ontology engineering process. The ORF targets different ontology reuse task; (1) Domain conceptualization, (2) Ontology Discovery, (3) Ontology Selection, (4) Ontology Merging and Integration, and (5) Ontology Evaluation to have the best represented existing ontology for the new domain ontology for simulation modeling.
- **The Domain conceptualization process (method)** The first process in ORF, grasped domain knowledge to assist in the ontology discovery process, by following the idea of conceptual model by Stachowiak (1973). The mapping, reductive and pragmatic features of the conceptual model contributed to the improvement of the modeler's understanding of the domain, and focused on only the intended purpose of the model, resulting in a well-defined specification of the domain.
- **The State-Time-Process Map (model)** is a novel generic method that enabled delineation of the state of space and the continuous progress of process existence in any

scenarios. This model contributed to a generic structured representation corresponding to Discrete-Event Simulation standard of modeling. It can be effectively applied prior to simulation modeling to identify and extract space-time-process relations given a domain pathway. The literature applied a heuristic strategy for business process generation as proposed by Sider et al. (2001) and De Cesare et al. (2014), which implemented the strategy with no intention of simulation modeling. The model contributed by this research is a systematic approach to map a real-world model to its abstraction.

- **The FCA-Relation Discovery Methodology (FDiMe) process (method)** is an effective method that can be applied to identify a process and its elements involved in a domain. The process was aimed at developing a set of transformation rules that can be easily applied in a process-based ontology development to extract process relations using FCA approach. The transformation of FCA attributes to ontology class properties using natural language statement is a novel method specifically tailored to map real-world phenomena to suitable ontology properties between classes.

- **The Event-Semantic Derivation Framework (ESDF) (method)** is another contribution made by this research and can be applied in different scenarios in an ontology development lifecycle. The ESDF was aimed at extracting event-semantic for a detailed ontological model. Apart from that, the event-semantic model derived is a novel model in identifying the simulation component for reuse. This model interpretation allowed the reuse of process component not only within the domain but also across the domain.

- **The Minimal Viable Simulation Ontology (instantiation)** is an application prototype that implemented the ORF, the FDiMe and the ESDF. The ontology can be applied to build a discrete-event simulation model in different scenario efficiently. In the context of this research, the focus were *Patient Checkin* and *Patient Assessment*, but it can also be applied in other scenarios based on the objectives of the simulation model. The learned ontological model can be easily pruned and modified by domain engineers and simulation modelers. The generality of the ORF, the FDiMe and the ESDF in extracting process, events and its semantic elements was clearly demonstrated through the achievement of similar evaluation results in comparison with the Cumberland Model, with 83.33% of similarity percentage.

7.3 Limitations and Areas for Future Research

Though the research has made a number of valuable contributions to the ontology reuse for simulation modeling in complex and heterogeneous domain both in the process and the tools, a number of limitations and challenges have been noted:

- It was noted in Chapter 4 that the concept of ontology reuse as part of ontology development by Pinto and Martins (2000) consisted of finding a suitable existing ontology to be merged and integrated into a new ontology, the Minimal Domain Ontology, MinDo. The ORF has successfully discovered and select the ontologies and proved its applicability by identifying its strength and weaknesses, despite no involvement of domain expert. As discussed in Chapter 2, the evaluation using OQuaRe quality model is used to demonstrate the difference between ontologies in the same domain, which is useful for supporting users in making informed decisions, thus no global score to the ontology if there is no comparison. Identifying the difference between ontology within the same domain by having a gold-standard ontology could lead to certain instances or criteria to evaluate. This area was not explored in this research and can be further investigated.
- A number of limitations have been noted in Chapter 5. First, it was observed that in the data pre-processing phase, the transformation of many-valued column to single-valued column was conducted manually. With a high frequency of data, the transformation could be performed automatically to save time and effort. This can be addressed by applying a set of pre-defined rules and generating a programming algorithm. Investigating the effectiveness of the algorithm on the transformation process is therefore an important area for a research. Second, the concepts were explored to generate class properties using concept lattice hierarchy notion by Wille (1992). The

concept lattice generated by ConExp tool was explored with minimal abstraction. A bigger context may lead to a bigger abstraction of the domain, thus more process relations can be extracted.

- Chapter 6 also noted several areas in which the approach may be improved. First, the analysis of the structural layer of MVSimO revealed that there were few elements in Cumberland model, that MVSimO has missed and also few concepts in Cumberland that can be improved by the MVSimO elements. This area could be investigated more. Secondly, identification of new scopes and scenarios for the intended model may lead to further refinement of MVSimO, which is also known as the Ontological Improvements (Sabou, 2005). Thirdly, learned outcomes from Chapter 4 and 5 to produce MVSimO has identified the need to extend it to adopt other processes in A&E domain and opens a possibility of identifying simulation components through clarifying the ontic commitment in the ESDF. This area was not explored in this research and can be further investigated. Finally, further development of the first-cut of MVSimO model using ontology software like Protégé could be performed to create more details and to be used effectively.

- This research was conducted manually due to time constraint, whereby automated tool for future research could open more possibilities for the research development in the area of ontology for simulation. The research aim to design an effective domain knowledge extraction method that supports the design and development the ontology for simulation using existing knowledge when access to expert's opinions are limited. Through out this research, the processes of design, development and evaluation were conducted by the same person, hence the evaluation of the final ontology, MVSimO, by another person could be useful. By following the blueprint (Figure 7.1), the evaluation can eliminate an inherent bias as other's knowledge about this research is limited.

Bibliography

Abinaya, V. Kumar, and Swathika

2015. Ontology based public healthcare system in internet of things (IoT). *Procedia Computer Science*, 50:99–102.

Akbari Haghighinejad, H., E. Kharazmi, N. Hatam, S. Yousefi, S. Ali Hesami, M. Danaei, M. Askarian, and A. H. Haghighinejad

2016. Using Queuing Theory and Simulation Modelling to Reduce Waiting Times in An Iranian Emergency Department. *IJCBNM January*, 44(11):11–26.

Aminuddin, W. M. W. M. and W. R. Ismail

2016. Integrated simulation and data envelopment analysis models in emergency department. In *AIP Conference Proceedings*, volume 1784, P. 050003.

Amith, M. F., Z. He, J. Bian, J. Antonio Lossio-Ventura, and C. Tao

2018. Assessing the Practice of Biomedical Ontology Evaluation: Gaps and Opportunities. *Journal of Biomedical Informatics*, 80(February):1–13.

Anagnostou, A., A. Nouman, and S. J. Taylor

2013. Distributed Hybrid Agent-based Discrete Event Emergency Medical Services Simulation. In *Proceedings of the 2013 Winter Simulation Conference - Simulation: Making Decisions in a Complex World, WSC 2013*, Pp. 3181–3190.

Andrews, S. and C. Orphanides

2010. Analysis of large data sets using formal concept lattices. *CEUR Workshop Proceedings*, 672:104–115.

Anutariya, C., R. Ungrangsi, and V. Wuwongse
2007. Sqore: a framework for semantic query based ontology retrieval. In *International Conference on Database Systems for Advanced Applications*, Pp. 924–929. Springer.

Arcelli, F. and L. Christina
2007. Enhancing software evolution through design pattern detection. In *Third International IEEE Workshop on Software Evolvability 2007*, Pp. 7–14. IEEE.

Aringhieri, R., M. E. Bruni, S. Khodaparasti, and J. T. van Essen
2017. Emergency medical services and beyond: Addressing new challenges through a wide literature review. *Computers and Operations Research*, 78(August 2016):349–368.

Aznag, M., M. Quafafou, and Z. Jarir
2014. Leveraging formal concept analysis with topic correlation for service clustering and discovery. *Proceedings - 2014 IEEE International Conference on Web Services, ICWS 2014*, Pp. 153–160.

Baboolal, K., J. D. Griffiths, V. A. Knight, A. V. Nelson, C. Voake, and J. E. Williams
2012. How efficient can an emergency unit be? A perfect world model. *Emergency Medicine Journal*, 29(12):972–977.

Baclawski, K. and T. Schneider
2009. The open ontology repository initiative: Requirements and research challenges. *CEUR Workshop Proceedings*, 514:1–10.

Baker, C.
2015. Accident and Emergency Statistics. *House of Commons Library Briefing Paper*, Pp. 1–34.

Balasubramaniam, K.

2015. Hybrid fuzzy-ontology design using FCA based clustering for information retrieval in semantic web. *Procedia Computer Science*, 50:135–142.

Baskerville, R. L. and M. D. Myers
2002. Information systems as a reference discipline. *Mis Quarterly*, Pp. 1–14.

Batres, R.
2017. Ontologies in Process Systems Engineering. *Chemie Ingenieur Technik*, 89(11):1421–1431.

Bau, C.-T., R.-C. Chen, and C.-Y. Huang
2014. Construction of a Clinical Decision Support System for Undergoing Surgery Based on Domain Ontology and Rules Reasoning. *Telemedicine and e-Health*, 20(5):460–472.

Beck, H., K. Morgan, Y. Jung, S. Grunwald, H.-y. Kwon, and J. Wu
2010. Ontology-based simulation in agricultural systems modeling. *Agricultural Systems*, 103(7):463–477.

Becker, P., J. Hereth, and G. Stumme
2002. ToscanaJ – An Open Source Tool for Qualitative Data Analysis. *Advances in Formal Concept Analysis for Knowledge Discovery in Databases. Proc. Workshop FCAKDD of the 15th European Conference on Artificial Intelligence (ECAI 2002)*, Pp. 1–2.

Bedoya-Valencia, L. and E. Kirac
2016. Evaluating alternative resource allocation in an emergency department using discrete event simulation. *Simulation*, 92(12):1041–1051.

Bell, D., C. Cordeaux, T. Stephenson, H. Dawe, P. Lacey, and L. O'Leary
2017. Designing effective hybridization for whole system modeling and simulation in healthcare. *Proceedings - Winter Simulation Conference*, Pp. 1511–1522.

Bell, D., S. De Cesare, N. Iacovelli, M. Lycett, and A. Merico
2007. A framework for deriving semantic web services. *Information Systems Frontiers*, 9(1):69–84.

Bell, D., A. Kashefi, N. Saleh, T. Turchi, and T. Young

2016. A data-driven agent based simulation platform for early health economics device evaluation. In *Proceedings of the Agent-Directed Simulation Symposium*, P. 5. Society for Computer Simulation International.

Benjamin, P. and M. Graul

2006. A framework for adaptive modeling and ontology-driven simulation (famos). In *Enabling Technologies for Simulation Science X*, volume 6227, P. 622705. International Society for Optics and Photonics.

Benjamin, P., M. Patki, and R. Mayer

2006. Using Ontologies for Simulation Modeling. *Constraints*, Pp. 1151–1159.

Berners-Lee, T., J. Hendler, O. Lassila, et al.

2001. The semantic web. *Scientific american*, 284(5):28–37.

Blunt, I., N. Edwards, and L. Merry

2015. What ' S Behind the A & E ' Crisis '?

Bonabeau, E.

2002. Agent-based modeling: Methods and techniques for simulating human systems. *Proceedings of the National Academy of Sciences*, 99(suppl 3):7280–7287.

Borst, W. N.

1997. *Construction of engineering ontologies for knowledge sharing and reuse.* PhD thesis, Faculty of Electrical Engineering, Mathematics & Computer Science.

Borza, P. V., O. Sabou, and C. Sacarea

2010. OpenFCA, an open source formal concept analysis toolbox. *2010 IEEE International Conference on Automation, Quality and Testing, Robotics (AQTR)*, Pp. 1–5.

Brown, A. L.

1992. Design experiments: Theoretical and methodological challenges in creating complex interventions in classroom settings. *The journal of the learning sciences*, 2(2):141–178.

Buitelaar, P., T. Eigner, and T. Declerck

2004. Ontoselect: A dynamic ontology library with support for ontology selection. In *In Proceedings of the Demo Session at the International Semantic Web Conference*. Citeseer.

Cantador, I., M. Fernández, and P. Castells

2007. Improving ontology recommendation and reuse in WebCORE by collaborative assessments. *CEUR Workshop Proceedings*, 273.

Carbonnel, J., K. Bertet, M. Huchard, and C. Nebut

2016. FCA for software product lines representation: Mixing configuration and feature relationships in a unique canonical representation. *CEUR Workshop Proceedings*, 1624:109–122.

Castellanos, A., J. Cigarran, and A. Garcia-Serrano

2017. Formal concept analysis for topic detection: A clustering quality experimental analysis. *Information Systems*, 66:24–42.

Chahal, K., T. Eldabi, and T. Young

2013. A conceptual framework for hybrid system dynamics and discrete event simulation for healthcare. *Journal of Enterprise Information Management*, 26(1/2):50–74.

Chandrasegaran, S. K., K. Ramani, R. D. Sriram, I. Horváth, A. Bernard, R. F. Harik, and W. Gao

2013. The evolution, challenges, and future of knowledge representation in product design systems. *Computer-Aided Design*, 45(2):204–228.

Charles Turnitsa, A. T.

2006. Ontology applied - techniques employing ontological representation for m&s. In *Proceedings of the 2006 Fall Simulation*.

Cigarran, J., A. Castellanos, and A. Garcia-Serrano

2016. A step forward for Topic Detection in Twitter: An FCA-based approach. *Expert Systems with Applications*, 57:21–36.

Cimellaro, G. P., M. Malavisi, and S. Mahin

2017. Using Discrete Event Simulation Models to Evaluate Resilience of an Emergency Department. *Journal of Earthquake Engineering*, 21(2):203–226.

Cimellaro, G. P. and M. Pique

2016. Resilience of a hospital emergency department under seismic event. *Advances in Structural Engineering*, 19(5):825–836.

Cimiano, P., A. Hotho, G. Stumme, and J. Tane

2004. Conceptual Knowledge Processing with Formal Concept Analysis and Ontologies. In *Lecture Notes in Computer Science*, volume 2961, Pp. 189–207.

Collarana, D., C. Lange, and S. Auer

2016. Fuhsen: a platform for federated, rdf-based hybrid search. In *Proceedings of the 25th International Conference Companion on World Wide Web*, Pp. 171–174. International World Wide Web Conferences Steering Committee.

Corcho, O., A. Gómez-Pérez, and M. Fernández-López

2004. Ontological engineering. *With examples from the areas of Knowledge Management, e-Commerce and the Semantic Web (Advanced Information and Knowledge Processing)*.

Coughlan, J., J. Eatock, and N. Patel

2011. Simulating the use of re-prioritisation as a wait-reduction strategy in an emergency department. *Emergency medicine journal : EMJ*, 28(12):1013–8.

Daga, A., S. De Cesare, M. Lycett, and C. Partridge

2005. An ontological approach for recovering legacy business content. In *Proceedings of the 38th Annual Hawaii International Conference on System Sciences*, Pp. 224a–224a. IEEE.

Dakin, H., N. Devlin, Y. Feng, N. Rice, P. O'Neill, and D. Parkin

2015. The influence of cost-effectiveness and other factors on nice decisions. *Health Economics (United Kingdom)*, 24(10):1256–1271.

D'Aquin, M., C. Baldassarre, L. Gridinoc, M. Sabou, S. Angeletou, and E. Motta
2007. Watson: supporting next generation semantic web applications. *Knowledge Creation Diffusion Utilization.*

Daraio, C., M. Lenzerini, C. Leporelli, P. Naggar, A. Bonaccorsi, and A. Bartolucci
2016. The advantages of an Ontology-Based Data Management approach: openness, interoperability and data quality. *Scientometrics*, 108(1):441–455.

Davey, B. A. and H. A. Priestley
2002. *Introduction to lattices and order.* Cambridge university press.

David Bell, Armin Kashefi, N. S. T. T. T. Y.
2016. A data-driven agent based simulation platform for early health economics device evaluation. In *Proceedings of the Agent-Directed Simulation Symposium*. ADS '16.

De Cesare, S., D. Juric, and M. Lycett
2014. Toward the automation of business process ontology generation. *Proceedings - 16th IEEE Conference on Business Informatics, CBI 2014*, 1:70–77.

Di Somma, S., L. Paladino, L. Vaughan, I. Lalle, L. Magrini, and M. Magnanti
2015. Overcrowding in emergency department: an international issue. *Internal and Emergency Medicine*, 10(2):171–175.

Ding, L., T. Finin, P. Reddivari, R. S. Cost, and J. Sachs
2004. Swoogle : A Search and Metadata Engine for the Semantic Web. *ACM conference on Information and knowledge management*, Pp. 652–659.

Duque-Ramos, A., J. T. Fernández-Breis, M. Iniesta, M. Dumontier, M. Egaña Aranguren, S. Schulz, N. Aussenac-Gilles, and R. Stevens
2013. Evaluation of the OQuaRE framework for ontology quality. *Expert Systems with Applications*, 40(7):2696–2703.

Duque-Ramos, A., J. T. Fernández-Breis, R. Stevens, and N. Aussenac-Gilles

2011. OQuaRE: A square-based approach for evaluating the quality of ontologies. *Journal of Research and Practice in Information Technology*, 43(2):159–176.

Durak, U., H. Oğuztüzün, C. Köksal Algin, and Ö. Özdikiş

2011. Towards interoperable and composable trajectory simulations: an ontology-based approach. *Journal of Simulation*, 5(3):217–229.

Durak, U. and T. Ören

2016. Towards an ontology for simulation systems engineering. In *Proceedings of the 49th Annual Simulation Symposium*.

d'Aquin, M. and N. F. Noy

2012. Where to publish and find ontologies? a survey of ontology libraries. *Journal of Web Semantics*, 11:96–111.

Eatock, J., M. Clarke, C. Picton, and T. Young

2011. Meeting the four-hour deadline in an A&E department. *Journal of Health Organization and Management*, 25(6):606–624.

Edelson, D. C.

2002. Design research: What we learn when we engage in design. *The Journal of the Learning sciences*, 11(1):105–121.

Escudero-Marin, P. and M. Pidd

2011. Using ABMS to simulate Emergency Departments. *Proceedings - Winter Simulation Conference*, Pp. 1239–1250.

Farinelli, F., M. B. Almeidaa, P. Elkin, and B. Smith

2016. OntONeo: The obstetric and neonatal ontology. *CEUR Workshop Proceedings*, 1747.

Feilmayr, C. and W. Wöß

2016. An analysis of ontologies and their success factors for application to business. *Data and Knowledge Engineering*, 101:1–23.

Fernández-López, M., A. Gómez-Pérez, and M. C. Suárez-Figueroa

2013. Methodological guidelines for reusing general ontologies. *Data and Knowledge Engineering*.

Ford, R., M. D. E. D. J. M.

2011. Ontologies and tools for analysing and composing simulation confederations for the training and testing domains. *Journal of Simulation*, 5:230–245.

Ganter, B. and S. Obiedkov

2016. *Conceptual Exploration*.

Garcia, R.

2009. *Semantic Web for Business*. IGI Global.

Ghanes, K., M. Wargon, Z. Jemai, O. Jouini, R. Hellmann, V. Thomas, and G. Koole

2014. a Comprehensive Simulation Modeling of an Emergency Department: a Case Study for Simulation Optimization of Staffing Levels. *Proceedings of the 2014 Winter Simulation Conference*, Pp. 1421–1432.

Gil, M., M. Beron, G. Montejano, D. Riesco, H. Lee, and N. Debnath

2016. An ontology for the interoperability of hydrological simulation model. *Proceedings of the IEEE International Conference on Industrial Technology*, 2016-May:1706–1711.

Giustozzi, F., J. Saunier, and C. Zanni-Merk

2018. Context Modeling for Industry 4.0: An Ontology-Based Proposal. *Procedia Computer Science*, 126:675–684.

Gómez-Pérez, A. and M. D. Rojas-Amaya

1999. Ontological reengineering for reuse. In *International Conference on Knowledge Engineering and Knowledge Management*, Pp. 139–156. Springer.

Goodale, P., P. D. Clough, N. Ford, M. M. Hall, M. Stevenson, S. Fernando, N. Aletras, K. Fernie,

P. Archer, and A. de Polo
2012. User-centred design to support exploration and path creation in cultural heritage collections. In *EuroHCIR*.

Grau, B. C., I. Horrocks, Y. Kazakov, and U. Sattler
2007. Ontology reuse: Better safe than sorry. *CEUR Workshop Proceedings*.

Grigore, B., J. Peters, C. Hyde, and K. Stein
2016. A comparison of two methods for expert elicitation in health technology assessments. *BMC Medical Research Methodology*, 16(1):85.

Grigoriev, P. A. and S. A. Yevtushenko
2004. QuDA: Applying Formal Concept Analysis in a Data Mining Environment. In *Concept Lattices: Second International Conference on Formal Concept Analysis, ICFCA 2004, Sydney, Australia, February 23-26, 2004. Proceedings*, P. Eklund, ed., Pp. 386–393. Berlin, Heidelberg: Springer Berlin Heidelberg.

Grolinger, K., M. A. M. Capretz, J. R. Marti, K. D. Srivastava, K. Grolinger, and M. A. M. Capretz
2012. Ontology – based Representation of Simulation Models Ontology – based Representation of Simulation Models. *The Twenty-Fourth International Conference on Software Engineering and Knowledge Engineering (SEKE)*, Pp. 432–437.

Gruber, T. R.
1993. A translation approach to portable ontology specifications. *Knowledge Acquisition*, 5(2):199–220.

Guizzardi, G.
2005. *Ontological Foundations for Structural Conceptual Model*, volume 015.

Guizzardi, G.
2007. On ontology, ontologies, conceptualizations, modeling languages, and (meta) models. *Frontiers in artificial intelligence and applications*, 155:18.

Guizzardi, G. and G. Wagner

2004. A unified foundational ontology and some applications of it in business modeling. In *CAiSE Workshops (3)*, Pp. 129–143.

Guizzardi, G. and G. Wagner

2010. Towards an ontological foundation of discrete event simulation. *Proceedings - Winter Simulation Conference*, Pp. 652–664.

Guizzardi, G., G. Wagner, J. P. A. Almeida, and R. S. Guizzardi

2015. Towards ontological foundations for conceptual modeling: The unified foundational ontology (ufo) story. *Applied ontology*, 10(3-4):259–271.

Gul, M. and A. F. Guneri

2015. A comprehensive review of emergency department simulation applications for normal and disaster conditions. *Computers and Industrial Engineering*, 83:327–344.

Günal, M. M. and M. Pidd

2009. Understanding target-driven action in emergency department performance using simulation. *Emergency medicine journal : EMJ*, 26:724–727.

Hay, A. M., E. C. Valentin, and R. A. Bijlsma

2006. Modeling emergency care in hospitals: A paradox - The patient should not drive the process. *Proceedings - Winter Simulation Conference*, Pp. 439–445.

Heath, B., R. Hill, and F. Ciarallo

2009. A Survey of Agent-Based Modeling Practices. *Journal of Artificial Societies and Social Simulation*, 12(4 (9)):1–46.

Hevner, A. R., S. T. March, J. Park, and S. Ram

2004. Design Science in Information Systems Research. *MIS Quarterly*, 28(1):75–105.

Hicks, A., J. Hanna, D. Welch, M. Brochhausen, and W. R. Hogan

2016a. The ontology of medically related social entities: recent developments. *Journal of biomedical semantics*, 7(1):47.

Hicks, A., J. Hanna, D. Welch, M. Brochhausen, W. R. Hogan, B. Smith, M. Ashburner, C. Rosse, J. Bard, W. Bug, W. Ceusters, L. Goldberg, K. Eilbeck, A. Ireland, C. Mungall, L. Obrst, P. Whetzel, N. Noy, N. Shah, P. Alexander, C. Nyulas, T. Tudorache, M. Musen, S. Federhen, M. Almeida, L. Slaughter, M. Brochhausen, B. Smith, R. Fleurence, L. Curtis, R. Califf, R. Platt, J. Selby, J. Brown, B. Smith, W. Ceusters, B. Klagges, J. Köhler, A. Kumar, J. Lomax, C. Mungall, F. Neuhaus, A. Rector, and C. Rosse
2016b. The ontology of medically related social entities: recent developments. *Journal of Biomedical Semantics*, 7(1):47.

House, E. R.
2010. *Evaluating with validity*. IAP.

Huang, Y.-L.
2016. Appointment standardization evaluation in a primary care facility. *International Journal of Health Care Quality Assurance*, 29(6):675–686.

Isern, D. and A. Moreno
2016. A Systematic Literature Review of Agents Applied in Healthcare. *Journal of Medical Systems*, 40(2):1–14.

Jahangirian, M., S. J. Taylor, J. Eatock, L. K. Stergioulas, and P. M. Taylor
2014. Causal study of low stakeholder engagement in healthcare simulation projects. *Journal of the Operational Research Society*, 66(3):369–379.

Jiang, G. and C. G. Chute
2009. Auditing the Semantic Completeness of SNOMED CT Using Formal Concept Analysis. *Journal of the American Medical Informatics Association*, 16(1):89–102.

Jiang, X. and A.-H. Tan
2006. Ontosearch: A full-text search engine for the semantic web. In *AAAI*, volume 6, Pp. 1325–1330.

Kaiya, H. and M. Saeki

2006. Using domain ontology as domain knowledge for requirements elicitation. In *14th IEEE International Requirements Engineering Conference (RE'06)*, Pp. 189–198. IEEE.

Kamdar, M. R., T. Tudorache, and M. A. Musen

2017. A systematic analysis of term reuse and term overlap across biomedical ontologies. *Semantic Web*, 8(6):853–871.

Kelly, A. E. and R. A. Lesh

2012. *Handbook of research design in mathematics and science education.* Routledge.

Kester, Q.-a.

2016. Application of Formal Concept Analysis to Visualization of the Evaluation of Risks Matrix in Software Engineering Projects. *International Journal of Science Engineering and Technology Research*, 2(January 2013).

Kettouch, M. S., C. Luca, M. Hobbs, and A. Fatima

2015. Data integration approach for semi-structured and structured data (linked data). In *2015 IEEE 13th international conference on industrial informatics (INDIN)*, Pp. 820–825. IEEE.

Khan, Z. and C. M. Keet

2012. Onset: Automated foundational ontology selection and explanation. In *International Conference on Knowledge Engineering and Knowledge Management*, Pp. 237–251. Springer.

Khanna, S., D. Sier, J. Boyle, and K. Zeitz

2016. Discharge timeliness and its impact on hospital crowding and emergency department flow performance. *EMA - Emergency Medicine Australasia*, 28(2):164–170.

Komashie, A. and A. Mousavi

2005. Modeling emergency departments using discrete event simulation techniques. *Proceedings - Winter Simulation Conference*, 2005:2681–2685.

Kontopoulos, E., C. Berberidis, T. Dergiades, and N. Bassiliades

2013. Ontology-based sentiment analysis of twitter posts. *Expert Systems with Applications*, 40(10):4065–4074.

Kuznetsov, S. O. and S. Obiedkov

2008. Some decision and counting problems of the Duquenne-Guigues basis of implications. *Discrete Applied Mathematics*, 156(11):1994–2003.

Kuznetsov, S. O., S. A. Obiedkov, and C. Roth

2007. Reducing the Representation Complexity of Lattice-Based Taxonomies. *Conceptual Structures: Knowledge Architectures for Smart Applications*, 4604:241–254.

Kuznetsov, S. O. and J. Poelmans

2013a. Knowledge representation and processing with formal concept analysis. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 3(3):200–215.

Kuznetsov, S. O. and J. Poelmans

2013b. Knowledge representation and processing with formal concept analysis. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 3(3):200–215.

La-Ongsri, S. and J. F. Roddick

2015. Incorporating ontology-based semantics into conceptual modelling. *Information Systems*, 52:1–20.

Labrousse, M. and A. Bernard

2008. Fbs-ppre, an enterprise knowledge lifecycle model. In *Methods and tools for effective knowledge life-cycle-management*, Pp. 285–305. Springer.

Lebcir, R., E. Demir, R. Ahmad, C. Vasilakis, and D. Southern

2017. A discrete event simulation model to evaluate the use of community services in the treatment of patients with Parkinson's disease in the United Kingdom. *BMC Health Serv Res*, 17(1):50.

Lee, S. K., K. R. Kim, and J. H. Yu

2014. BIM and ontology-based approach for building cost estimation. *Automation in Construction*, 41:96–105.

Lee, W., C. Blais, and L. W. Lacy

2004. Semantic Web : Implications for Modeling and Simulation System Interoperability.

Lee, Y., S. Jana, T. Mylavarapu, D. Dinakarpandian, and D. Owens

2011. MindFlow: Intelligent workflow for clinical trials in mental healthcare. *Proceedings of the Annual Hawaii International Conference on System Sciences*, Pp. 2779–2788.

Leidig, J., E. A. Fox, K. Hall, M. Marathe, and H. Mortveit

2011. Simdl: a model ontology driven digital library for simulation systems. In *Proceedings of the 11th annual international ACM/IEEE joint conference on Digital libraries*, Pp. 81–84. ACM.

Li, H., J. Wu, L. Liu, and Q. Li

2015. Adoption of Big Data Analytics in Healthcare : The Efficiency and Privacy. *PACIS 2015 Proceedings*.

Lincke, R. and W. Löwe

2007. Compendium of software quality standards and metrics - version 1.0. `http://www.arisa.se/compendium/`. Accessed: 2019-03-30.

Liu, Z., D. Rexachs, F. Epelde, and E. Luque

2017. A simulation and optimization based method for calibrating agent-based emergency department models under data scarcity. *Computers and Industrial Engineering*, 103:300–309.

Lonsdale, D., D. W. Embley, Y. Ding, L. Xu, and M. Hepp

2010. Reusing ontologies and language components for ontology generation. *Data & Knowledge Engineering*, 69(4):318–330.

Luppa, P. B., C. Müller, A. Schlichtiger, and H. Schlebusch
2011. Point-of-care testing (POCT): Current techniques and future perspectives. *TrAC - Trends in Analytical Chemistry*, 30(6):887–898.

Malone, J., R. Stevens, S. Jupp, T. Hancocks, H. Parkinson, and C. Brooksbank
2016. Ten Simple Rules for Selecting a Bio-ontology. *PLoS Computational Biology*, 12(2):1–6.

Manley, K., A. Martin, C. Jackson, and T. Wright
2016. Using systems thinking to identify workforce enablers for a whole systems approach to urgent and emergency care delivery: a multiple case study. *BMC Health Services Research*, 16(1):368.

March, S. T. and D. Smith
1995. Design and natural science research on information technology. *Decision Support Systems*, 15(4):251–266.

March, S. T. and V. C. Storey
2008. Design science in the information systems discipline: An introduction to the special issue on design science research. *MIS Q.*, 32(4):725–730.

Marshall, D. A., L. Burgos-Liz, K. S. Pasupathy, W. V. Padula, M. J. IJzerman, P. K. Wong, M. K. Higashi, J. Engbers, S. Wiebe, W. Crown, and N. D. Osgood
2016. Transforming Healthcare Delivery: Integrating Dynamic Simulation Modelling and Big Data in Health Economics and Outcomes Research. *PharmacoEconomics*, 34(2):115–126.

Maull, R. and P. Smart
2009. An evaluation of 'fast track'in A&E: a discrete event simulation approach. *The Service Industries*, 44(06).

Mielczarek, B. and J. Uzialko-Mydlikowska
2012. Application of computer simulation modeling in the health care sector: a survey. *Simulation*, 88(2):197–216.

Minyaoui, A. and F. Gargouri

2012. Iterative ontology selection guided by user for building domain ontologies. In *Twenty-Fifth International FLAIRS Conference*.

Mizoguchi, R.

2010. Yamato: yet another more advanced top-level ontology. In *Proceedings of the Sixth Australasian Ontology Workshop*, Pp. 1–16.

Mizoguchi, R., J. Vanwelkenhuysen, and M. Ikeda

1995. Task ontology for reuse of problem solving knowledge. *Towards Very Large Knowledge Bases: Knowledge Building & Knowledge Sharing*, 46(59):45.

Mohiuddin, S., J. Busby, J. Savovic, A. Richards, K. Northstone, W. Hollingworth, J. L. Donovan, and C. Vasilakis

2017a. Patient flow within UK emergency departments: a systematic review of the use of computer simulation modelling methods. *BMJ Open*, 7(5).

Mohiuddin, S., J. Busby, J. Savović, A. Richards, K. Northstone, W. Hollingworth, J. L. Donovan, and C. Vasilakis

2017b. Patient flow within UK emergency departments: a systematic review of the use of computer simulation modelling methods. *BMJ Open*, 7(5):e015007.

Moreira, J., P. Sales, J. Guerson, F. B. B. Braga, F. Brasileiro, and V. Sobral

2016. Menthor Editor : an ontology - driven conceptual modeling platform.

Mould, G., J. Bowers, C. Dewar, and E. Mcgugan

2013. Assessing the impact of systems modeling in the redesign of an Emergency Department. *Health Systems*, 2(1):3–10.

Newell, A. and H. A. Simon

1975. Computer science as empirical inquiry: Symbols and search. *PHILOSOPHY OF PSYCHOLOGY*, P. 407.

Neznanov, A. A. and S. O. Kuznetsov

2013. Information retrieval and knowledge discovery with FCART. *CEUR Workshop Proceedings*, 977:74–82.

Nicola Guarino, C. W.

1998. Formal ontology in information systems. In *Proceedings of FOIS98*. IOS Press.

Nicola Guarino, D. O. and S. Staab

2009. What is an ontology? *Handbook on Ontologies, International Handbooks on Information Systems*, Pp. 1–17.

Noy, N. F., D. L. McGuinness, et al.

2001. Ontology development 101: A guide to creating your first ontology.

Nunamaker Jr, J. F., M. Chen, and T. D. Purdin

1990. Systems development in information systems research. *Journal of management information systems*, 7(3):89–106.

Obitko, M., V. Snášel, and J. Smid

2004. Ontology Design With Formal Concept Analysis. *Proceedings of the International Workshop on Concept Lattices and their Applications (CLA 2004)*, Pp. 111–119.

Okhmatovskaia, A., D. L. Buckeridge, A. Shaban-Nejad, A. Sutcliffe, J. A. Kopec, P. Fines, and M. C. Wolfson

2012. Simpho: An ontology for simulation modeling of population health. In *Proceedings of the 2012 Winter Simulation Conference (WSC)*, Pp. 1–12. IEEE.

Okoye, K., S. Islam, and U. Naeem

2018. Ontology: core process mining and querying enabling tool. In *Ontology in Information Science*. IntechOpen.

Orlikowski, W. J. and S. Iacono

2001. Desperately seeking the 'IT'in IT research: A call to theorizing the IT artifact. *Information Systems Research*, 12(2):121–134.

Owen, C. L.

1998. Design research: Building the knowledge base. *Design Studies*, 19(1):9–20.

Park, J., S. Oh, and J. Ahn

2011. Ontology selection ranking model for knowledge reuse. *Expert Systems with Applications*.

Partridge, C.

1996. *Business objects: re-engineering for re-use*. Butterworth-Heinemann.

Peñalver-Martinez, I., F. Garcia-Sanchez, R. Valencia-Garcia, M. Á. Rodríguez-García, V. Moreno, A. Fraga, and J. L. Sánchez-Cervantes

2014. Feature-based opinion mining through ontologies. *Expert Systems with Applications*, 41(13):5995–6008.

Pidd, M.

2014. The ways forward : A personal view of system dynamics and discrete-event simulation. *Discrete-Event Simulation and System Dynamics for Management Decision Making*, Pp. 318–336.

Pinto, H.S., G.-P. A. M. J.

1999a. Some issues on ontology integration. In *CEUR Workshop Proceedings*. IJCAI'99.

Pinto, H. S.

1999b. Towards ontology reuse. In *Proceedings of AAAI99's Workshop on Ontology Management, WS-99*, volume 13, Pp. 67–73.

Pinto, H. S. and J. P. Martins

2000. Reusing Ontologies. *Proceedings of the AAAI 2000 Spring Symposium on Bringing Knowledge to Business Processes*, Pp. 77–84.

Pinto, H. S. and J. P. Martins

2001. A methodology for ontology integration. *Proceedings of the international conference on Knowledge capture - K-CAP 2001*, P. 131.

Poelmans, J., D. I. Ignatov, S. O. Kuznetsov, G. Dedene, D. I. Ignatov, and G. Dedene
2013. Formal concept analysis in knowledge processing: A survey on models and techniques. *Expert Systems with Applications*, 40(16):6601–6623.

Qu, Y. and G. Cheng
2011. Falcons concept search: A practical search engine for web ontologies. *IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans*, 41(4):810–816.

Ries, E.
2009. Minimum Viable Product: a guide. *Startup Lessons Learned*.

Robinson, S.
2006. Conceptual modeling for simulation: Issues and research requirements. *Proceedings - Winter Simulation Conference*, Pp. 792–800.

Robinson, S.
2008a. Conceptual modelling for simulation Part I: Definition and requirements. *Journal of the Operational Research Society*, 59(3):278–290.

Robinson, S.
2008b. Conceptual modelling for simulation Part II: A framework for conceptual modelling. *Journal of the Operational Research Society*, 59(3):291–304.

Robinson, S.
2013. Conceptual modeling for simulation. *2013 Winter Simulations Conference (WSC)*, Pp. 377–388.

Roongsangjan, S., T. Sunetnanta, and P. Mongkolwat
2017. Using FCA Implication to Determine the Compliance of Model Practice Implementation for Software Process. *Proceedings of the 2017 International Conference on Management Engineering, Software Engineering and Service Sciences - ICMSS '17*, Pp. 64–70.

Sabou, M.

2005. Learning web service ontologies: an automatic extraction method and its evaluation. *Ontology learning from text: methods, evaluation and applications*, 123(0):125.

Saghafian, S., G. Austin, and S. Traub

2015. Operations Research Contributions to Emergency Department Patient Flow Optimization: Review and Research Prospects. *Available at SSRN 2420163*, Pp. 1–36.

Sarno, R. and F. P. Sinaga

2016. Business process anomaly detection using ontology-based process modelling and Multi-Level Class Association Rule Learning. *Proceeding - 2015 International Conference on Computer, Control, Informatics and Its Applications: Emerging Trends in the Era of Internet of Things, IC3INA 2015*, Pp. 12–17.

Schulz, S., N. Grewe, J. Röhl, D. Schober, M. Boeker, and L. Jansen

2012. Guideline on Developing Good Ontologies in the Biomedical Domain with Description Logics.

Scriven, M.

1991. *Evaluation thesaurus*. Sage.

Sernadela, P., A. Pereira, and R. Rossetti

2015a. DISim: Ontology-driven simulation of biomedical data integration tasks. *2015 10th Iberian Conference on Information Systems and Technologies, CISTI 2015*.

Sernadela, P., A. Pereira, and R. Rossetti

2015b. DISim: Ontology-driven simulation of biomedical data integration tasks. *2015 10th Iberian Conference on Information Systems and Technologies, CISTI 2015*.

Shadbolt, N., T. Berners-Lee, and W. Hall

2006. The semantic web revisited. *IEEE Intelligent Systems*, 21(3):96–101.

Shafee Kalid, K.

2017. The Investigation of Knowledge Management Practices Among Healthcare Practitioners in Malaysia The Investigation of Knowledge Management Practices Among Healthcare Practitioners in Malaysia.

Sider, T. et al.
2001. *Four-dimensionalism: An ontology of persistence and time.* Oxford University Press on Demand.

Silver, G. A., O. A.-H. Hassan, and J. A. Miller
2007. From domain ontologies to modeling ontologies to executable simulation models. In *Proceedings of the 39th conference on Winter simulation: 40 years! The best is yet to come*, Pp. 1108–1117. IEEE Press.

Silver, G. A., J. A. Miller, M. Hybinette, G. Baramidze, and W. S. York
2011. DeMO: An Ontology for Discrete-event Modeling and Simulation. *Simulation*, 87(9):747–773.

Simon, H. A.
1996. *The sciences of the artificial.* MIT press.

Simperl, E.
2009. Reusing ontologies on the Semantic Web: A feasibility study. *Data & Knowledge Engineering*, 68(10):905–925.

SIMPERL, E.
2010. Guidelines for Reusing Ontologies on the Semantic Web. *International Journal of Semantic Computing*, 04(02):239–283.

Sklar, D. P., D. A. Handel, J. Hoekstra, J. M. Baren, B. Zink, and J. R. Hedges
2010. The Future of Emergency Medicine: An Evolutionary Perspective. *Academic Medicine*, 85(3):490–495.

Stachowiak, H.
1973. *Allgemeine modelltheorie [General Model Theory].* Springer.

Stecher, R., C. Niederée, W. Nejdl, and P. Bouquet

2008. Adaptive ontology re-use: Finding and re-using sub-ontologies. *International Journal of Web Information Systems*, 4(2):198–214.

Stevens, R. and P. Lord

2009. Application of ontologies in bioinformatics. In *Handbook on ontologies*, Pp. 735–756. Springer.

Steward, D., T. F. Glass, and Y. B. Ferrand

2017. Simulation-Based Design of ED Operations with Care Streams to Optimize Care Delivery and Reduce Length of Stay in the Emergency Department. *Journal of Medical Systems*, 41(10).

Stroetmann, B. and A. Aisenbrey

2012. Medical Knowledge Management in Healthcare Industry. *World Academy of Science, Engineering and Technology*, 6(4):507–512.

Studer, R., V. R. Benjamins, and D. Fensel

1998. Knowledge engineering: principles and methods. *Data & knowledge engineering*, 25(1-2):161–197.

Stumme, G. and A. Maedche

2001. FCA-MERGE: Bottom-up merging of ontologies. *IJCAI International Joint Conference on Artificial Intelligence*, Pp. 225–230.

Sure, Y., S. Staab, and R. Studer

2002. Methodology for development and employment of ontology based knowledge management applications. *Sigmod Record*, 31(4):18–23.

Takeda, H., P. Veerkamp, and H. Yoshikawa

1990. Modeling design process. *AI magazine*, 11(4):37–37.

Tao, M., K. Ota, and M. Dong

2017. Ontology-based data semantic management and application in IoT- and cloud-enabled smart homes. *Future Generation Computer Systems*, 76:528–539.

Thomas, O. and M. Fellmann M.A.

2009. Semantic Process Modeling – Design and Implementation of an Ontology-based Representation of Business Processes. *Business & Information Systems Engineering*, 1(6):438–451.

Tolk, A., S. Y. Diallo, I. O. Ryzhov, L. Yilmaz, S. Buckley, and J. A. Miller

2014. Tutorial: information and process modeling for simulation. *Proceedings of the 2014 Winter Simulation Conference - , WSC 2014*, Pp. 103–117.

Tolk, A., M. D. Haase, J. Bricio, A. Tolk, and M. D. Haase

2015. Ontological Support to address the Multi-Dimensionality of Complex Systems Engineering Challenges. *ModSim World 2015*, Pp. 1–11.

Tolk, A. and J. A. Miller

2011. Enhancing simulation composability and interoperability using conceptual/semantic/ontological models. *Journal of Simulation*, 5(3):133–134.

Traoré, M. K., G. Zacharewicz, R. Duboz, and B. Zeigler

2018. Modeling and simulation framework for value-based healthcare systems. *Simulation*.

Trokanas, N. and F. Cecelja

2016. Ontology evaluation for reuse in the domain of Process Systems Engineering. *Computers and Chemical Engineering*, 85:177–187.

Uschold, M. and R. Jasper

1999. A framework for understanding and classifying ontology applications. In *Proceedings of the IJCAI-99 Workshop on Ontologies and Problem-Solving Methods (KRR5), Stockholm, Sweden*, volume 2.

Vaishnavi, V. and B. Kuechler

2004. Design Science Research in Information Systems. *Ais*, P. 45.

Van Heijst, G., A. T. Schreiber, and B. J. Wielinga

1997. Using explicit ontologies in kbs development. *International journal of human-computer studies*, 46(2-3):183–292.

Vannieuwenborg, F., F. Ongenae, P. Demyttenaere, L. Van Poucke, J. Van Ooteghem, S. Verstichel, S. Verbrugge, D. Colle, F. De Turck, and M. Pickavet

2015. Techno-economic evaluation of an ontology-based nurse call system via discrete event simulations.

Varga, A.

2001. Discrete event simulation system. In *In Proceedings of the European Simulation Multiconference*. ESM2001.

Venable, J., J. Pries-Heje, and R. Baskerville

2016. FEDS: A Framework for Evaluation in Design Science Research. *European Journal of Information Systems*, 25(1):77–89.

Verdonck, M., F. Gailly, S. De Cesare, and G. Poels

2015. Ontology-driven conceptual modeling: A'systematic literature mapping and review. *Applied Ontology*, 10(3-4):197–227.

Verdonck, M., F. Gailly, R. Pergl, G. Guizzardi, B. Martins, and O. Pastor

2018. Comparing traditional conceptual modeling with ontology-driven conceptual modeling: An empirical study. *Information Systems*, 81:92–103.

Virtue, A., T. Chaussalet, and J. Kelly

2011. Using Simplified Discrete-Event Simulation Models for Health Care Applications. *Proceedings of the 2011 Winter Simulation Conference*, Pp. 1154–1165.

Weber, E. J., S. Mason, J. V. Freeman, and J. Coster

2012. Implications of England's four-hour target for quality of care and resource use in the emergency department. *Annals of Emergency Medicine*, 60(6):699–706.

Whetzel, P. L., N. F. Noy, N. H. Shah, P. R. Alexander, C. Nyulas, T. Tudorache, and M. A. Musen
2011. BioPortal: Enhanced functionality via new Web services from the National Center for Biomedical Ontology to access and use ontologies in software applications. *Nucleic Acids Research*, 39(SUPPL. 2):541–545.

Wilk, S., W. Michalowski, D. O'Sullivan, K. Farion, J. Sayyad-Shirabad, C. Kuziemsky, and B. Kukawka
2013. A task-based support architecture for developing point-of-care clinical decision support systems for the emergency department. *Methods of Information in Medicine*, 52(1):18–32.

Wille, R.
1982. Restructuring Lattice Theory: An Approach Based on Hierarchies of Concepts. In *Ordered Sets: Proceedings of the NATO Advanced Study Institute held at Banff, Canada, August 28 to September 12, 1981*, I. Rival, ed., Pp. 445–470. Dordrecht: Springer Netherlands.

Wille, R.
1992. Concept lattices and conceptual knowledge systems. *Computers & Mathematics with Applications*, 23(6):493–515.

Winter, R.
2008. Design science research in europe. *European Journal of Information Systems*, 17(5):470–475.

Wray, T. and P. Eklund
2014. Using Formal Concept Analysis to Create Pathways through Museum Collections. *Proceedings of the 3rd International Workshop "what Can Fca Do for Artificial Intelligence"? (fca4ai 2014)*, Pp. 9–16.

Wu, R., C. Chen, S. Ahmad, J. M. Volk, C. Luca, F. C. Harris, and S. M. Dascalu

2016. A real-time web-based wildfire simulation system. In *IECON 2016-42nd Annual Conference of the IEEE Industrial Electronics Society*, Pp. 4964–4969. IEEE.

Xiang, Z., C. Mungall, A. Ruttenberg, and Y. He
2011. Ontobee: A linked data server and browser for ontology terms. *CEUR Workshop Proceedings*, 833:279–281.

Xiao Hang Wang, Da Qing Zhang, Tao Gu, and Hung Keng Pung
2004. Ontology based context modeling and reasoning using OWL. *IEEE Annual Conference on Pervasive Computing and Communications Workshops, 2004. Proceedings of the Second*, Pp. 18–22.

Yan, H., C. Zou, J. Liu, and Z. Wang
2015. Formal concept analysis and concept lattice: Perspectives and challenges. *International Journal of Autonomous and Adaptive Communications Systems*, 8(1):81–96.

Yao, W., C. Chu, A. Kumar, and Z. Li
2009. Using ontology to support context awareness in healthcare. *19th Workshop on Information Technologies and Systems, WITS 2009.*

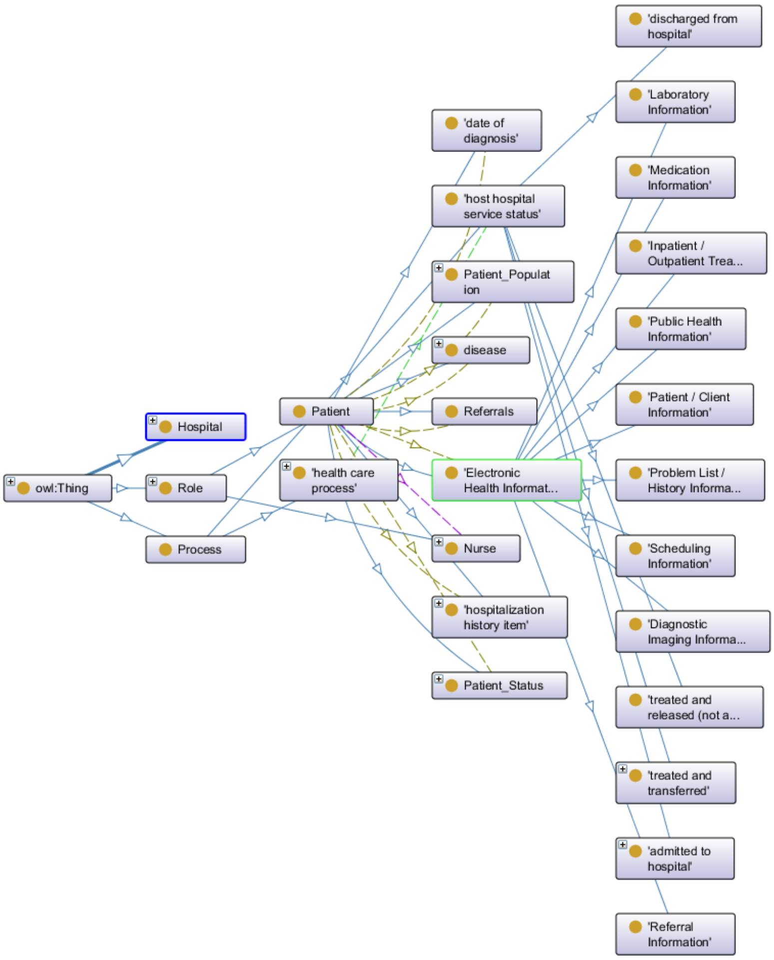
Yevtushenko, S. A.
2000. System of data analysis "Concept Explorer". In *Proceedings of the 7th national conference on Artificial Intelligence KII*, volume 2000.

Zeigler, B. P.
1984. Multifacetted modelling and discrete event simulation. *Journal of the Operational Research Society.*

Zeigler, B. P., T. G. Kim, and H. Praehofer
2000. *Theory of modeling and simulation.* Academic press.

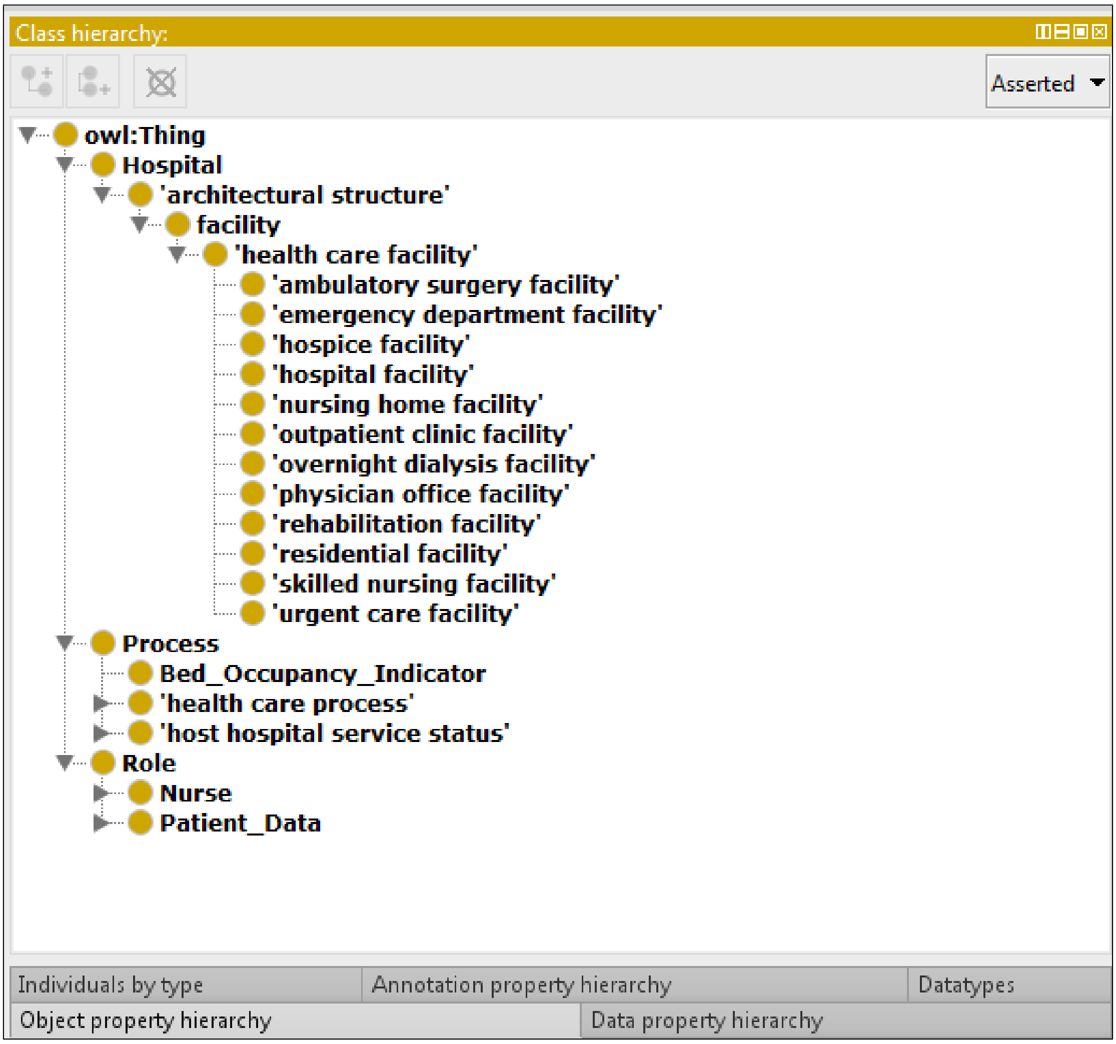
Appendix

A.1 Appendix A - Minimal Domain Ontology (MinDO)



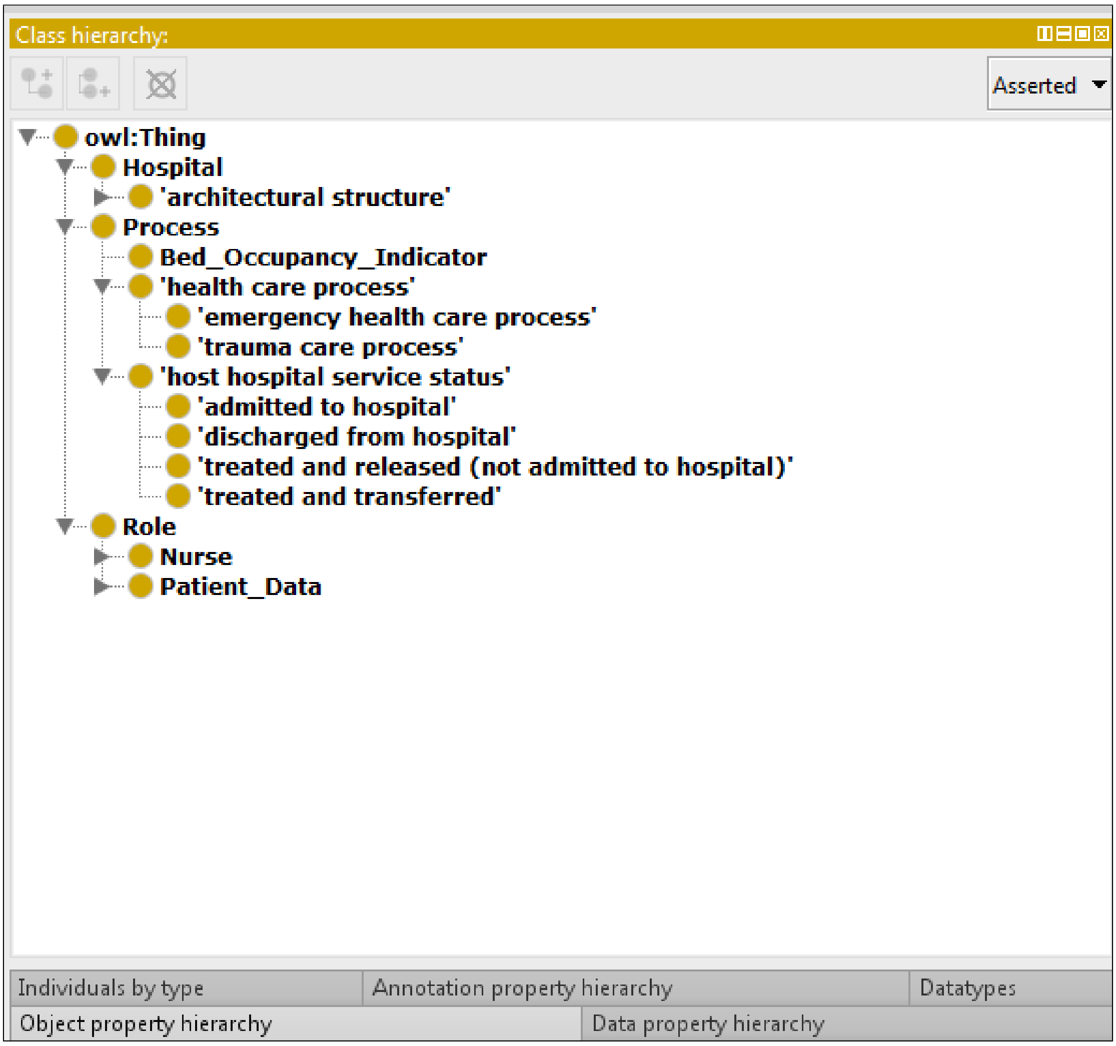
Minimal Domain Ontology

A.2 Appendix A - MinDO Class: Hospital



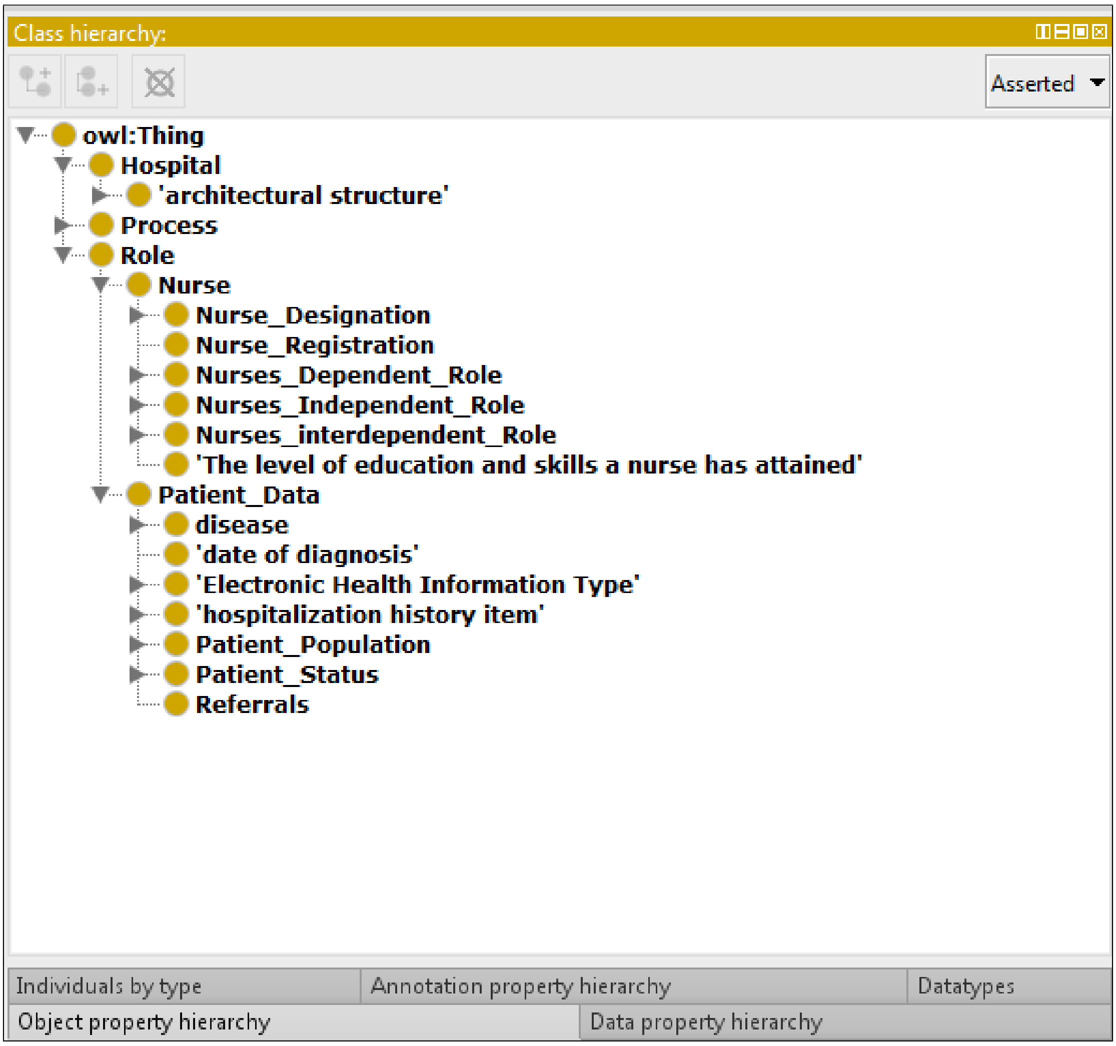
MinDO Class: Hospital

A.3 Appendix A - MinDO Class: Process



MinDO Class: Process

A.4 Appendix A - MinDO Class: Role



MinDO Class: Role

B.1 Appendix B - Normalized Data Element Name (DEN)

ID	Column Name	Normalized Data Element Name (DEN)	Description
1	zNHSNumberPseudo	NHS Number	The NHS NUMBER, the primary identifier of a PERSON, is a unique identifier for a PATIENT within the NHS in England and Wales.
2	ProviderCode	Provider code	ORGANISATION CODE (CODE OF PROVIDER) is the ORGANISATION CODE of the Organisation acting as a Health Care Provider.
3	ProviderName	Provider name	
4	zArrivalDateTime	Arrival date time	
5	zEMAttendanceConclusionDateTime	Attendance conclusion time	A and E ATTENDANCE CONCLUSION TIME is the same as attribute ACTIVITY TIME where the ACTIVITY TIME TYPE is National Code 'Accident and Emergency Attendance Conclusion Time'
6	zEMDateTimeSeenforTreatment	Time seen for treatment	A and E TIME SEEN FOR TREATMENT is the same as attribute ACTIVITY TIME where the ACTIVITY TIME TYPE is National Code 'Accident and Emergency Time Seen For Treatment'.
7	Arrivalmonth	Arrival month	ARRIVAL DATE is the same as attribute ACTIVITY DATE where the ACTIVITY DATE TYPE is National Code 'Arrival Date At Accident and Emergency Department'
8	YearMonth	Year month	ARRIVAL DATE is the same as attribute ACTIVITY DATE where the ACTIVITY DATE TYPE is National Code 'Arrival Date At Accident and Emergency Department'
9	FiscalYearLabel_1	Fiscal year	
10	CoreHRG	Healthcare resource group	ACTIVITY in terms of the types of PATIENTS they care for and the treatments they undertake. They enable the comparison of ACTIVITY within and between different Organisations and provide an opportunity to benchmark treatments and services to support trend analysis
11	hrg_desc	HRG description	
12	EMModeofArrival	Mode of arrival	ACCIDENT AND EMERGENCY ARRIVAL MODE CODE is the same as attribute ACCIDENT AND EMERGENCY ARRIVAL MODE.
13	EMModeofArrivaldesc	Mode of arrival description	ACCIDENT AND EMERGENCY ARRIVAL MODE CODE is the same as attribute ACCIDENT AND EMERGENCY ARRIVAL MODE.
14	EMReferralSource	Referral source	The source of a REFERRAL REQUEST for a Screening Test
15	EMReferralSourcedesc	Referral source description	The source of a REFERRAL REQUEST for a Screening Test
16	A&EDepartmentDescription	A &E Department description	The type of Accident and Emergency Department according to the ACTIVITY performed.
17	GPPracticeCodeGPPracticeName	GP ode GP name	
18	Age	Age	AGE AT ATTENDANCE DATE is derived as the number of completed years between the PERSON BIRTH DATE of the PATIENT and the ATTENDANCE DATE or the estimated age of the PATIENT.
19	CCG	Clinical Commissioning Grou	Clinical Commissioning Groups
20	EMDiagnosisFirst	First diagnosis	ACCIDENT AND EMERGENCY DIAGNOSIS - FIRST is the first recorded PATIENT DIAGNOSIS for an Accident and Emergency Attendance.
21	diag_desc	Diagnosis description	A PATIENT DIAGNOSIS should be classified, where possible, using International Classification of Diseases (ICD) or other classification codes approved centrally for mapping to International Classification of Diseases (ICD) codes.
22	EMAttendanceDisposal	EM attendance disposal	
23	Attendance_disposal	Attendance disposal	
24	PostcodeSectorofUsualAddress	Postcode of usual address	POSTCODE OF USUAL ADDRESS is the POSTCODE of the ADDRESS nominated by the PATIENT where the ADDRESS ASSOCIATION TYPE is 'Main Permanent Residence' or 'Other Permanent Residence'

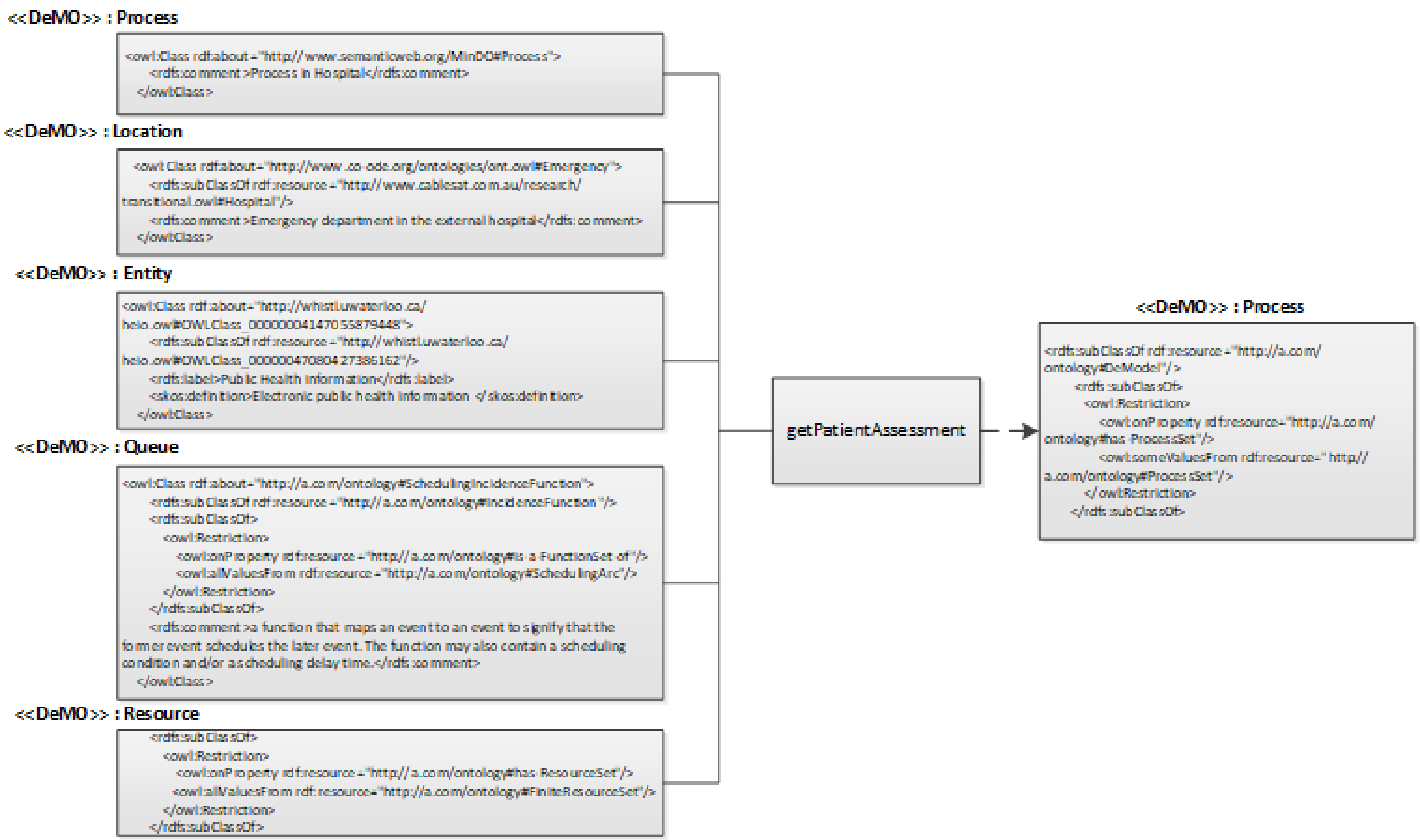
Normalized Data Element Name (DEN)

C.1 Appendix C - A&E Data Instances

zNHSNumberPseudo	zEMAttendanceConclusi	Arrivalmo	YearMont	CoreH	hrg_desc	EMReferralSourcedesc	A&E_DepartmentDescripti	A	Attendance_disposal
86B4-74AB-C17D-4540	27/04/2014 14:53	201404	201401 (Apr)	VB11Z	Emergency Medicine, No Investigation with No Significant Treatment	Self referral	THH UCC	77	Discharged - follow up treatment to be provided by GENERAL PRACTITIONER
FE8B-A764-4ED1-4AA8	04/03/2015 03:38	201503	201412 (Mar)	VB11Z	Emergency Medicine, No Investigation with No Significant Treatment	Self referral	THH UCC	33	Discharged - follow up treatment to be provided by GENERAL PRACTITIONER
F4B4-4517-7222-4F44	05/07/2014 04:10	201407	201404 (Jul)	VB11Z	Emergency Medicine, No Investigation with No Significant Treatment	Self referral	THH UCC	58	Discharged - follow up treatment to be provided by GENERAL PRACTITIONER
B749-DED5-6BBE-43BB	16/12/2014 18:32	201412	201409 (Dec)	VB11Z	Emergency Medicine, No Investigation with No Significant Treatment	Emergency services	THH UCC	32	Discharged - follow up treatment to be provided by GENERAL PRACTITIONER
5D8A-A434-6FE0-48AA	16/03/2015 12:45	201503	201412 (Mar)	VB08Z	Emergency Medicine, category 2 Investigation with category 1 Treatment	GENERAL MEDICAL PRACTITIONER	THH A&E	58	Admitted to a Hospital Bed /became a LODGED PATIENT of the same Health Care Provider
7805-0DDD-DD5B-4771	25/03/2015 16:13	201503	201412 (Mar)	VB09Z	Emergency Medicine, category 1 Investigation with category 1-2 Treatment	Self referral	THH A&E	10	Discharged - did not require any follow up treatment
9B98-4D2D-2B5B-4294	14/12/2014 23:37	201412	201409 (Dec)	VB09Z	Emergency Medicine, category 1 Investigation with category 1-2 Treatment	Health care provider: same or other	THH A&E	4	Discharged - did not require any follow up treatment
7C3A-DFB9-E6EA-4A7E	19/07/2014 02:20	201407	201404 (Jul)	VB09Z	Emergency Medicine, category 1 Investigation with category 1-2 Treatment	Other	THH A&E	80	Admitted to a Hospital Bed /became a LODGED PATIENT of the same Health Care Provider
A09A-AA5C-6603-4182	24/09/2014 08:09	201409	201406 (Sep)	VB11Z	Emergency Medicine, No Investigation with No Significant Treatment	Self referral	THH UCC	27	Discharged - follow up treatment to be provided by GENERAL PRACTITIONER
347C-E92D-0BE6-416D	29/05/2014 22:02	201405	201402 (May)	VB11Z	Emergency Medicine, No Investigation with No Significant Treatment	Self referral	THH UCC	4	Discharged - did not require any follow up treatment
EE27-06F5-D521-4D06	12/05/2014 02:40	201405	201402 (May)	VB11Z	Emergency Medicine, No Investigation with No Significant Treatment	Self referral	THH UCC	1	Discharged - follow up treatment to be provided by GENERAL PRACTITIONER
F0C1-7D92-09E3-4354	26/10/2014 21:26	201410	201407 (Oct)	VB08Z	Emergency Medicine, category 2 Investigation with category 1 Treatment	Self referral	THH A&E	37	Discharged - follow up treatment to be provided by GENERAL PRACTITIONER
31CC-FC44-2A4A-4729	10/12/2014 23:33	201412	201409 (Dec)	VB11Z	Emergency Medicine, No Investigation with No Significant Treatment	Self referral	THH UCC	3	Discharged - follow up treatment to be provided by GENERAL PRACTITIONER
74D3-BE94-39E1-4FC3	20/08/2014 02:10	201408	201405 (Aug)	VB11Z	Emergency Medicine, No Investigation with No Significant Treatment	Self referral	THH UCC	18	Discharged - did not require any follow up treatment
DB19-140C-1DA7-4866	12/01/2015 22:38	201501	201410 (Jan)	VB08Z	Emergency Medicine, category 2 Investigation with category 1 Treatment	GENERAL MEDICAL PRACTITIONER	THH A&E	55	Discharged - follow up treatment to be provided by GENERAL PRACTITIONER
B7A3-FB42-80A1-4308	13/07/2014 23:51	201407	201404 (Jul)	VB05Z	Emergency Medicine, category 2 Investigation with category 3 Treatment	Self referral	THH A&E	49	Referred to other Out-Patient Clinic
9A7E-1B1A-3F7E-401C	02/11/2014 13:24	201411	201408 (Nov)	VB11Z	Emergency Medicine, No Investigation with No Significant Treatment	Self referral	THH Mount Vernon MIU	57	Discharged - did not require any follow up treatment
6DEC-8997-FFD8-4CBF	18/11/2014 17:27	201411	201408 (Nov)	VB09Z	Emergency Medicine, category 1 Investigation with category 1-2 Treatment	Self referral	THH UCC	19	Discharged - did not require any follow up treatment
FF2F-74A0-6D69-4270	03/09/2014 14:30	201409	201406 (Sep)	VB08Z	Emergency Medicine, category 2 Investigation with category 1 Treatment	Self referral	THH A&E	59	Discharged - did not require any follow up treatment
9871-C8F1-FE72-440E	13/01/2015 11:13	201501	201410 (Jan)	VB09Z	Emergency Medicine, category 1 Investigation with category 1-2 Treatment	Self referral	THH Mount Vernon MIU	54	Discharged - did not require any follow up treatment
B082-6686-AD9F-4022	22/09/2014 14:12	201409	201406 (Sep)	VB07Z	Emergency Medicine, category 2 Investigation with category 2 Treatment	Self referral	THH A&E	41	Discharged - did not require any follow up treatment
7D46-67F4-05EA-4F27	30/05/2014 02:13	201405	201402 (May)	VB09Z	Emergency Medicine, category 1 Investigation with category 1-2 Treatment	Other	THH A&E	33	Admitted to a Hospital Bed /became a LODGED PATIENT of the same Health Care Provider
5B5B-A995-7326-4D06	09/02/2015 20:46	201502	201411 (Feb)	VB11Z	Emergency Medicine, No Investigation with No Significant Treatment	Self referral	THH UCC	20	Referred to other Out-Patient Clinic
11AA-D108-25F1-4D57	25/04/2014 23:29	201404	201401 (Apr)	VB08Z	Emergency Medicine, category 2 Investigation with category 1 Treatment	Self referral	THH UCC	45	Referred to Fracture Clinic
CE4E-1183-47A3-4F41	15/12/2014 11:33	201412	201409 (Dec)	VB11Z	Emergency Medicine, No Investigation with No Significant Treatment	Self referral	THH UCC	21	Discharged - did not require any follow up treatment
87FA-068A-6788-45BD	21/01/2015 22:18	201501	201410 (Jan)	VB11Z	Emergency Medicine, No Investigation with No Significant Treatment	Self referral	THH UCC	26	Referred to other Out-Patient Clinic
2AFE-A95D-DB40-4B5E	14/04/2014 19:05	201404	201401 (Apr)	VB07Z	Emergency Medicine, category 2 Investigation with category 2 Treatment	Health care provider: same or other	THH A&E	47	Referred to Fracture Clinic
74C1-F9D5-0A11-4E11	16/01/2015 18:30	201501	201410 (Jan)	VB11Z	Emergency Medicine, No Investigation with No Significant Treatment	Self referral	THH UCC	2	Discharged - follow up treatment to be provided by GENERAL PRACTITIONER
1300-9A0E-B48C-41AB	02/01/2015 01:45	201501	201410 (Jan)	VB08Z	Emergency Medicine, category 2 Investigation with category 1 Treatment	Self referral	THH A&E	87	Admitted to a Hospital Bed /became a LODGED PATIENT of the same Health Care Provider
C7E1-C625-C4DA-42F3	29/12/2014 00:00	201412	201409 (Dec)	VB11Z	Emergency Medicine, No Investigation with No Significant Treatment	Self referral	THH UCC	33	Discharged - did not require any follow up treatment
9AE8-A087-8AD9-49D5	22/12/2014 11:03	201412	201409 (Dec)	VB11Z	Emergency Medicine, No Investigation with No Significant Treatment	Self referral	THH UCC	4	Discharged - did not require any follow up treatment
BC92-E566-A297-4B47	19/08/2014 17:35	201408	201405 (Aug)	VB08Z	Emergency Medicine, category 2 Investigation with category 1 Treatment	Self referral	THH Mount Vernon MIU	20	Discharged - did not require any follow up treatment
23CF-3D73-8982-4560	08/08/2014 23:22	201408	201405 (Aug)	VB08Z	Emergency Medicine, category 2 Investigation with category 1 Treatment	Health care provider: same or other	THH A&E	47	Admitted to a Hospital Bed /became a LODGED PATIENT of the same Health Care Provider
09B3-7C41-5E60-44EA	28/07/2014 17:08	201407	201404 (Jul)	VB11Z	Emergency Medicine, No Investigation with No Significant Treatment	Self referral	THH UCC	39	Discharged - follow up treatment to be provided by GENERAL PRACTITIONER
6867-8FBF-9552-4E62	26/11/2014 19:09	201411	201408 (Nov)	VB08Z	Emergency Medicine, category 2 Investigation with category 1 Treatment	Self referral	THH UCC	12	Referred to Fracture Clinic
02CA-3DAA-5CA9-4BCC	30/11/2014 05:40	201411	201408 (Nov)	VB11Z	Emergency Medicine, No Investigation with No Significant Treatment	Other	THH A&E	15	Discharged - did not require any follow up treatment
2C8B-3222-416B-4CED	25/03/2015 22:40	201503	201412 (Mar)	VB07Z	Emergency Medicine, category 2 Investigation with category 2 Treatment	Self referral	THH A&E	46	Discharged - did not require any follow up treatment
4DFD-9FD6-A8F7-4B54	01/03/2015 16:39	201503	201412 (Mar)	VB11Z	Emergency Medicine, No Investigation with No Significant Treatment	Self referral	THH Mount Vernon MIU	50	Discharged - did not require any follow up treatment
B35E-F89E-1922-48EA	25/09/2014 21:50	201409	201406 (Sep)	VB11Z	Emergency Medicine, No Investigation with No Significant Treatment	Self referral	THH UCC	26	Discharged - follow up treatment to be provided by GENERAL PRACTITIONER
B26E-4660-83DA-4487	30/11/2014 03:39	201411	201408 (Nov)	VB11Z	Emergency Medicine, No Investigation with No Significant Treatment	Self referral	THH UCC	1	Discharged - follow up treatment to be provided by GENERAL PRACTITIONER
11AA-D108-403F-42D0	16/05/2014 16:22	201405	201402 (May)	VB09Z	Emergency Medicine, category 1 Investigation with category 1-2 Treatment	Health care provider: same or other	THH A&E	46	Admitted to a Hospital Bed /became a LODGED PATIENT of the same Health Care Provider

Snippet of A&E Data Instances

D.1 Appendix D - MVSimO Ontological Model



MVSimO Ontological Model