BIM-BASED AUTOMATED FAULT DETECTION AND DIAGNOSIS OF HVAC SYSTEMS USING KNOWLEDGE MODELS

Arash Hosseini Gourabpasi

A Thesis

In the Department of

Department of Building, Civil and Environmental Engineering (BCEE)
Gina Cody School of Engineering and Computer Science

Presented in Partial Fulfillment of the Requirements

for the Degree of

Doctor of Philosophy (Civil Engineering) at

Concordia University

Montreal, Québec, Canada

November 2024

CONCORDIA UNIVERSITY

SCHOOL OF GRADUATE STUDIES

This is to certify that the thesis prepared

By: Arash Hosseini Gourabpasi

Entitled: BIM-Based Automated Fault Detection and Diagnosis of HVAC Systems Using Knowledge Models

and submitted in partial fulfillment of the requirements for the Degree of

Doctor Of Philosophy (Civil Engineering)

complies with the regulations of the University and meets the accepted standards with respect to originality and quality.

Signed by the final Examining Committee:

______________________________ Chair

Dr. Ferhat Khendek

______________________________ External Examiner

Dr. Ahmad Jrade

______________________________ Examiner

Dr. Amin Hammad

______________________________ Examiner

Dr. Sang Hyeok Han

______________________________ Examiner

Dr. Fuzhan Nasiri

______________________________ Supervisor

Dr. Mazdak Nik-Bakht

Approved by ______________________________

Dr. Po-Han Chen, Graduate Program Director

08/12/024 ______________________________

Dr. Mourad Debbabi, Dean

Gina Cody School of Engineering and Computer Science

Abstract

BIM-Based Automated Fault Detection and Diagnosis in HVAC Systems Using Knowledge Models

Arash Hosseini Gourabpasi, Ph.D

Concordia University, 2024

Automated Fault Detection and Diagnosis (AFDD) of building mechanical systems, including HVAC (Heating, Ventilation, and Air Conditioning), has received substantial attention recently from both research and application angles. The reasons are attributed to potential savings in energy consumption and maintenance. Various methods, including simulation and Grey-Box, are offered, but data-driven ones have received the most attention due to reduced manual effort, integrability, and scalability. Accordingly, to enhance energy efficiency and reduce operational costs, various Machine Learning (ML) models have been developed for AFDD of HVAC systems. However, the implementation of such data-driven approaches has often translated into a loss of contextual data. This study integrates operational data with building information and its various disciplines, linking the two to facilitate AFDD model development. BIM (Building Information Model) and BAS/BMS (Building Automation System/Building Management System) data are the repositories utilized for this integration.

The proposed solution integrates bottom-up (data-driven via Machine Learning) and top-down (knowledge-oriented via Semantic Web Technologies) AI approaches to generate an effective AFDD knowledge model. The study materializes a two-way flow of data and knowledge between the BIM and BMS by utilizing an ontology named AFDDOnto, which integrates building components with fault types, methods, and parameters. The solution enables AFDD algorithms to utilize static and dynamic information related to HVAC and building spaces to develop enriched AFDD models. It incorporates building spatial information and stores analytics to represent the facility's as-is state. The proposed BIM-based knowledge solution can be used for AFDD model development, tracking changes, and analysis and visualization in two ways. Firstly, to integrate the BIM features with BMS features for creating ‘context-aware’ AFDD models. Secondly, to semantically store BIM-based AFDD performance analytics through AFDDOnto that can be used for model comparison, reproduction and visualization through knowledge graphs.

The knowledge stored in the repository can be queried, which enables access to contextual information (knowledge graphs, images, videos, project snippets); spatial data (locations, states); and apriori knowledge (configuration and analytics) to enable development, application, and visualization of context-aware AFDD models. Additionally, the proposed solution can maintain access to external project files and databases to enable interoperability between BIM and BAS/BMS. The potential users include HVAC operators, BIM Managers, and Facility Managers tasked with the operation and maintenance of HVAC systems.

Acknowledgments

My most heartfelt thanks go to my beloved Zhinous for her unwavering companionship and support. She gives meaning to my life and makes it truly worthwhile. She has been there at every step of my personal and professional journey, and this would not have been possible without her.

I would also like to express my gratitude to Dr. Mazdak Nik-Bakht, who served as my research supervisor. His guidance and expertise have helped me grow and mature throughout the planning, designing, development, and execution of this research effort. I am grateful to have shared this journey with him.

I also want to thank my committee members, Dr. Amin Hammad, Dr. Fuzhan Nasiri, and Dr. Sang Hyeok Han, for their valuable feedback, time, effort, and encouragement. Additionally, I would like to extend my appreciation to the GCS faculty for providing financial support.

Furthermore, I appreciate the support from my colleagues at the COMPLECCiTY lab: Abdelhady Ossama Hosny, Sobhan Kouhestani, Farzaneh Zarei, Leila Rafati, Abraham Martinez, and all the other members who provided valuable insights and support.

My parents, Esmail Hosseini, Gita Dhingra, Fatemeh Mirzaei, Mehdi Asadi, and Santosh Dhingra, deserve my gratitude for their unconditional support. My family members Neha Hosseini, Mahdis Asadi, Pardis Asadi, and Soheil Hamidi were there when I needed them the most.

Preface

This thesis is submitted for the Degree of Doctor of Philosophy in Civil Engineering at Concordia University. The research was conducted under the supervision of Dr. Mazdak Nik-Bakht in the Department of Building, Civil, and Environmental Engineering.

The results of the conducted research for this thesis have been published in the form of 5 technical papers as follows:

Journal papers:

1. Hosseini Gourabpasi A, Nik-Bakht M. Knowledge discovery by analyzing the state of the art of data-driven fault detection and diagnostics of building HVAC. CivilEng. 2021; 2(4):986-1008. https://doi.org/10.3390/civileng2040053
2. Hosseini Gourabpasi A, Nik-Bakht M. BIM-based automated fault detection and diagnostics of HVAC systems in commercial buildings; Journal of Building Engineering, Volume 87, 15 June 2024, 109022. https://doi.org/10.1016/j.jobe.2024.109022
3. Hosseini Gourabpasi A, Nik-Bakht M (01 Jul 2024): An ontology for automated fault detection & diagnostics of HVAC using BIM and machine learning concepts, Science and Technology for the Built Environment, https://doi.org/10.1080/23744731.2024.2363104

Conference article:

1. Hosseini Gourabpasi A, Nik-Bakht M. Live BIM for capturing dynamism of physical spaces, occupants and assets through linked data. ASHRAE and IBPSA-USA. Sep 1, 2021.

Book Chapter:

2. Hosseini Gourabpasi A, Nik-Bakht M. Stateful BIM(BIM) for Digital twinning of the built environment, in Next-Generation Cities Encyclopedia. 2024 (in press).

Table of Contents

List of Figures

List of Tables

Nomenclature

Abbreviation	Full Name
ABOX	Assertion Box
AEC	Architecture, Engineering, and Construction
AEC/FM	Architecture, Engineering, Construction, and Facility Management
AFDD	Automated Fault Detection and Diagnosis
AFDDOnto	Automated Fault Detection and Diagnosis Ontology
AI	Artificial Intelligence
ANN	Artificial Neural Network
API	Application Programming Interface
ARM	Association Rule Mining
ASHRAE	American Society of Heating, Refrigerating and Air-Conditioning Engineers
ASCE	American Society of Civil Engineers
BAS	Building Automation System
BEMS	Building Energy Management Systems
BIM	Building Information Model
BIMQL	Building Information Model Query Language
Blazegraph	Blazegraph Database
BMS	Building Management System
BOT	Building Topology Ontology
CAD	Computer-Aided Design
CMMS	Computerized Maintenance Management System
COBie	Construction-Operations Building information exchange

Abbreviation	Full Name
CRISP-DM	Cross-Industry Standard Process for Data Mining
CSV	Comma-Separated Values
DT	Digital Twin
ETL	Extract, Transform, Load
FD	Fault Detection
FDD	Fault Detection and Diagnosis
FOI	Feature of Interest
FP-growth	Frequent Pattern Growth
FRP	Flexible Research Platform
gbXML	Green Building Extensible Markup Language
GUID	Globally Unique Identifier
HVAC	Heating, Ventilation, and Air Conditioning
HVAC/R	Heating, Ventilation, Air Conditioning, and Refrigeration
IDM	Information Delivery Manual
IFC	Industry Foundation Classes
IFC4	Industry Foundation Classes version 4
IFCOWL	Industry Foundation Classes Web Ontology Language
IoT	Internet of Things
ISOM	Inductive Self-Organizing Map
JDBC	Java Database Connectivity
JSON	JavaScript Object Notation
JSDAI	Java Standard Data Access Interface
LOD	Level of Development

Abbreviation	Full Name
LBD	Linked Building Data
ML	Machine Learning
MVD	Model View Definition
NoSQL	Non-SQL
O&M	Operation and Maintenance
ODBC	Open Database Connectivity
OPM	Ontology for Property Management
OWL	Web Ontology Language
QTO	Quantity Take-off
RDF	Resource Description Framework
ReLU	Rectified Linear Unit
RTU	Roof-Top Unit
SAREF	Smart Appliances REFerence
SVM	Support Vector Machine
SQL	Structured Query Language
SPARQL	Simple Protocol and RDF Query Language
SSN	Semantic Sensor Network
STEP	Standard for the Exchange of Product model data
SWRLS	Semantic Web Rule Language
TBOX	Terminological Box
TTL	Terse RDF Triple Language
URI	Uniform Resource Identifier
UUID	Universally Unique Identifier

Abbreviation	Full Name
VAV	Variable Air Volume
W3C	World Wide Web Consortium
XGBoost	Extreme Gradient Boosting

Chapter 1: Introduction

1.1 Motivation and Background

The heating, ventilation, air conditioning (HVAC), and refrigeration systems are arguably use up the most energy out of all a building's physical assets. HVAC/R systems regulate the temperature, humidity, quality, and air movement in buildings, making them critical for occupant comfort, health, and productivity. In Canadian commercial stores, HVAC and lighting combined contribute to 90% of energy consumption [1]. In 2011, heating systems, particularly furnaces (57%), followed by electric baseboards (27%) and boilers (5%), were the primary type of heating system used by Canadian households [2]. This energy consumption indicates the dependency of Canadian households and commercial buildings on the HVAC system, and hence emphasizes the importance of timely and accurate identification of its faults.

Performance in HVAC systems and sub-systems are negatively affected by system degradation, operational misuse, reduced maintenance, and sensor issues [3] [4]. Many HVAC faults that require repair or immediate attention go unnoticed and cause progressive damages. The most common components where faults occur are the damper, fan, filter, and other parts such as sensors [5]. Furthermore, faults in HVAC systems affect the HVAC's energy consumption. For example, when the refrigerant charge is less than 25% of the design value, it can reduce the energy efficiency by 15%. Moreover, 20% capacity loss is also reported in such situations [6]. The reasons mentioned above can lead to increased energy usage in addition to user discomfort, shorter equipment life, and less reliability [4]. Malfunctioning sensors, components, and control systems and degrading systems in HVAC and lighting systems are the main reasons for energy wastage and an unsatisfactory indoor environment [7].

Fault detection and diagnostics in HVAC systems enable asset managers to promptly identify and locate faults, a process known as fault detection, and to further specify the type of fault, a process known as fault diagnosis. Current advancements in the Internet of Things (IoT) have led to the application of big data for creating automated fault detection and diagnostic (AFDD) models, which can be developed using machine learning (ML) techniques. The sensory data available in building automation systems (BAS) and building management systems (BMS) are used to detect the HVAC's faults and perform diagnostics. In asset management of buildings, energy management and maintenance models differ in scope and structure. While models for energy management describe continuous states (energy, temperature, etc.) and usually assume the HVAC to be in a healthy condition, the models used for maintenance do not consider human factors such as comfort and only describe discrete states, such as faulty/non-faulty states of equipment and fault typology [8].

Data-driven AFDD methods that have garnered significant attention from practitioners and researchers often lack context related to HVAC and its faults in terms of spatial information and their relationship with surroundings. In this regard, BIM (Building Information Modelling) presents a promising solution for managing assets and improving facility operations. BIM as a solution contains geometric, spatial, and semantic relationships which AFDD models lack. However, BIM in its current implementation lacks the ability to gather and display up-to-date information about building components.

Typically, BIM data is only exported for analysis without updating the BIM model, resulting in a disconnect between the digital model and the actual condition of assets. This limitation arises from the perception of BIM as a static model, primarily reflecting the as-built or as-designed state of components rather than the ongoing state of the facility. Consequently, BIM is underutilized for asset management. To harness BIM's full potential, it must transform to become a dynamic model capable of representing the facility in near real-time or periodically. This involves linking entity-related data and ensuring connectivity with sensor-generated data streams.

In this research Experts' knowledge was also directly extracted twice during the study; first time for association rules identified in the literature, through a survey, completed by 13 experts from Mechanical and Machine Learning areas with expertise in FDD of HVAC; and second time, through semi-structured interviews with seven experts in the domain of HVAC, Machine Learning and asset management with expertise in FDD of HVAC to validate the developed AFDDOnto.

1.2 Problem Statement

Buildings account for one-third of worldwide energy consumption, while HVAC systems consume half of all building energy. HVAC-related issues can decrease energy efficiency by as much as 15% [6]. In addition to lowering energy efficiency, HVAC problems can cause discomfort for building occupants and drive-up operating expenses. A building's annual energy costs can be reduced by as much as 10% when AFDD (Automated Fault Detection and Diagnostics) models are used to restore functionality to its original state [9] [10]. Hence, detecting faults accurately is essential as it allows it to control, mitigate, and further allow its users to understand energy consumption patterns.

AFDD models provide means of automated fault discovery by detecting anomalies. HVAC performance can be negatively impacted by system degradation over time, reduced maintenance, and sensor-related issues such as availability and condition [3]. The absence of a suitable AFDD method in HVAC can cause false negatives and false positive alarms or allow faults to go unnoticed, which can contribute to increased building energy consumption and affect the reliability of the HVAC system, leading to user discomfort and shortening the HVAC equipment life expectancy [4].

There has been a significant increase in the utilization of data-driven techniques for FDD [11] [12]. The primary motivation behind opting for these methods is that AFDD models can be created mainly by relying on the sensory data that is stored in BAS/BMS (Building Automation System/ Building Management System).

Typically, such systems collect near real-time data from sensors available within a building environment or system at a predefined and periodic interval of time. The data types can vary depending on the type of sensors used in the HVAC system and building type. For example, multiple sensors can be installed in a building facility to collect temperature, humidity, and pressure data. The data collection process involves continuous monitoring of the system and logging of sensory data to capture building systems' behavior. However, building facilities are often equipped with a limited number of sensors.

A major challenge for AFDD models is the lack of semantic information and contextual data related to buildings and HVAC systems. This absence of detailed context often results in AFDD solutions being perceived as 'black-boxes' which complicates interpretation and decision-making compared to 'white-box' models that are based on physical principles. Without comprehensive contextual information, these models struggle to accurately diagnose faults and are frequently underutilized as decision support systems. This gap in semantic understanding significantly impairs the effectiveness and reliability of AFDD models in real-world applications, limiting their ability to fully support fault localization and analysis.

BIM offers a valuable source of contextual information to overcome these challenges. A BIM-based approach can cover the entire lifecycle of a building, including the Operation and Maintenance (O&M) phase. While traditionally utilized the most during the building design and construction phase; BIM's potentials extend beyond. BIM integrates 3D geometry models with the semantics of the building and physical asset data and metadata such as HVAC and its peripherals like sensors.

Additionally, BIM fosters collaboration during the entire lifecycle of the building and is utilized for HVAC asset management during the operation stage phase. By integrating BIM with BMS, it becomes possible to create a decision support system, which enables capturing essential knowledge that can be leveraged for AFDD model development [13] [14] [15]. The AFDD model can benefit from access to building contextual data and utilizing BIM data to generate additional features for AFDD model development.

However, to fully unlock its potential for use cases such as HVAC AFDD, the technology needs to evolve beyond its current state. This evolution entails representing and providing access to an accurate "as-is" model of the building, facility, and its components. Such an evolved BIM framework can enable feature generation and diagnostic capabilities and serve as a knowledge model to facilitate the AFDD of HVAC. By incorporating knowledge models in the form of ontologies, BIM modules can be extended to support HVAC AFDD analytics, facilitating the bi-directional flow of data and knowledge between BIM and BAS. Existing literature suggests that BIM-based knowledge models can address the limitations of HVAC AFDD models by effectively using BIM-compatible ontologies [16] [17] [18] [19] [20]. However, a comprehensive methodology for implementation is currently lacking, with existing solutions still in the early stages of research or conceptual form.

1.3 Research Objectives

The primary objective of this research is to use BIM-based semantic knowledge related to building and physical assets to improve the AFDD of HVAC in commercial buildings by adding dynamic features created using BIM and BMS data. By doing so, the research aims to create a method that can use semantic information using BIM to enhance the AFDD of HVAC, which is conventionally supported by information sources such as BAS and BMS. The following five sub-objectives are defined along with their corresponding tasks to achieve the expected outcome.

Research Sub-objective 1: To identify the requirements of sensory data needed for creating a BIM-based knowledge model to apply AFDD into HVAC systems. Extensive literature analysis is performed to identify the sensory data required from BIM and BMS/BAS, as well as the algorithms, fault types, and HVAC types that need to be identified.

Research Sub-objective 2: To examine integration methods for BAS/BMS data and BIM for AFDD of HVAC based on the IFC schema. Integration strategies and methods between BIM and BAS/BMS are tested for the use case of AFDD of HVAC by extending the IFC Schema to capture and store AFDD data analytics in BIM.

Research Sub-objective 3: To develop a BIM-based knowledge model to facilitate interoperability between BIM and BAS/BMS, enabling a bi-directional flow of data and analytics.

A knowledge model in the form of an ontology is created to apply AFDD of HVAC using dynamic BIM concepts and capture model parameters and model evaluation.

Research Sub-objective 4: To implement the proposed methodology for developing a BIM-based knowledge model for AFDD of HVAC. A test facility is utilized as a case study in which the BIM model is used to enrich the AFDD model using additional BIM-based features. Conversely, the analytics from the AFDD model are captured to represent the present state of the building asset.

1.4 Thesis Organization

This thesis is based on the core ideas addressed in the publications in the form of a book chapter, a conference paper, and three journal papers. Figure 1.1 , Shows the components and their overlap to achieve the objectives and sub-objectives of the research work. The rest of this thesis is organized as follows:

Chapter 2: Literature Analysis – This chapter introduces AFDD for HVAC systems and discusses the role of knowledge management through ontologies. It begins by presenting the inputs, algorithms, and types of HVAC faults. Subsequently, association mining techniques are employed to form rules that are utilized in the proposed knowledge model. The final section identifies gaps in the existing literature and outlines the scope of the study investigated in this thesis.

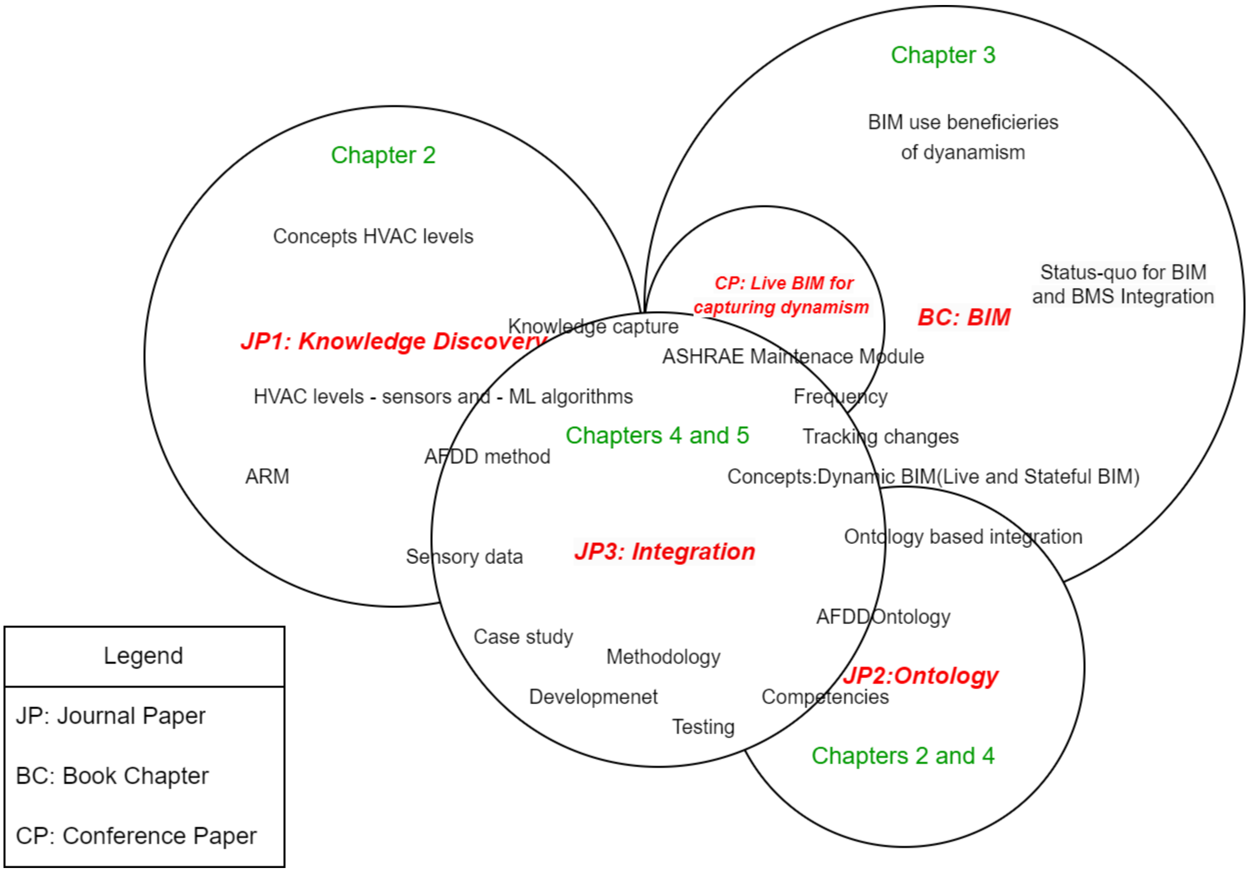


Figure 1.1: Thesis Organization

Chapter 3: This chapter provides an in-depth analysis of the role of dynamism in BIM, explores the dynamic BIM enablers, and explains how the data required for AFDD of HVAC systems can be integrated with BIM.

Chapter 4: Methods – This chapter covers two main topics: the development of a knowledge model in the form of an ontology, and the integration methods between BIM and BAS/BMS. The integration method extends BIM using the proposed ontology, enabling a bi-directional flow of data to facilitate the development of a BIM-based AFDD model.

Chapter 5: Implementation – In this chapter, a case study is used to demonstrate the applicability of the proposed integration method. A BIM model is employed to create additional dynamic features to facilitate the AFDD model. Furthermore, the analytics stored in the form of knowledge are accessed through knowledge graphs and queries.

Chapters 6 and 7: The final two chapters present the findings of the case study and discussions on the topic. The last chapter comprises the research conclusions, contribution to the field's knowledge, and suggestions for future work.

Chapter 2: Literature Analysis – AFDD of HVAC System and Application of Knowledge Models [1]

The literature review and analysis conducted in this study aims to provide a comprehensive understanding of the current state of research on AFDD models for HVAC systems. The review begins by exploring the existing body of knowledge on FDD methods and analyzing their strengths and limitations. It then focuses on the potential application of knowledge models in the form of ontologies applicable to building operation. By examining the available literature, this review highlights the current gaps in research and sheds light on the potential benefits and challenges.

2.1 Automated FDD of the HVAC System

At a holistic level, HVAC can be studied at a system or local level [21]. Local-level classification can be divided into (i) sub-system level and (ii) equipment/component level [22], which we refer to as 'HVAC levels' in this thesis. The full HVAC system consists of the sub-systems and/or components coupled together. In the past two decades, fault detection has been mainly applied to the HVAC at the sub-system level, and very few researchers have looked at detecting faults at the whole building level [7].

Hereafter in this thesis, we use the term HVAC system in a general sense, by which we also refer to sub-systems and pieces of equipment in HVAC. System-level faults refer to the occurrence of a fault in one sub-system or equipment and its consequence at the system level [7]. Previous literature reviews in the domain of FDD have focused on overall FDD modeling methods [23] [24] [25] [26] [27] or data-driven methods [12]. Another group of review studies focuses on a specific step of the procedure, such as algorithms [28] or fault types [29].

However, the current thesis is different in the way that it analyzes the models developed in the literature by looking at the features used, fault types identified, corresponding HVAC systems, and algorithms used for data-driven FDD models. Through an affinity analysis of these studies, we extract knowledge in the form of association rules and deploy them in the form of a recommender system. The scope of this study and the recommender system is mainly

---

[1] The material of this chapter is published in form of the following publications:

1. Hosseini Gourabpasi A, Nik-Bakht M. Knowledge Discovery by Analyzing the State of the Art of Data-Driven Fault Detection and Diagnostics of Building HVAC. CivilEng. 2021; 2(4):986-1008. https://doi.org/10.3390/civileng2040053
2. Hosseini Gourabpasi A, Nik-Bakht M (01 Jul 2024): An ontology for automated fault detection & diagnostics of HVAC using BIM and machine learning concepts, Science and Technology for the Built Environment, https://doi.org/10.1080/23744731.2024.2363104

commercial buildings since the majority of AFDD models developed in the literature have been of this type.

One classification for AFDD methods is the top-down versus bottom-up approach. The top-down approach detects faults that manifest themselves at the whole building level, whereas the bottom-up approach focuses on the component or sub-system level. In both approaches, models of ideal operation conditions are compared with actual measurements to detect faulty or abnormal behavior [30]. Whole-building fault detection usually makes use of a top-down fault detection strategy [7]; the top-down approach is comparatively more difficult than the bottom-up. In the top-down approach, further analysis is required to locate faults because of the system-level effect that causes the faults' symptoms to spread across the system [31] [32].

Based on ASHRAE's recommendations [33] as found in [34], the two main modeling methods are the forward (classical) approach and the data-driven (inverse) approach. The forward approach is known as white box/engineering methods. Forward approaches usually require detailed knowledge of various system processes and interactions. Most simulation software tools use such approaches. The data-driven and model-based classifications have been found to be the most common FDD classification approaches [35] [36] [37] [38], which have also been referred to as model-free methods and model-based methods [39] in the literature.

The other common classification found in the literature categorizes the AFDD techniques into model-based methods, rule-based methods, and data-driven methods [5] [40]. The data-driven methods are also called process history-based [32] [41]. In some cases, knowledge-based is also included in data-driven models [36]. Further, model-based classification is also referred to as quantitative [41] [42]. The other classifications found are analytical model-based, signal-based, and data-driven methods [43].

On the other hand, the data-driven models are classified by ASHRAE into calibrated simulation models, grey-box models, and black-box models/empirical approaches. In simple terms, calibrated simulation is similar to forward approaches and requires detailed knowledge of the system and processes, but black-box models are data-driven and use statistical or artificial intelligence approaches to develop models. Grey-box models, on the other hand, are formulated using training data and physical principles and are a mix of simulation and black-box model [34].

The AFDD techniques (black-box methods) reviewed in the literature are broadly grouped and categorized into supervised and unsupervised learning. Most of the reviewed studies implementing AFDD are supervised methods and treat the FDD as essentially a classification problem. Unsupervised methods are mainly adopted in the pre-processing phase or are used for fault detection through clustering.

Extensive research has been conducted on HVAC FDD at various levels, including systems [44] [45], sub-systems , and components/equipment [46] [47]. Three main categories of HVAC FDD are: Calibrated Simulation [48] [49], Grey-Box models [50] [51], and Empirical Approach.

As illustrated in Figure 2.1 . Calibrated Simulation models can predict HVAC system behavior with high accuracy, but developing such models requires significant time and involves domain expertise. This is because such models often require calibration against real-world data. In contrast, Grey-Box models integrate and utilize data in addition to utilizing expert knowledge and physical system models to improve the accuracy of FDD models [52].

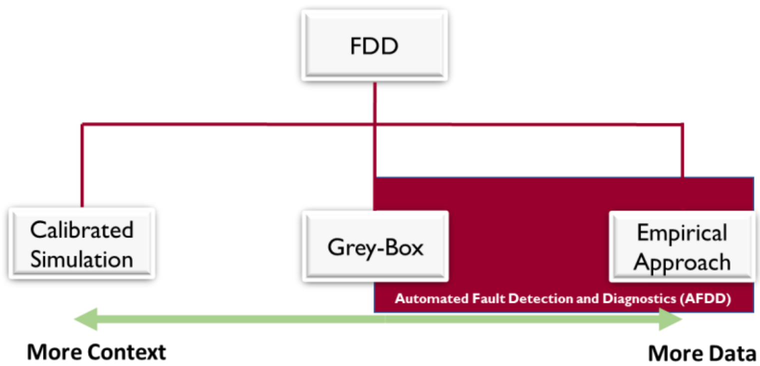


Figure 2.1: Fault Detection and Diagnostics of HVAC

On the other hand, the Empirical Approach prioritizes and emphasizes a more data-driven model development strategy, hence making it easier to develop but potentially lacking in-depth contextual information [52]. This study specifically concentrates on Automated Fault Detection and Diagnosis (AFDD) within the broader scope of FDD. The AFDD emphasizes data-dependent models, i.e., mathematical models that aim to detect and diagnose faults in the mechanical systems by analyzing their sensory data.

The common source of sensory data needed for AFDD model development is found in BMS and BAS systems. AFDD models leverage diverse machine learning algorithms designed for the FDD of HVAC systems. The selection of a specific machine learning model is contingent upon the characteristics of the HVAC and facility, such as the type of HVAC, faults considered, and sensory data available.

A comprehensive approach is necessary to ensure accurate and timely diagnosis of faults in HVAC systems, incorporating contextual information and knowledge. This involves considering expert opinions, historical data, industry guidelines, and best practices due to these systems' complex and interconnected nature [53] [54]. However, the trend shifts towards adopting more data-centric approaches in AFDD models, such as empirical models, due to lesser dependability on expert opinion. As a result, models dependent on data usually lack contextual information about the building, environment, and occupants.

Therefore, striking a balance between data-driven approaches and the inclusion of contextual information is crucial for reliable and effective HVAC diagnosis. One valuable source of information is Building Information Modeling (BIM), which contains detailed data often available in calibrated simulation models but frequently absent in AFDD models of HVAC systems. This wealth of information from BIM can provide the contextual and semantic data necessary to bridge the gap and enhance the utility of AFDD models.

AFDD models for HVAC systems rely heavily on the type of data available and the choice of algorithms suitable for the specific HVAC equipment. The effectiveness of AFDD models is contingent upon the quality and diversity of sensory data accessible from the BMS or other sources. Additionally, selecting appropriate algorithms plays a crucial role in accurately detecting and diagnosing faults in HVAC components [52]. Therefore, to enhance the performance of AFDD models, it is essential to consider the availability of relevant data types and employ algorithms tailored to the specific HVAC systems being monitored.

2.2 Features, Fault Types and Algorithms for AFDD

A total of 109 papers (Appendix A), were initially collected and reviewed to focus on those providing a complete FD (Fault detection) or FDD model. From this review, 82 studies were selected. Machine learning algorithms used in pre-processing and post-processing stages were excluded; only data-driven techniques applied during HVAC fault detection and diagnosis were considered. Supplementary information collected includes HVAC type, data sources for the AFDD process (both synthetic and real), and data collection frequency. Features and faults for each HVAC system were gathered and visualized with a Sankey diagram.

Figure 2.2, summarizes the features used in the analyzed literature for AFDD. They are ranked based on their frequency of use in the HVAC FDD models reported in the literature. It is evident that ‘temperature’ is the single most crucial feature used for AFDD, as its application also extends to the second most frequently used feature, i.e., the ‘calculated measure’. This feature commonly uses arithmetic operations such as subtraction and often uses features such as ‘temperature’ or ‘pressure’ as the calculation component; for example, the calculated measure is used to show the temperature difference between the supply and return air or pressure difference between the entrance and the exit (inlet and outlet) to indicate pressure drop or increase. Other frequently used parameters include the 'pressure' and 'flow rate'. State-representative information and energy-related parameters such as 'Opening/position', which represents physical characteristics such as position or percentage of a valve being open or closed, and 'Load' and 'Energy' categories are among other attributes frequently used by AFDD models.

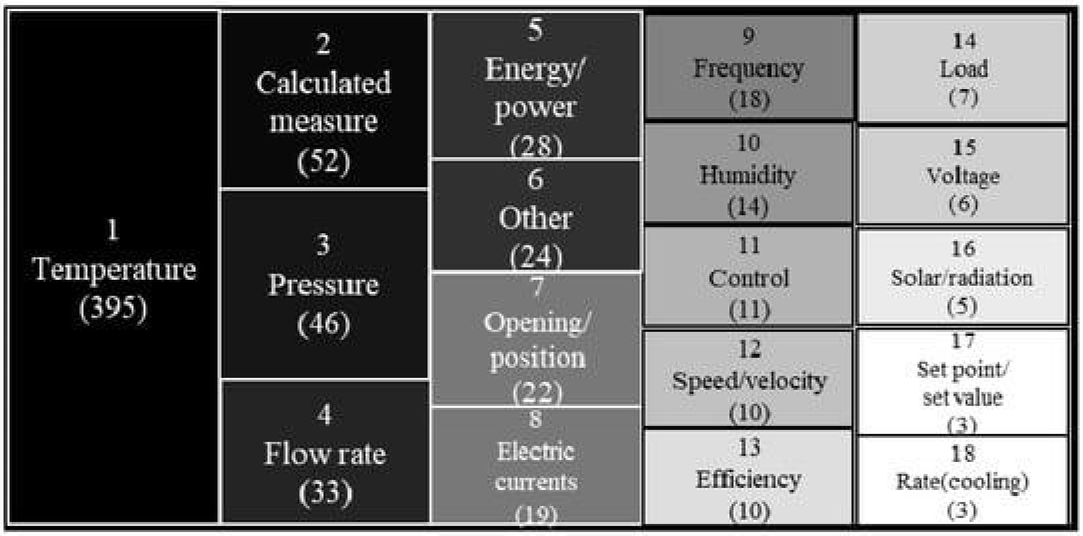


Figure 2.2: Common Features Selected for AFDD of HVAC (18 categories formed for 706 features used in the analyzed literature and the numbers in brackets represent the frequency of occurrence in the literature)

Several fault classification systems exist in the literature, such as lists of prioritized HVAC faults or the faults targeted more specifically, such as those for chillers [55] [56] [57] [58]; however, they cannot be used in this study. Some classifications are specific to a particular sub-system, such as chillers [55], or, if they cover the whole HVAC system [56] [57] [58] , they are too detailed and elaborate and cannot support the abstraction required for rule mining. Accordingly, in this study, eighteen (18) fault categories were created and introduced to solely organize, categorize and analyze more than 400 faults reported in the literature investigated for HVAC's most common faults detected using data-driven methods in this thesis.

It must be noticed that the categories shown in Table 1, are not meant to provide a comprehensive classification of all fault types. The faults considered apply to the HVAC system, sub-system, and/or components. The faults are categorized based on the following procedure. The categories are created using a hypernym keyword. The faults are hyponym and belong to only one of the eighteen hypernyms created. Then, logical reasoning is performed to assign each fault to the category that it best represents. In cases where a hyponym consists of more than one word in its description, the first word will be selected, and the assignment is carried out based on that word. For example, for the fault type referred to as ‘control unstable’, the term ‘control’ is considered the primary word, and ‘unstable’ is a condition associated with controlling. Hence, the fault is assigned to the ‘control’ category. The only exception applies to faults that include bias/drift. In particular, for sensor faults, we skip the sensor type, even if it is the first word of the fault description and look at the following term in the description.

Table 1: Category of Faults Identified for Data-driven Techniques

Rank	Fault category	Count
1	Limit issue	68
2	Stuck/Partially closed	67
3	Flow problems	54
4	Bias/Drift/Calibration	49
5	Leakage	41
6	Foul	38
7	Other faults	20
8	Non-functioning	20
9	Non-condensable	18
10	Control	18
11	Temperature issue	12
12	Speed	12
13	Set point	8
14	Performance	8
15	Capacity reduction	5
16	Blockage	4
17	Schedule	3
18	Sizing issue	3

The categories of the faults are sorted in Table 1 in descending order of occurrence frequency in our database. The ‘Limit issue’, which is the dominant category, comprises faults related to over/undercharge, excess oil, or reduced evaporator. The second category, ‘stuck/partially closed’, includes faults such as exhausted air, damper stuck (fully open), or cooling coil valve partially closed (15% open). The other categories’ names such as ‘temperature issue’,

’blockage’, ’speed’, and ‘non-functioning’ are self-explanatory. ‘Flow problems’ and sensor-related faults, which are categorized as ‘bias/drift/calibration’ alongside ‘leakage’ and ‘foul’-related faults, comprise the top six frequent categories of HVAC faults. The ‘other faults’ comprises of different types of faults that did not form a category due to limited appearance in the database.

The fault categories such as ‘set point’ and ‘non-condensable’ belong to a particular type of fault, and on the other hand, fault categories such as ‘control’ and ‘performance’ belong to a more diverse pool of faults pertaining to their respective categories. It is evident that most studies have relied purely on sensory data and very few categories with a small occurrence in our database represent faults that may be detected given static information such as ‘schedule’ and ‘sizing issue’, which are categories with the lowest counts in the table.

3.3 Systematic Association of HVAC Faults and AFDD algorithms

In order to understand the latent relationships and associations between the common HVAC faults and/or AFDD techniques, association rule mining (ARM) has been used. ARM is an unsupervised machine learning procedure in which the aim is to observe the frequently occurring patterns, correlations, and associations in a dataset. Association mining is performed in two steps. The first step is to generate ’frequent itemsets’. The second is generating rules, where rules are generated and filtered based on set constraints. Two models were trained in this study: one for detecting affinity between various fault types and a second model to investigate the association between the FDD techniques and the HVAC’s most common faults.

The FP-growth algorithm is an improved affinity analysis algorithm, in which the number of scans of the database is reduced to find the frequent itemsets [59]. In this study, FP-growth was implemented in the model to generate frequent itemsets of fault types and then extract relationships of a high level of support and confidence as rules. The rules take the form of a ‘premise’, followed by a ‘conclusion’. The metrics considered in the model development are support and confidence, where confidence is used as a measure of the strength of the rule and support correlates to statistical significance. The equation for support of a rule and confidence of a rule are as shown below:

Rule: $(X \rightarrow \mathrm{Y})$ (1)

Support$(X \rightarrow \mathrm{Y}) = \Pr(\mathrm{X}, \mathrm{Y}) = \frac{\mathrm{n}(\mathrm{X},\mathrm{Y})}{\mathrm{N}}$ (2)

Confidence$(X \rightarrow \mathrm{Y}) = \Pr(\mathrm{Y}|\mathrm{X}) = \frac{\Pr(\mathrm{X},\mathrm{Y})}{\Pr(\mathrm{X})}$ (3)

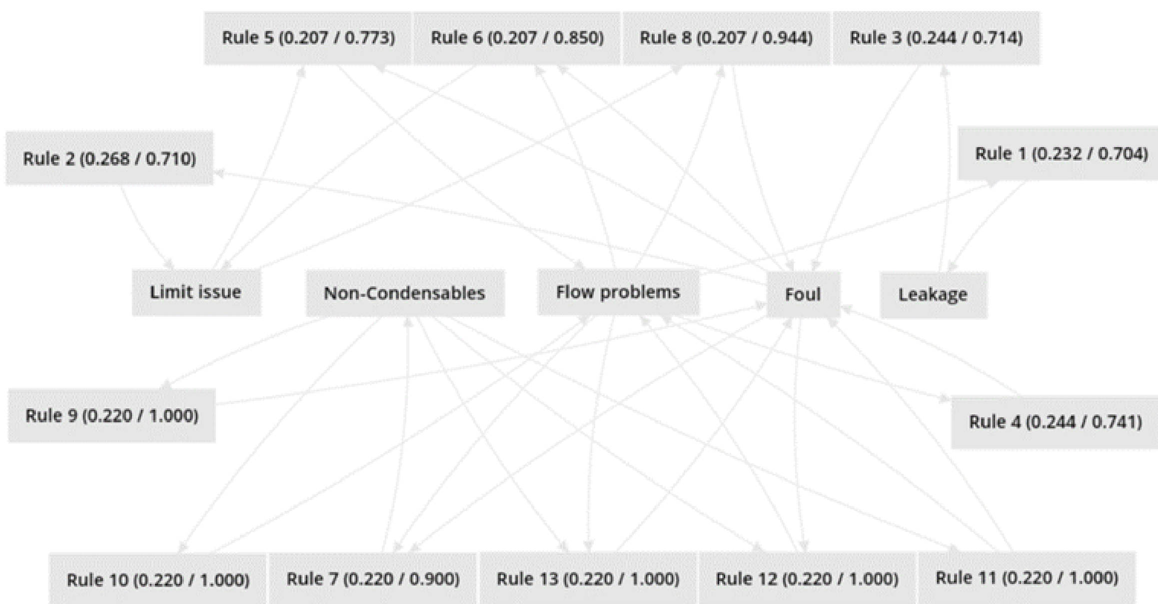
where X and Y are independent items or itemsets, n is the relative frequency of occurrence and N is the total transaction numbers.

Minimum support andminimum confidence are needed to eliminate the unimportant association rules [59] [60]. Syntactic constraints were enforced for the second model to add restrictions on items that can be included in the rule.

The frequent itemsets are created using the FP-growth algorithm, which has been assigned minimum support of 20% for the frequent itemsets and minimum confidence of 70% for detecting the association rules. By applying these criteria, five frequent fault types and 13 rules are identified through the FP-growth algorithm as shown in Figure 2.3. On the left-hand side of each rule are the premises and, on the right, after, the arrow is the conclusion. For example, rule #5 indicates that “if ‘limit issue’ fault and ‘foul’ fault are found simultaneously in the HVAC system for the designed FDD algorithm found in the literature, then it is likely that the system is also designed to detect ‘flow problem’-related issues”, with 77.3% confidence.

The rules mined have either one or two fault categories in their premises. In seven of the rules that have one fault category in their premises, rule #9 and rule #10 have equal support of 22%, and confidence of 100%, which indicates “if ‘non-condensables’ fault occurs then there are equal chances that ‘foul’ and ‘flow problems’ related issues can be existing separately” or as per rule #11 the ’foul’ and ’flow problems’ can appear simultaneously. The first rule mined indicates that when ‘flow problems’ are found using the particular FDD algorithm, then it is likely that the FDD algorithm can detect ‘leakage’ in the HVAC system considered, which shows the correlation among these two fault categories in the database recorded.

The other six rules mined show how the faults in the HVAC can be interrelated as they have two premises. Rules #12 and #13 have the highest confidence and represent how different combinations of fault in their premises and respective conclusion can be indicative of the correlation between ‘foul’, ’non-condensable’, and ‘flow problems’ fault categories in the given database. Model 1′s mined rules only indicate correlations found in the database for the select FDD algorithms specifically designed to detect the faults investigated and cannot be used to investigate causality or indicate that FDD algorithms were designed to detect faults simultaneously.




```
Association Rules:
1- [Flow problems] --> [Leakage] (confidence: 0.704)
2- [Foul] --> [Limit issue] (confidence: 0.710)
3- [Leakage] --> [Foul] (confidence: 0.714)
4- [Flow problems] --> [Foul] (confidence: 0.741)
5- [Limit issue, Foul] --> [Flow problems] (confidence: 0.773)
6- [Foul, Flow problems] --> [Limit issue] (confidence: 0.850)
7- [Foul, Flow problems] --> [Non-Condensables] (confidence: 0.900)
8- [Limit issue, Flow problems] --> [Foul] (confidence: 0.944)
9- [Non-Condensables] --> [Foul] (confidence: 1.000)
10- [Non-Condensables] --> [Flow problems] (confidence: 1.000)
11- [Non-Condensables] --> [Foul, Flow problems] (confidence: 1.000)
12- [Foul, Non-Condensables] --> [Flow problems] (confidence: 1.000)
13- [Flow problems, Non-Condensables] --> [Foul] (confidence: 1.000)
```

Figure 2.3: Association Rules for the Co-occurrence of Common HVAC Faults for the Literature Analyzed

A second ARM model is developed for faults and FDD techniques found in the literature to determine the association between the faults and the methods used to detect and diagnose HVAC faults. The accuracy and performance of the FDD methods are not considered, and only their quantitative adoption in the literature is considered as a measure for the effectiveness of an FDD algorithm for detecting certain fault types. The support for FDD techniques was determined and selected to understand how frequently the items for the methods under investigation appear in the dataset.

For the complete list of algorithms considered in this thesis please refer to Appendix B. The followings are the frequency of occurrence, i.e., the support, of different analysis methods in our database: 20% for SVM (Support vector machine), 19% for ANN (Artificial Neural Network) and 17% for dimensionality reduction techniques, and 11% for Bayesian networks. A minimum support of 2% and minimum confidence of 50% were selected for model 2, which is appropriate when compared to the highest support (20%) found, which is indicative of a limited number of algorithms in our database. Setting lower thresholds for the second model leads to the generation of a large number of rules that need syntactic constraints to prune and only show the associated faults and methods. The rules found for 100% confidence are removed at the 2% support, as this was considered an indicator of the availability of a few examples, and hence may not represent useful rules. We further limited the rules to those with a single item in their conclusion, which should belong to one of the FDD techniques.

A total of 16,703 rules were mined before being pruned (an excerpt of which is shown in Figure 2.4). Four methods, namely SVM, ANN, dimensionality reduction, and decision tree (with a confidence of 50%, 67%, 50%, and 67%, respectively) resulted in forming 12 rules where eight rules belong to SVM; two rules were found for ANN, and dimensionality reduction and decision tree have one rule each. Other than rule #9, which belongs to ANN, all rules have more than one item in their premises.

SVM Confidence = 50%

```
1- [Leakage, Control] --> [SVM]
2- [Stuck/partially closed, Temperature issue] --> [SVM]
3- [Stuck/partially closed, Control] --> [SVM]
4- [Control, Speed] --> [SVM]
5- [Leakage, Stuck/partially closed, Control] --> [SVM]
6- [Leakage, Control, Speed] --> [SVM]
7- [Stuck/partially closed, Control, Speed] --> [SVM]
8- [Leakage, Stuck/partially closed, Control, Speed] ->[SVM]
```

Neural Networks/Deep Learning Confidence = 67%

```
9- [Set point] --> [Neural Networks/Deep Learning]
10- [bias/drift/calibration, Control] --> [Neural Networks/Deep
Learning])
```

Dimensionality reduction Confidence = 50%

```
11- [Foul, Other faults] --> [Dimensionality reduction]
```

Decision tree Confidence = 67%

```
12- [Non-functioning, Speed] --> [Decision tree]
```

Figure 2.4: Excerpt of a Deductive Self-organizing Graph (ISOM) for Rules Generated for Techniques Used for Different HVAC Fault Types (minimum support = 2%)

At 50% confidence, the rules consist of the following fault categories, namely ‘leakage’, ‘control’, ‘stuck/partially closed’, and ‘speed’, which when combined form rule #9. The initial four rules have two premises made up of the combination of these fault categories, and rules #6 and #7 have three fault categories in their premises which are detected using the SVM algorithm.

The ANN algorithms are found to be utilized for diagnosing the ‘set point’ faults or a combination of sensor-related issues and the ‘control’ category of faults. Rule #11 indicates the applicability of dimensionality reduction techniques when ‘Foul’ or ‘other faults’ are found together in our database. The decision tree technique, which has the highest joint confidence of 67%, is used for detecting the ‘non-functioning’ and ‘speed’ categories of faults.

The findings of this study merely indicate how specific types of faults are often addressed in the sampled research literature, using specific types of algorithms, and they do not provide information on the actual co-occurrence of the faults in building mechanical systems, nor on the performance of the data-driven algorithms with respect to faults.

The rules discovered through the first model, i.e., the association among HVAC’s most common faults, were validated by taking experts’ opinions through structured surveys and are validated in the context of the data gathered from the academic papers reviewed to reflect experts’ opinions. The survey contained thirteen questions corresponding to the rules detected by the first model.

The correspondents were given the Likert scale anchors for the frequency of use, i.e., ‘never’, ‘almost never’, ‘occasionally/sometimes’, ‘almost every time’, and ‘every time’. In addition to these, an ‘I do not know’ option was added to reduce the uncertainty resulting from enforcing the respondents to answer all the questions in the survey. The survey was made available to respondents with HVAC and FDD experience in the industry or those with relevant research background expertise and was made available for two months. It must be noted that later in the study, another survey is conducted to validate the proposed method (ontology) using a semi-structured survey.

Additionally, model 2 provides four FDD algorithms and their associated faults in the form of a set of rules that allows the asset managers to decide on the type of algorithm that can be selected for AFDD of the HVAC system faults. For example, the SVM algorithm is found to be effective in FDD when fault types belong to ‘leakage’, ‘stuck/partially closed’, and ‘control’ issues. It was found that some algorithms are used more often for detecting particular faults. The algorithms that can be utilized for each category of the top six common HVAC faults separately are shown in below Table 2 and are organized in descending order.

Table 2: Recommend Algorithms for an Individual Category of HVAC Systems Most Common Faults

Fault category	Recommended algorithms
Limit issue	SVM – ANN - BN
Stuck/Partially closed	ANN – SVM - DT
Flow problems	ANN – SVM - BN
Bias/Drift/Calibration	ANN - Dimensionality reduction methods - SVM
Leakage	SVM – ANN - Dimensionality reduction methods
Foul	SVM – ANN - Dimensionality reduction methods

The set of rules extracted through association mining techniques, combined with taxonomies developed from the literature analysis on HVAC types, algorithms, and features, serve as the input for the proposed knowledge model for Automated Fault Detection and Diagnostics (AFDD) in HVAC systems. While this approach effectively captures the intricacies of HVAC systems and associated faults, it lacks spatial and semantic information related to the building. The next section of this thesis addresses the Building Information Modeling (BIM) approach adopted to access such information, which is then integrated to capture the relationships comprehensively. This integrated model is discussed in greater detail in a subsequent section of this chapter.

2.4 Knowledge Management with Ontologies - Case of BIM-based AFDD of HVAC

The semantic web is a vision advocated by W3C as a web of linked data. Semantic web technologies allow the creation and linking of data collections known as user data stores, building vocabularies, and writing rules for handling data [61] [62]. The Resource Description Framework (RDF) is the standard model for data interchange on the web that allows for modeling graphs of resources over the web.

The RDF triples consist of three parts: subject, object, and predicate, as shown in Figure 2.5. When considering relationships, the subject is the source node, and the object is the destination node for a directed edge; the label of the edge is the name of the relationship node [63]. For example, the class “Algorithm” is *Subjec*t, and the class “Fault” is the *Object*. The Algorithm and Fault Classes are connected using an *object property* “canDetect” which is the predicate in the triple.

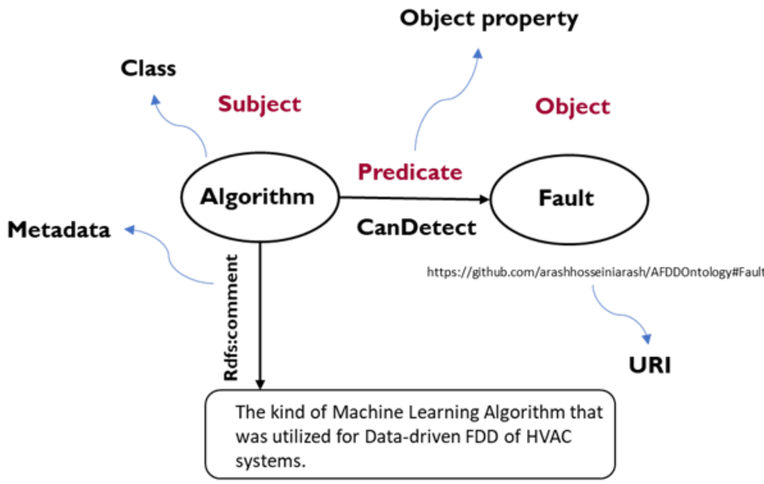


Figure 2.5: Resource Description Framework (RDF) Triple

Ontologies are structured representations of knowledge that consist of individual entities, called instances, and the relationships that exist between them. Figure 2.6. illustrates the instances denoted as circles and the relationships denoted as arrows, which are used as a building block of an ontology. To begin with, an entity (instance) can be considered an abstraction of any physical, logical, or virtual item, the actual ‘things’ in a building. Sets of attributes show what kind of entity it is and what concepts (classes) it represents [63] [64].

Secondly, Terminology is a set of terms that belongs to one unique language which denotes a particular discipline. Finally, Taxonomy is not only a unified vocabulary but a semantic model of how to categorize domain concepts and their interdependencies. Taxonomy is a hierarchy of terms and ontologies set up as a network [65] [66] [67]. Ontology is a set of formal names, concepts, definitions, and relationships that constitute the knowledge domain. Ontology class hierarchy allows defining rules to resolve ambiguities, such as specifying synonyms in the ontology as “equivalent classes” [63] [64] [68].

The OWL (Web Ontology Language) is an ontology language for the Semantic Web with a formally defined meaning. OWL provides classes, properties, individuals, and data values that can be stored and used with information written in RDF. OWLs themselves are primarily exchanged as RDF documents [69] [70]. An ontology can also imply relationships not explicitly expressed in the model and influence a query processor’s interpretation of relationships. To populate a newly developed ontology, instances can be brought in from other ontologies by a procedure called mapping and matching [71] [72], or their entire class can be brought in, which is often referred to as ontology reuse [66].

Implementing semantic web technologies will allow the server host to combine the building and HVAC asset data to respond to requests through the proposed ontology. The applications that utilize linked data define the relationships among readable and accessible data by semantic web tools [73] and can follow the RDF links by accessing the ontology developed through their unique identifiers [74]. The ABOX (Assertion BOX) represents the knowledge model with assertions and individuals, and the TBOX (Terminological Box) refers to the ontology representation of concepts and relationships [75].

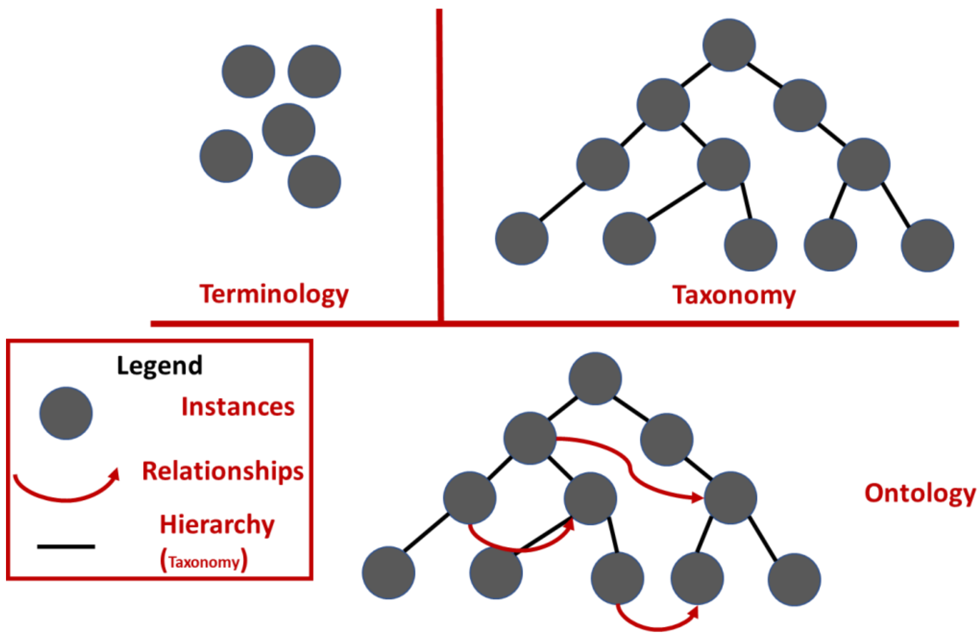


Figure 2.6: Structure of an Ontology

BIM is proposed by ASHRAE to be applicable during the entire lifecycle of the building and defines how to incorporate BIM requirements within the project using standards [76]. The IFC schema is commonly utilized in ontology development studies as the primary source of BIM data [77] [78] [79]. During the facility management phase of a building, the IFC format is used to store objects and related information or COBie (Construction-Operations Building information exchange) data which is a subset of IFC for facility management and maintenance work [77].

BIM provides a basis for human reasoning through visual analytics to facilitate fault detection and diagnostic processes [80] [81]. The schema from IFC can be translated to OWL [77] [79] or IFCOWL [82]; however, the semantic relationship is limited to inferred knowledge representation as different schemas are being utilized [77]. The spatial classes residing in the IFC schema can be extracted from the IFC file to a proposed ontology automatically and systematically using the value of property ID [77]. Ontology approaches for BIM allow the development of the Information

Delivery Manual (IDM), which is manually defined by industry experts and is transferred and stored in OWL/XML that can be ultimately translated to a specific Model view definition (MVD) [79].

According to the literature [83], the current applications of semantic web technologies in AEC (Architecture, Engineering, and Construction) can be broken down into three distinct groups: Interoperability, Linking domain, and Logical inferences. Within the scope of this study, which investigates the ontologies and their application for HVAC FDD, the main areas of research are automation, configuration, selection, deployment, and evaluation of services for FDD [82] [84] [85] [86]. The following sections will explore ontologies pertinent to the application of HVAC FDD in buildings. Ontologies and rule sets enable mechanisms to reason in time, deduce spatial information, and detect and diagnose faults [69]. Ontologies can be broadly classified into domain-specific and domain-independent ontologies [69]. The ontologies that capture HVAC and building concepts provide modeling support for HVAC systems, sensor systems, and spatial information, which are vital for AFDD [87].

Project Haystack, as a domain-specific example, is one of the initial ontology attempts in the building domain that allowed its users to tag objects such as HVAC or other systems. It is an open-source initiative that enables its users to work with IoT data [88]. An example of a domain-independent ontology is the SAREF (Smart Appliances REFerence) ontology [89], which can be used to define HVAC and sensor data.

The Brick ontology [76] is an example of an ontology explicitly made for the building's operation stage and contains most of the building's asset information, including HVAC. However, there was no automatic conversion from IFC to Brick at the time of research. Google's Digital Buildings [90] is an example of an ontology specifically created for an organization, comprising more than one existing ontology. This ontology is inspired by the Brick ontology and Project Haystack, aiming for human readability and cross-compatibility.

Domain-specific ontologies, such as the industry EXPRESS IFC schema, are the primary source of building information throughout the lifecycle. Industry Foundation Classes (IFC) is a conceptual schema language that defines the building and construction domain and is introduced by buildingSMART International. The schema serves as an open data standard for BIM, enabling the exchange of information between various software tools through a software-agnostic data schema.

To ensure interoperability, scalability, and readability, the IFC schema can be encoded in different formats, including IFCOWL, which is the expression of IFC in the Ontology Web Language (OWL). This leverages semantic web and linked data technologies to represent and share building data. The SPC schema can be converted to IFCOWL using existing converters [91]. While IFC is a general schema, modular ontologies represent a subset of a broader concept that can be reused in other ontologies.

For example, the Building Topology Ontology (BOT) is a lightweight ontology that includes only essential topological concepts of buildings. BOT uses IFC subset concepts, such as site, building, story, space, and zone classes, as well as the relationships between them [92]. Another relevant ontology is the Ontology for Property Management (OPM), which enables tracking the history, reliability, and provenance of a property of some feature of interest (FOI) in three main categories: spatial elements (spaces, zones, stories), physical elements (walls, windows, heaters, and sensors), and abstract elements (interfaces, systems, and concepts) to manage interrelated projects in AEC [86]. Modular ontologies are preferred for ontology development to maintain compatibility.

Table 3. compares relevant ontologies reviewed in this thesis for the research work, focusing on their application for FDD in HVAC systems. The comparison is based on different modeling supports, indicating which ontologies cover the necessary concepts such as HVAC, spatial, operational, state, and FDD-related concepts. The symbol (✓) represents the presence of modeling support by an investigated ontology, while the symbol (X) indicates a lack of modeling support from the investigated ontology.

Table 3: Comparison of Ontologies Applicable to the Operation Phase of Buildings. Adapted from [93]

Modeling Support	IFC	Brick	Digital Buildings	BOT	OPM
HVAC Systems	✓	✓	✓	X	X
Spatial Information	✓	✓	✓	✓	X
Operational Relationships	✓	✓	✓	X	X
State	X	X	X	X	✓
FDD	X	X	X	X	X

The analysis signified that no ontology during this research could provide the modeling support required for AFDD. Some modeling supports, such as spatial information, are widely adopted, and some, such as state-related information, are found exclusive. No existing ontology covers the modeling support for AFDD of HVAC. Hence, an ontology is required for AFDD model development and analytics.

Ontologies can be developed in response to competency questions identified and be further extended to include new competency questions such as a description of property reliability, reasoning logic for derived properties, and parametric queries [94]. As such BIM is identified as a solution that encompasses the missing modelling support required for AFDD of HVAC.

The primary reason for adopting IFCOWL in this study is to establish connections between building and HVAC-related data. By utilizing IFCOWL, the proposed model can capture and extend the knowledge necessary for developing AFDD models, which is currently missing from the existing IFC schema [95] [96]. The proposed solution should enable users to perform AFDD of HVAC systems using BIM and machine learning concepts, encompassing the information needed for the development or comparison of HVAC AFDD models.

The FDD modeling support captures knowledge of FDD analytics and parameters, such as the type of algorithm for FDD, the parameters used for model development, and the corresponding values required for model comparison and development. Presently, none of the ontologies investigated are aimed specifically at the FDD of HVAC or AFDD and generally have a broader scope, making them unsuitable for direct adoption for HVAC FDD.

The analytical tools can access the BIM and BAS data by performing a suitable query from the ontology, enabling access through graph-based queries such as SPARQL (Simple Protocol and RDF Query Language). Further, the SPARQL server can perform queries, export, add/remove data over the web [82] [84], or build applications that can be developed based on query engines that support OWL-based ontology. For example, the ARQ query engine can be used with Jena API, which can return results of FM information, such as work types, IFC objects (e.g., spaces and building elements), and facility properties from COBie in XML, which is easy to use [77].

Further, BIM can be enriched with rules (expressed in SWRLS) to allow reasoning and deriving logical inferences; logical IF-THEN rules are specified and executed by a reasoner. This approach is reported to be suitable for identifying faults' root causes and assessing their consequences[85]. Evidence from the literature suggests that data and information from BIM can be helpful for FDD in HVAC; however, the information is often disconnected from other data systems like BMS/BAS, which contain data needed for AFDD model development and do not get updated with the results of the analytics.

BIM provides access to the knowledge model, enabling examining the as-is model and gathering maintenance-related information. Sensory data stored in BMS/BAS can be integrated with BIM using different methods [97], which can be broadly categorized as Integration using databases, Linking to resources, and ontology-based integration [45] [46] [47]. Using database integration methods, the BIM and BAS/BMS data is stored in a central repository [101] [102]. In contrast, in linking to resources, the databases are provided with connections to retrieve data as needed [97] [103].

The ontology-based approach offers a solution in the form of a knowledge model to capture semantic relationships and context, enabling intelligent data integration, knowledge representation, and reasoning [104] [105]. This approach can enhance the accuracy and effectiveness of AFDD systems by facilitating context-aware fault detection and diagnosis. Ontologies can be created or reused and extended to meet the user-identified requirements.

This chapter outlines the development of the AFDD model, beginning with feature inputs and fault types, from which association rules and a taxonomy were derived. These elements will be captured using ontologies as part of a knowledge management system. Chapter 3 will explore how dynamic BIM can be utilized for AFDD of HVAC systems, where the integration of BIM concepts and relationships will provide essential contextual information.

Subsequently, Chapter 4 will develop an ontology specifically tailored for AFDD of HVAC systems, leveraging BIM and Machine Learning concepts. This ontology will draw on information gathered and classified through literature analysis, enabling AFDD systems to incorporate building context in the form of spatial semantics.

2.5 Gaps in the Literature

The primary challenge facing current AFDD models for HVAC systems is the lack of contextual information and limited sensory data. This deficiency hinders the ability of HVAC modelers to use these models to understand the complex interactions within buildings. Additionally limited sensory data means only certain faults can be detected while others may go un-noticed. Without incorporating crucial data related to building layout, assets, and spatial relationships, AFDD models struggle to provide meaningful insights. Their reliance on only sensory data (Specific types) from BMS and BAS, which often lack this contextual and semantic depth, only exacerbates the problem. Furthermore, the absence of a methodology to enable the bi-directional flow of contextual information limits the broader adoption and effectiveness of AFDD solutions.

In contrast, as shown in Figure 2.7, BIM can contain information such as HVAC schedules, building parameters, and supplementary information pertaining to building and HVAC that can help with AFDD of HVAC. At present, the concepts required by BIM to create an effective AFDD model are not integrated with BMS data sources and, hence, do not communicate with each other. As a result, AFDD models built using sensory data often lack semantics related to building spaces, HVAC information, and environment.

The AFDD of HVAC can be facilitated through knowledge models that contain the information needed for AFDD model development using BIM and BMS-infused features. Such a knowledge model can capture the dynamism of the facility to attain connectivity between time-series data and context-representative data. Moreover, the data stored in a knowledge model can be leveraged to create and access AFDD models for buildings that have comparable HVAC systems or similar building envelopes. Such models can be employed to keep a record of modifications, which can help to assist in diagnostic tasks, such as visual inspections.

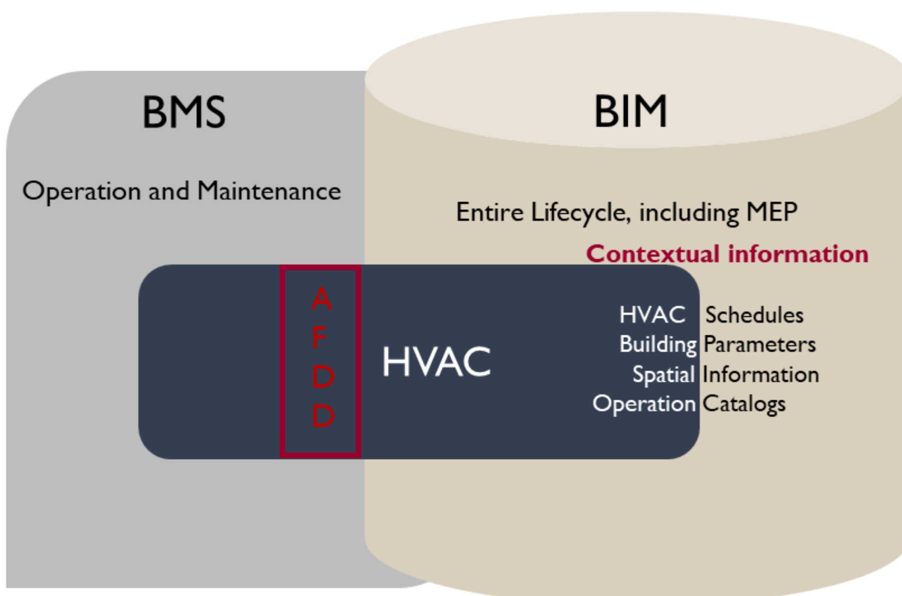


Figure 2.7: Data-driven FDD (AFDD) of HAVC Enriched with BIM Contextual Information

To address these challenges, literature analysis indicates one effective approach can be integrating BIM features with BMS/BAS to enhance AFDD by utilizing BIM-based knowledge models. BIM-based knowledge models are advocated by literature as a viable solution to overcome the shortcoming of HVAC AFDD models by effectively using BIM-compatible ontologies for HVAC FDD [16] [17] [18] [19] [20].

However, a complete methodology is lacking since the solutions are either in the early stages of research or in conceptual form. To bridge these gaps, a comprehensive approach involving the integration of BIM throughout the AFDD process is required. BIM can be extended to support HVAC AFDD analytics by the addition of dynamic BIM features and facilitating the bi-directional flow of data and knowledge between BIM and BAS. Effective integration of BIM and BMS systems is essential to unlocking valuable contextual resources, improving fault detection and diagnostics capabilities, and ultimately enhancing building energy efficiency and occupant comfort.

2.6 Scope of the Study

The scope of this study is confined to purely data-driven techniques that utilize sensory data available in BMS or BAS systems to build their models and do not pose context related to building, HVAC, environment, spatial, and occupant-related information. i.e., a subset of AFDD models for HVAC is investigated (based on literature analyzed), which pertains to data-driven models utilized for fault detection and diagnostics of faults at the system, sub-system, equipment, and parts levels in HVAC, in which machine learning and data mining techniques are commonly used.

In this study, BIM is selected as the primary contextual resource to enrich AFDD models. To enhance interoperability, BIM is extended through an IFC schema-based ontology, specifically developed as AFDDOnto [121]. AFDDOnto functions as a BIM-based knowledge management system designed to support data-driven FDD for HVAC systems, integrating concepts from IFC, BOT, and Brick.

As defined earlier, scope of this research is primarily focused on HVAC systems of medium to large commercial buildings, as most of the relevant literature pertains to this category. The proposed methodology is tested and validated in a controlled facility due to limitations in the availability of a BIM model and suitable data for AFDD. However, it is important to note that, as the ontology is based on existing literature, not all types of faults, HVAC systems, and their relationships may be fully captured.

Chapter 3: AFDD of HAVC as a Dynamic BIM Use

In response to the gap identified in the previous chapter regarding the lack of access to contextual information and limited sensory data in AFDD models, this chapter focuses on three main areas: the role of dynamic BIM, integration approaches for sensory data typically found in BMS/ BAS, and the methods to achieve such dynamism. To address these areas, various integration methods and schemas are identified to facilitate the development of a BIM-based solution for AFDD of HVAC systems.

Specifically, this chapter explores the integration of sensory data from BMS/BAS, the utilization of dynamic BIM models, and the schemas required to develop and implement a dynamic BIM solution for AFDD of HVAC systems. The findings from this chapter, along with those from Chapter 2, will be further utilized in Chapters 4 and 5.

3.1 Dynamic BIM – Case of FDD of HVAC

In this study, BIM is categorized into static and dynamic models, with dynamic BIM further divided into Live BIM and Stateful BIM [122] [123]. Live BIM represents an evolving model in near-real time, while Stateful BIM tracks the entities of interest over a predefined term. The Live BIM refers to the real-time overlay of sensory data that represents dynamic changes in the BIM model. This includes changes to entities such as objects, spaces, and occupants.

Conversely, stateful BIM uses the data captured from Live BIM over time to represent the current state of entities, such as the state of HVAC equipment. The sensory data for capturing the dynamism is sourced from BMS/BAS systems. This integration creates a unified system that reduces duplication of data for different use cases and also prevents overpopulating the BIM by only storing stateful information and retaining links for live data, which prevents the need for large BIM files [108].

Within the context of FDD, BAS/BMS sources focus on data-centric information, while BIM sources provide context-oriented information. The realization of dynamic BIM involves capturing analytics performed on BAS/BMS and BIM data and storing it as DT entities. For example, as shown in Figure 3.1, linking BMS-recorded temperature sensory data to associated BIM spaces enables the capture of dynamic information as Live BIM. In addition to the stream of data (Time-series), some sensors are state representatives (Binary) and can indicate states such as air conditioning status, closure/opening of doors, or windows at any given time by maintaining links between BIM elements and corresponding sensors being in the BMS/BAS. Additionally, Stateful BIM is achieved by tracking the state of entities such as rooms and associated states (e.g., cooling, ventilation, heating) over predefined intervals (e.g., HVAC maintenance schedule) and recording them in the model. Access to dynamic BIM facilitates various use cases, including diagnostics, life expectancy estimation, predictive maintenance, HVAC condition assessment, and feature engineering purposes.

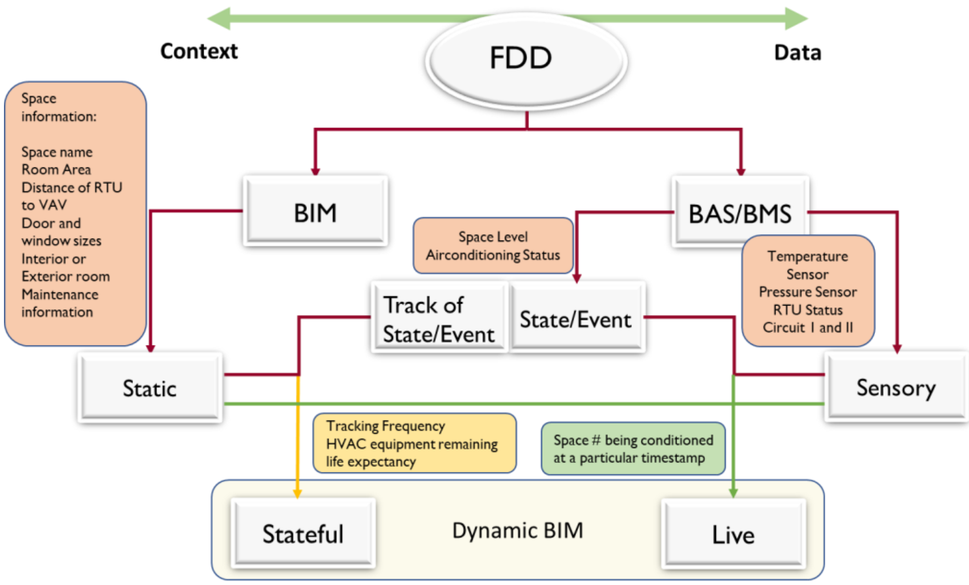


Figure 3.1: Dynamism in BIM for FDD of HVAC

To fully utilize BIM for asset management use cases, it is necessary to augment BIM with provisions that capture and store knowledge about the state or events for a defined time interval. The specific categories, types, and frequencies of states depend on the identified use case and the defined competency or task. The transfer rate of analytics from the knowledge model must be predetermined to create a dynamic BIM of a facility. This transfer frequency is crucial for upgrading BIM from a static model to an accurate representation of the facility's current state, enabling case-specific solutions during the operation and maintenance phase. When coupled with real-time data, this dynamic form of BIM can provide access to the most recent state of the facility essential for various operational applications, including AFDD of HVAC systems.

However, in practice, BIM is mostly used as a static model, and BAS/BMS systems that store sensor data are utilized independently, without any connection between them. Hence, BIM in its current form cannot be utilized to add additional contextual, static, or dynamic features to that of BMS features for AFDD model development. In this study, AFDDOnto [109], which is a BIM-based ontology, is used as a knowledge model to capture AFDD analytics and is further used for BIM-based feature engineering to facilitate the effectiveness of AFDD for building HVAC systems.

The overarching goal of this study is to develop a BIM-based knowledge model of the building, specifically the mechanical and architectural systems, to facilitate the fault detection and diagnosis process. This goal is followed under two major research directions. The first direction is (*BIM-to-AFDD*) aims to enrich the operation data with static (and dynamic) information of the building systems, from BIM, and upgrade the data-driven AFDD for HVAC systems.

The second direction is (*AFDD-to-BIM*) uses the BIM, as a semantic and granular model of the building elements and spaces, to visualize the results of detected and diagnosed faults. To achieve these objectives, a methodology will be introduced that enables BIM to contribute additional context-aware features and link them with the BMS data for an improved data-driven AFDD. Furthermore, the study uses ontology concepts in the form of a knowledge model to capture knowledge related to fault detection and diagnostics of HVAC.

3.2 Dynamic BIM Enablers

The past ten years (2009-2019) studies were reviewed and analyzed that have utilized BIM in its dynamic and have demonstrated the use case using a case study or approach. The studies are sourced from Concordia University library that allows access to multiple databases such as ASCE, Science direct and Google scholar, to name a few that are used in this study. Further different attributes such as application, technology, variable and building type, frequency and Year are recorded. Based on the most significant application of study, the BIM uses are selected and recorded.

Overall, eighty-nine studies were selected that have implemented dynamic BIM. In Appendix C, the table is provided. BIM use is defined by Penn State as “a method of applying Building Information Modeling during a facility’s lifecycle to achieve one or more specific objectives” [110]. In this study, the methodology encompasses several detailed steps to ensure a comprehensive analysis of BIM and its dynamic applications as shown in Figure 3.2: .

Initially, specific terminology used throughout the study is defined for clarity. A survey of BIM use guidelines is conducted to gather existing practices. Using these guidelines, a unified BIM use list is developed (Appendix D), considering the entire building lifecycle. The competency of dynamic BIM uses is examined through a review of literature and guidelines, and further established through brainstorming sessions with BIM practitioners. Sensory data that make each BIM use dynamic and stateful are identified through literature review.

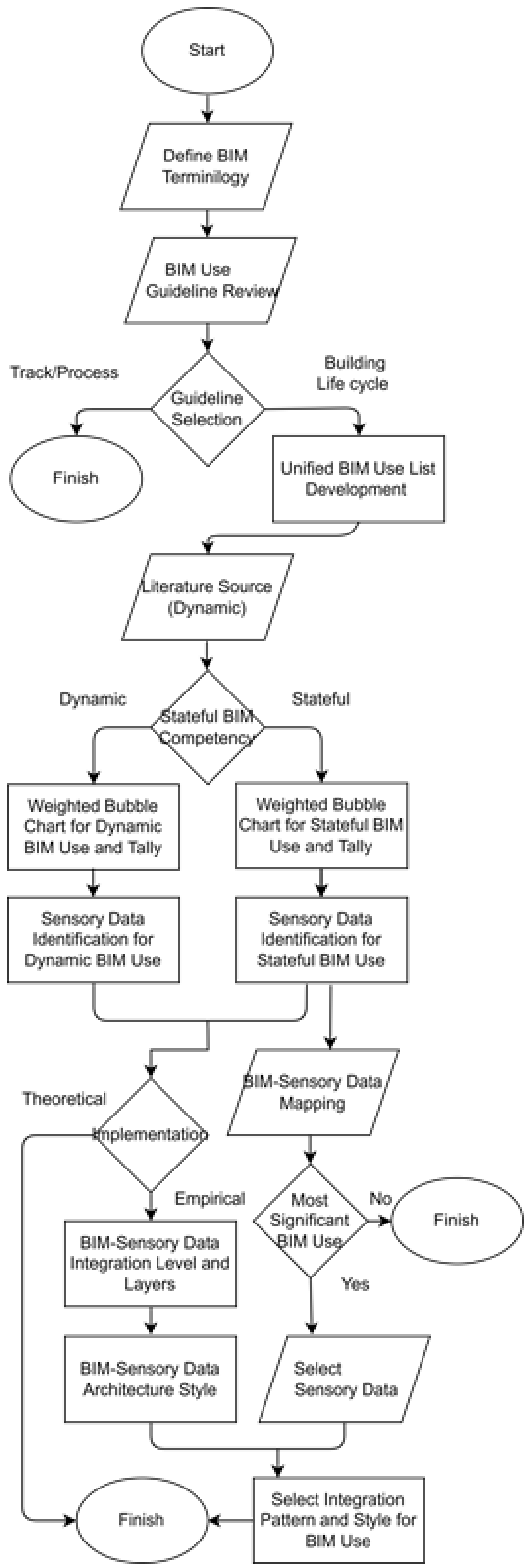


Figure 3.2: The Framework of the Methodology Adopted to Identify Dynamic BIM-sensory Data Integration

The sensory data that are the enablers of dynamic BIM are identified for each of BIM uses. The BIM Uses identified are further investigated for types of data (variables) that are related and required to add dynamism to BIM. As shown in Figure 3.3, The most significant enablers are Temperature followed by Location/Traction and Humidity.

Sensory data such as CO2 and CO are associated with time-series data types and are suitable for Live BIM applications where the latest reading of sensory data can be beneficial. In contrast, Dynamic BIM uses can use sensory data enablers such as Close/Open or Count for both Live BIM and Stateful BIM applications depending on the intended use case. The requirements can range from needing only the most recent value to tracking the changes over a specified period. Calculated measures effectively transform time-series data to states that can be further utilized by dynamic BIM applications.

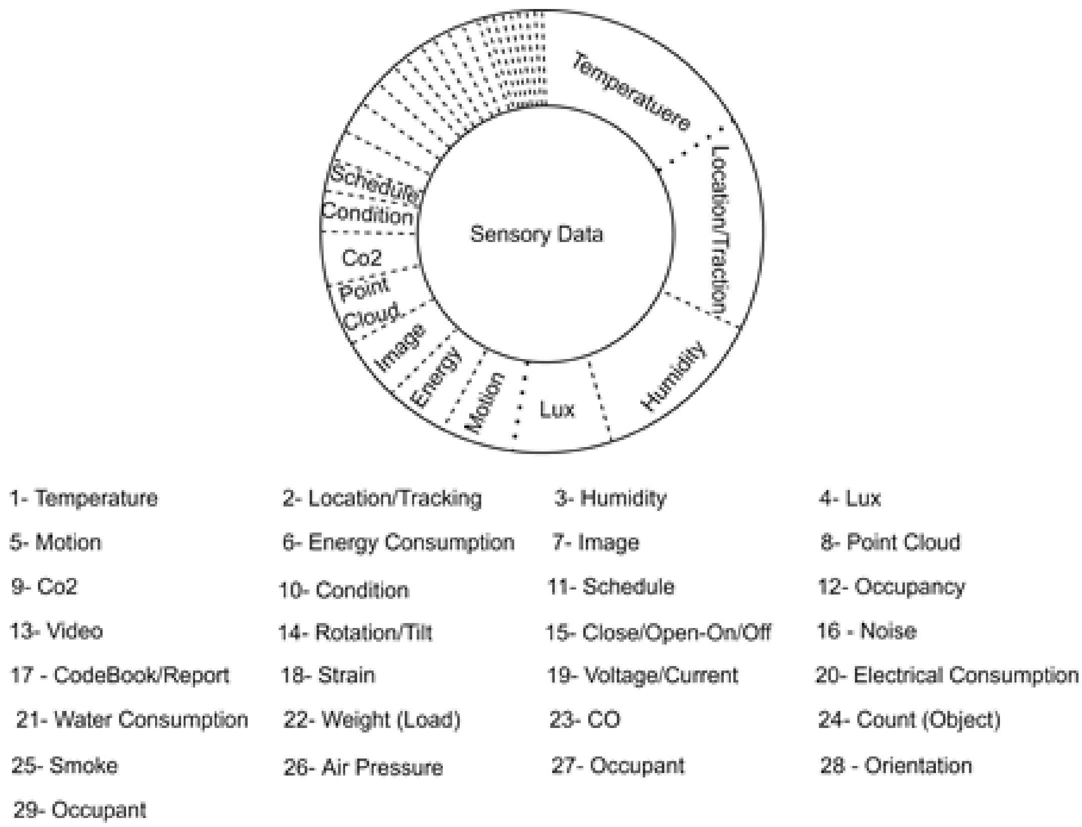


Figure 3.3: Dynamic BIM Enablers (Built Environment Sensory data)

In practice, sensory data may be integrated into BIM by a variety of direct and indirect sources, including IoT devices and existing dedicated systems like BMS, building energy management systems (BEMS), and BAS [127] [128] [129] [130] [131] [132] [133]. Sensory information is the primary enabler of the dynamic building information model. The below graphic

(Figure 3.4) illustrates an interactive Sankey diagram for dynamic BIM applications and the sensory data enablers for the required dynamism.

This graph captures the overall relationship between BIM uses and sensory data. The BIM uses significance is shown concerning the variety and the percentage data type, meaning that each link is weighted to reflect how frequently the sensory data is used for the specified BIM use; this can be viewed as an indicator of the value or significance of the sensory data for the selected BIM use.

Asset management, which accounted for 35% of the total examined papers with possible dynamic BIM applications, was found to use temperature, location, and humidity as the top three sensory data to achieve the BIM application. According to the research, temperature (17%) is the most prevalent sensory data enabling the dynamic BIM, with location data (16%) as the second significant enabler. The interactive Sankey diagram is accessible online[2].

[2] Interactive Sankey Diagram of an aggregated weighted view of dynamic BIM and their enablers (sensory data)

[https://public.tableau.com/profile/arash4461#!/vizhome/Book2_15676945815550/Dashboard1?publish=yes].

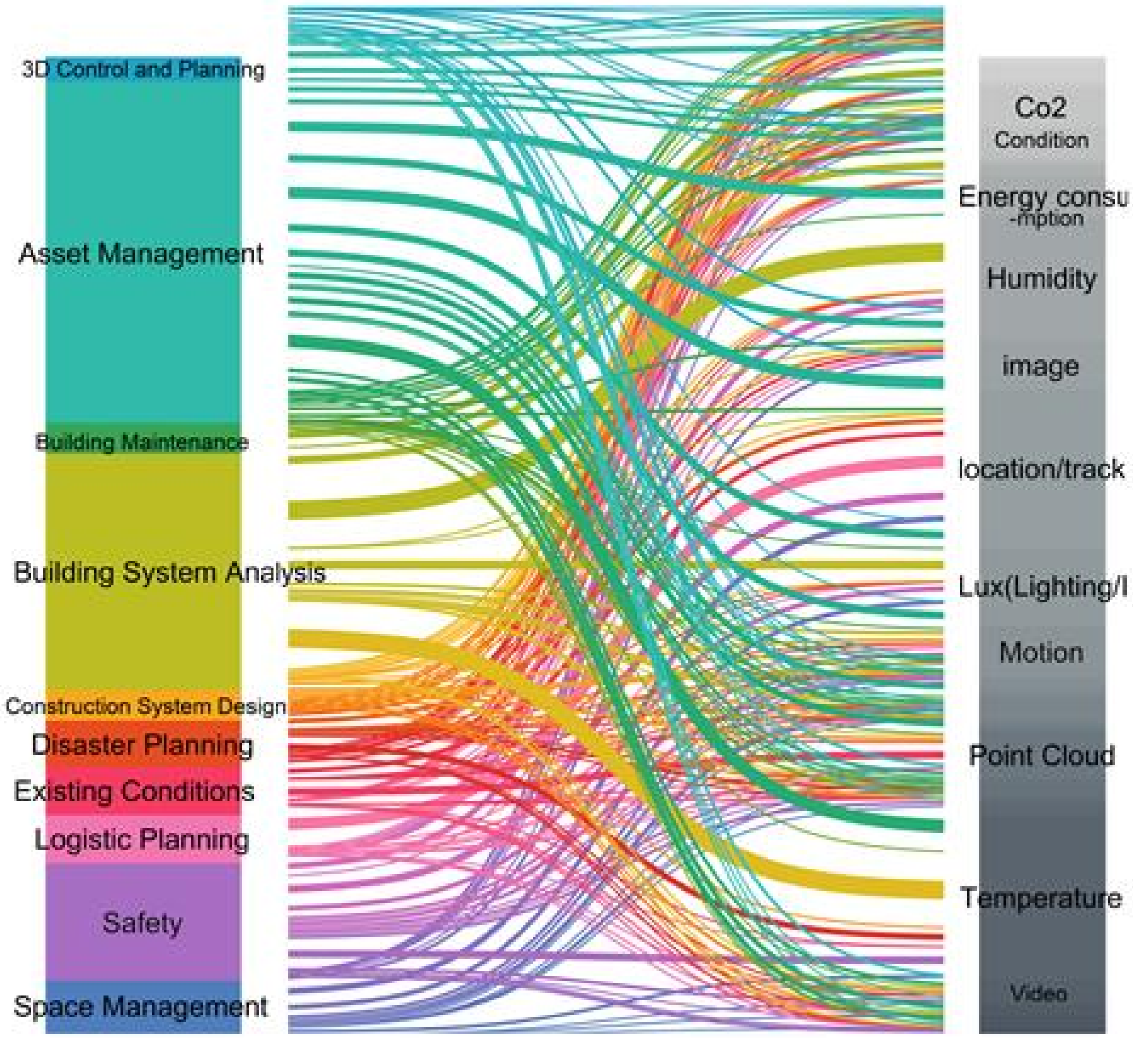


Figure 3.4: Sankey Diagram of an Aggregated Weighted View of Dynamic BIM and Their Enablers (sensory data)

3.3 Integration Methods – Towards Dynamism Utilizing Data from BIM and BAS

In practice, data integration in BIM can be a one-way function [134] [135] [136] [137] [138]; or bi-directional [97]. This functionality implies that in one-way function systems, the information is not retained on the BIM model; therefore, the BIM information is extracted, exported, and used for specific BIM purposes outside the BIM [139] [140]. Most of the reviewed research has examined the one-way function. However, realizing a digital twin of the facility would require the BIM to be stateful and allow for bi-directional communication and operation between BIM and IoT through adopting effective integration mechanisms. Contextual information residing in BIM and state representative data being captured by sensors are the enablers of dynamic BIM.

Integrating information between the building's context and sensory data is accomplished using services to store, integrate, and retrieve data. This integration can facilitate access to data and information or knowledge capture, useful in decision-making. The current data integration practices for BIM-IoT fusion are depicted in Figure 3.5. The system's four key components are BIM, sensory data, query, and visualization. The users can interact with the integrated system through Query and Visualization. In order to get at the data, a query must be executed, and various query languages are available.

For the most part, there are two distinct varieties of query: those explicitly designed for the construction industry, such as BIMQL (BIM Query Language), and more general-purpose ones, like SQL (structured query language), NoSQL (Non-SQL), and SPARQL (SPARQL Protocol and RDF Query Language). Data visualization can be accomplished via tools, plug-ins, or a graphical user interface (GUI). An additional layer, such as a web service layer, is necessary to implement and allow the protocols for data transfer over the web. Three tiers or layers are required for integrating BIM and sensory data, namely Static Data tier, Dynamic data tier, and Service layer tier, that result in one of the integration methods shown in Figure 3.5.

Data integration methods: As indicated in Figure 3.5, there are three methods of integration defined. In the first method identified, the static and dynamic tier data is saved in databases. These databases could be relational [141] [142] or non-relational [127]. Microsoft Access and SQL Servers are the most commonly utilized databases for storing BIM contextual and sensory data [128]. Depending on the objectives and sources, the number of databases used could range from one, to many databases of similar or different structures.

The second method of integration, Involves an ontology to bring all the data from various sources together [145] [146] [147]. At this level of integration, RDF triples are used to turn the data files into ontologies. In this method, each data source, like BIM or sensor data, is stored in RDF format to make a domain ontology. One of the most common ontology schemas for BIM data is IFCOWL [132], which is a formalization of the IFC's subset in OWL (Ontology Web Language). The most common ontology for sensor data is SSN (Semantic Sensor Network). SPARQL can be used to access and retrieve semantic data. Literature indicates such solutions are suitable when interclasses are involved, for example building and physical assets such as HVAC system. However, it is less useful for dealing with data that is streaming in real time [133].

The third integration method is not a separate integration method, but rather a combination of the ontology approach and the database [134]; this decreases the load of data conversion. In a level three integration, time-series data from sensors and semantically characterized building contextual data from BIM are combined. To accomplish this, it is necessary to construct ontologies that contain the existing database schema information [135].

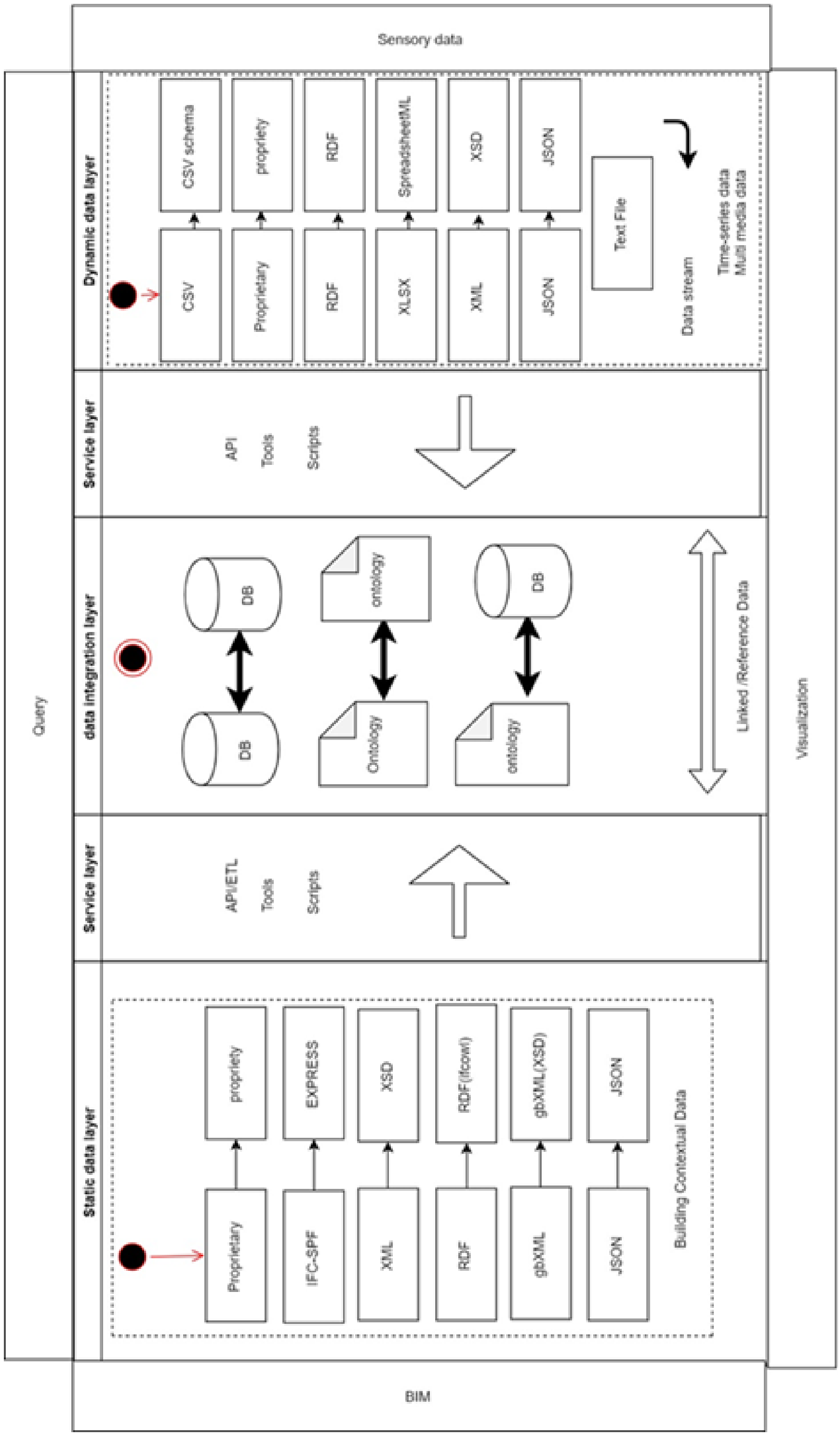


Figure 3.5: BIM-IoT Integration Methods

The literature demonstrates the suitability of such approach for creating web tools or solutions involving data mining solutions for energy management through ontologies [113]. This approach can promote connectivity by connecting BIM with existing resources through linked data, such as data on materials and building systems, profiles of occupants, and information on weather patterns, where are related, connected data can be queried and retrieved [129] [147] [150].

Static data tier: This layer contains the BIM contextual information. Various file formats are identified that can hold the data at the static data tier. However, this diversification often means files are associated with multiple schemas. The files within the static data layer can be either proprietary or open source in nature. Typically, proprietary file formats are only compatible with specific applications, whereas neutral file formats can be transmitted more easily between tools that support such formats. Contextual data can be stored in more than one file based on the intended use. In fact, multiple file types can be used to store the data [136] in one or more files containing overlapping data and information. The adoption of multiple files enables the retention of complete contextual information and sharing of only the information the user requires.

The most common proprietary design authoring tool format used in the studies investigated is Revit with (.rvt) file extension [137] format, which Autodesk develops; additional files from the same vendor are namely (.fbx),(.dwg) and (.dxf) [138] [139] which are used for design data, metadata, and interoperability.

The standard neutral format endorsed by building alliances is IFC (Industry Foundation Classes). The IFC file is based on the EXPRESS schema and is used extensively for data interchange and collaboration [152] [156] [157] [158] [159] [160] [161] [162] [163] [164] [165] [166] [167] [168] [169] [170] [171] [172]. IFC includes definitions for all areas, volumes, and elements of a building [106]. However, additional support is still needed to cover other definitions that BIM uses require. The other prevalent open-source file format is gbXML (green building Extensible Markup Language), which allows engineering analysis tool access to building data.

Regarding the usage of multiple files, gbXML and IFC files can be used by engineering analysis tools to detect HVAC faults [157] by linking HVAC and building envelope elements, surfaces, and zones. Also, users can use IFC alongside Collada (.dae) files, an ISO standard used as an open-source interchange file format for interactive 3D applications to facilitate geometric data sharing. XML [158] and JSON (JavaScript Object Notation) [159] are the other file formats used for data interchange. These schemas offer more flexibility and extensibility and are a choice of communication over the web. However, the reason for the lack of wide adoption is often, such files are lesser expressive when compared to schemas such as EXPRESS [160] .

In practice, presently, proprietary environments are more often used than neutral file formats to avoid sharing information that is not necessary [161] and prevent Information loss. For example, CAD and 3D modeling tools use the (.dxf) file format to exchange information about objects, drawings, and material details. They also use the (.fbx) file format to integrate data such as ambient information with the BIM model to provide a 3D view.

Dynamic data tier: The sensory input can be directly transmitted or temporarily stored for integration purposes. Although time-series data predominates in the construction domain, this layer's data contain other multimedia data types. The type of the data and the frequency depend on the use case intended. The BIM-IoT integrated system can save sensory data in various file formats with relevant schemas. The reviewed literature identifies several prevalent file formats for sensory data, which can be classified into the following categories: data frame formats, text file formats, meta-language formats, and semantic web formats.

The primary difference among these formats lies in how the data is structured and accessed. Data frame storage is similar to databases but is more suitable for subsets of data. In contrast, flat files come with various structures and schemas, employing different mechanisms for data access, retrieval, and storage.

The most commonly used format is CSV, which is used for storing time-series data [106], [161] or schedule information [149]. Proprietary spreadsheet formats, such as XLS format, are used to store data coming from BEMS (Building Energy Management Systems) and CMMS (Computerized Maintenance Management System) [152]. More recently, the studies have used XML-based files called (.XLSX) as a replacement for (.XLS) for storing QTO (Quantity Take-off) data [119].

Alternatively, data can be stored in simple text files (.txt) [148] and RDF(Resource Description Framework), Terse RDF Triple Language (.TTL) [134] The proposed integrated solution flat-files. Dynamic data tier can utilize the exchange data in XML or JSON formats [162]. Instead of storing sensory information in a file or spreadsheet, data can be saved to a database. For example, data from temperature and oxygen sensors, together with their respective timestamps, can be logged in a database [163].

Service layer tier: The service layer facilitates data integration, parsing, and transformation from the static and dynamic data layers to the data integration layer via various means such as APIs (Application programming interface), ETL (Extract, transform, load) processes, tools, and scripts that are used to achieve data parsing, transformation, and mapping.

API serves as a gateway for the data access [141] [152] [166]. The Autodesk DB connection API allows data to be transferred to and from a database [154] [180]. Similarly, JDBC (Java Database Connectivity) is used for applications built in the Java programming language to connect to a database using the JDBC-ODBC (Open Database Connectivity) bridge [148].IFC file can be parsed using an API called JSDAI (Java Standard Data Access Interface) for a STEP (Standard for the Exchange of Product model data) based application; this allows the creation of IFC objects via the parsing and exporting of BIM applications [148].

ETL (Extract, transform, load) is often utilized when there are multiple sources of data or data from one database to a data warehouse needs to be transferred. [171] [172]. Identifiers such as UUID (universally unique identifier), GUID (Globally Unique Identifier), uniqueID (unique identifier), and URI (Uniform Resource Identifier) or unaltered ID are utilized to store and retrieve data at different integration levels [161] [181] [182].

Tools and Scripts can be authored and developed using numerous programming languages such as Python, C#, and JavaScript, as well as graphical programming interfaces such as Dynamo [137], which enables users to write scripts for a proprietary tool. Custom GUI (Graphical user interface)can be developed using programming languages to display room information [163].

3.4 Summary and Discussion

In this thesis based on the scope defined, AFDD of HVAC systems is investigated to highlight its growing adoption by academia and industry, as well as to present the shortcomings that may arise for system users. The goal is to facilitate context-aware fault detection and diagnostics of HVAC systems in specific buildings. The literature analysis of AFDD methods identified the types of features, algorithms, and their associations with HVAC types and fault types.

Additionally, the application of ontologies as knowledge management tools is examined, emphasizing their significance in contributing to the missing context, such as spatial information pertaining to HVAC systems. As such the taxonomy of concepts required for model development and model evaluation are fed to the knowledge model proposed for knowledge management using the analysis performed in chapter 2 of this thesis.

Further, in chapter 3 of this thesis, the use of BIM is explored as an existing solution that contains comprehensive building information and is applicable to the operation and maintenance phases of buildings. As a semantic model, BIM can compensate for the lack of contextual information in AFDD models [110] [111] by the addition of dynamic BIM features. Additionally, BIM can provide spatial relationships, construction details, and system configurations that facilitate the AFDD of HVAC systems. BIM's 3D visualization aids fault localization, while its historical data repository enables baselining, deviation detection, and pattern recognition [78] [112].

Previous studies have leveraged BIM for FDD and combined it with machine learning techniques for enrichment, automation, and comparison [113] [114] [115]. BIM can be utilized throughout different phases of AFDD model development [18] [101]. It supports knowledge model development and configuration before AFDD implementation [13] [116], and during the AFDD development phase, BIM data integration plays a crucial role [117] [118]. The most common application of BIM for AFDD in literature is reported during the implementation phase and involves inferences and diagnostics of HVAC faults [119] [120]. The challenges identified include the black-box nature of current AFDD techniques that lead to loss of information pertaining to relationships between HVAC and the building, the unavailability of BIM in a dynamic form, and the lack of an existing comprehensive ontology. These challenges necessitate the development of a comprehensive framework that utilizes BIM to add context by creating dynamic features that integrate both BAS and BIM.

The availability of BIM as an open standard in the form of an ontology presents an opportunity to develop an ontology that can utilize BIM concepts existing in IFCOWL, along with Machine Learning concepts identified through the analysis. This can be achieved by employing the integration methods identified in the thesis to facilitate a BIM-based solution aimed at AFDD of HVAC systems. A knowledge model connecting BIM and BMS can greatly aid the development of the AFDD model. As an existing solution, BIM can serve as a basis for an efficient knowledge management system. It enables the integration of new applications without the need for a completely new model to be built from scratch.

This framework can be used for AFDD model development and for creating a knowledge model that enables the flow of data between isolated systems. This, in turn, facilitates context-aware model development and provides a mechanism to store model configurations for comparison and reuse. The next chapter will explore how dynamic BIM can be utilized for AFDD of HVAC systems.

Following this, Chapter 4 will focus on the development and deployment of the proposed knowledge model to develop an ontology specifically aimed at AFDD of HVAC systems. This ontology will leverage BIM and Machine Learning concepts, using information collected and classified through literature analysis to facilitate AFDD of HVAC systems. The goal is to enable these systems to benefit from building context in the form of spatial semantics and allow BIM to be used for generating additional context-aware dynamic features that can be used for the AFDD model development of HVAC systems.

Chapter 4: Methodology – BIM-based Knowledge Modeling for AFDD of HVAC [3]

In this chapter, an ontology named "AFDDOnto" is proposed and developed, drawing from BIM and BAS/ BMS resources. AFDDOnto encompasses concepts related to buildings, physical HVAC assets, and AFDD model development. The necessary relationships are captured within AFDDOnto to enable access to competencies identified in this chapter, facilitating the AFDD of HVAC systems. The taxonomy, concepts, and relationships are derived from Chapters 2 and 3, which focus on AFDD of HVAC and BIM, respectively.

Additionally, a framework is introduced to enable the flow of information from BIM to BMS datasets, which are used for AFDD model development. This framework also allows for the capture of knowledge in terms of FDD analytics and model development. This chapter aims to realize a BIM-based semantic model that facilitates the AFDD of HVAC systems in commercial buildings by incorporating dynamic features created using BIM and BMS data

4.1 Knowledge Model Development – AFDDOnto Overview

In this section of thesis, the intention is to create a semantically enabled knowledge model using BIM that can be used to generate AFDD models for an HVAC system, reusable in other projects that can be used for model comparison. The proposed ontology is intended to supplement the present BIM schema (IFC).

Three sub-objectives have been outlined to achieve this goal. 1- To develop a taxonomy for the information required for a semantically enabled AFDD knowledge model using BIM and BMS concepts. 2- To create a BIM-based knowledge model with the necessary axioms for AFDD. 3- To construct data access queries for the proposed AFDDOnto for select competencies defined by literature analysis.

Initially, relevant ontologies are analyzed, and further gaps and scopes of the ontology are defined. Additional concepts and classes are required to ensure the connectivity between BIM and BMS/BAS for AFDD is achieved. Predefined 'competency questions' (or 'competencies' for short) are required that are identified through the literature analysis. The information captured will serve as the foundation for developing an ontology that can be tested and validated. The proposed

[3] The material of this chapter is published in form of the following publications:
1. Hosseini Gourabpasi A, Nik-Bakht M (01 Jul 2024): An ontology for automated fault detection & diagnostics of HVAC using BIM and machine learning concepts, Science and Technology for the Built Environment, https://doi.org/10.1080/23744731.2024.2363104
2. Hosseini Gourabpasi A, Nik-Bakht M. BIM-based automated fault detection and diagnostics of HVAC systems in commercial buildings; Journal of Building Engineering, Volume 87, 15 June 2024, 109022. https://doi.org/10.1016/j.jobe.2024.109022

solution aims to assist its users in two ways. Firstly, integrating BIM with BAS/BMS concepts in the form of a knowledge model to facilitate AFDD model development; Secondly, capturing and updating the BIM-based model with configuration and analytics derived from AFDD models in the AFDDOnto.

Hereafter, descriptions of the terminologies, related definitions, and concepts employed in this study are discussed to make the present work comprehensible without resorting to extra resources. This study adopts and implements a methodology proposed by Fernández et al. (1997) called "Methontology" [178]. The following initial questions are identified and investigated: the ontology's purpose, use cases, and the stakeholders benefiting from them [179]. The results of the analysis were validated using brainstorming sessions. Also, white papers, books, and manuals were used to enrich the knowledge captured.

HVAC system is chosen as the scope of the proposed ontology since it is the highest contributor to energy consumption and maintenance budget and a significant factor affecting occupant comfort within physical asset management. The proposed system's end users could be HVAC Controllers, AI engineers, asset managers, facility managers, BIM managers, and owners.

Given that no single ontology can be relied upon to address all questions pertaining to a given area, through a review of the relevant literature, it was possible to identify the most pressing competency problems for which the proposed AFDDOnto offers a potential remedy; these questions were then translated into the setting of a specific use case to illustrate the ontology's practicality.

Figure 4.1. depicts the procedure followed for developing the proposed ontology, which involves an iterative approach where a pool of competencies is defined based on which taxonomy is developed in parallel to identify the existing ontologies that can be used to answer the competencies that can be re-used from other existing ontologies. In the next stage, ontology axioms in the form of concepts and relationships are added to the ontology, which is then tested and validated. Only necessary axioms in each cycle are retained, and based on the competencies, this development cycle needs to undergo several iterations before a complete ontology is formed that can answer the competencies identified.

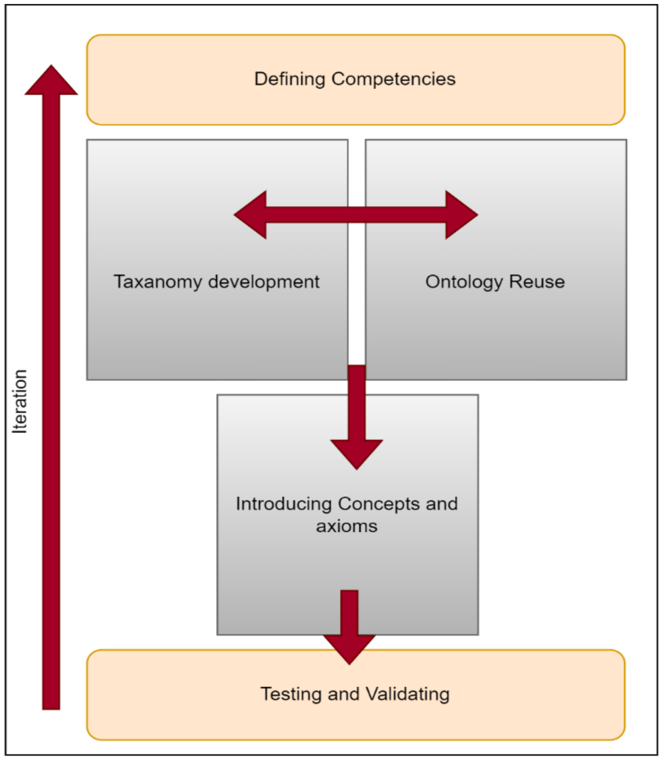


Figure 4.1: The methodology Implemented for the Development of AFDDOnto (generated based on Methontology)

4.2 AFDDOnto Development

The literature analysis conducted in chapter 2 of this thesis is used to identify and categorize the characteristics needed by literature for effective AFDD model development, fault detection, and diagnostics. Further, the areas where BIM can contribute to the AFDD of HVAC were analyzed and used from chapter 3. The characteristics identified by the analysis required by facility managers to develop AFDD models that can utilize BIM to Facilitate AFDD in terms of semantics for fault detection and diagnostics are namely Location, Features, and Parameters used for AFDD, Type of HVAC and Fault, and History. Which can answer to questions such as the following:

- *Location* - Where is the faulty component situated?
- *Type – What type of algorithm(s) is used for detecting and diagnosing the HVAC system?*
- *Parameters -What model parameters are used for AFDD model development?*

- *History* – What changes are made to the AFDD model? The track of evolving entities such as AFDD analytics, including the accuracy and configuration of the produced models (if dynamic)
- *Features* -What type of sensory data is used for detecting the faulty component?

These are the main categories identified as required by facility managers to develop and perform AFDD of HVAC. Hence, the proposed AFDDOnto includes the modeling support required for AFDD of HVAC by capturing the classes and relationships between HVAC types, building spaces, and the AFDD algorithm used for model development to retrieve this information.

As evident, most of the information needed for AFDD indicates that present AFDD models lack semantic information related to space, context, and model information. Hence, a case study will be devised to include the characteristics identified by the literature to allow the capture of such information with the proposed ontology. These characteristics are framed to case-specific competencies to evaluate the proposed AFDDOnto.

The proposed ontology follows a bottom-up approach that defines the ontology classes from a group of instances. The Protégé editor environment [180] for knowledge management is used to create and map the alignments. In this research, the 'subsumption approach is adopted for defining alignments as they offer more flexibility [181]. The knowledge extracted is then transformed into TBOX. The OWL language is used to formalize the ontology since the OWL language can be easily translated to languages that use directed graph structure, such as RDF, that can be easily searched.

The proposed AFDDOnto uses two main sources of information, namely BIM and BMS, which are used for AFDD model development. As shown in Figure 4.2, BIM populates five concepts: Track, Element, Zone, State, and Information. Additionally, BMS populates the Algorithm, Feature, State, and Fault super-classes through the AFDD model analytics. The State concept is dependent on both BIM and BMS. The concepts construct the AFDD modules needed for HVAC FDD. The ontologies reused for the development of AFDDOnto are the Brick (Element Concept) and BOT (Zones), which are reused ontologies, and OPM (State) is used indirectly as adapted for AFDDOnto. The BOT is used as a subset of IFCOWL in the proposed ontology.

For example, BIM information of a building facility can capture HVAC and maintenance-related information using the Element concept that captures building semantic information such as building spaces and zones. Further BMS data used for AFDD model development, and its analytic results are stored in the classes Algorithm, Fault, and Feature. By having AFDD model configuration and analytics integrated with building semantics, the AFDDOnto can retrieve information such as fault type, location concerning the building facility, and features and parameters used for AFDD model development. Additionally, in the case of the dynamic BIM model, the historical information and state changes can be tracked which can enable the facility manager for the AFDD model to be used for comparison, development, and management.

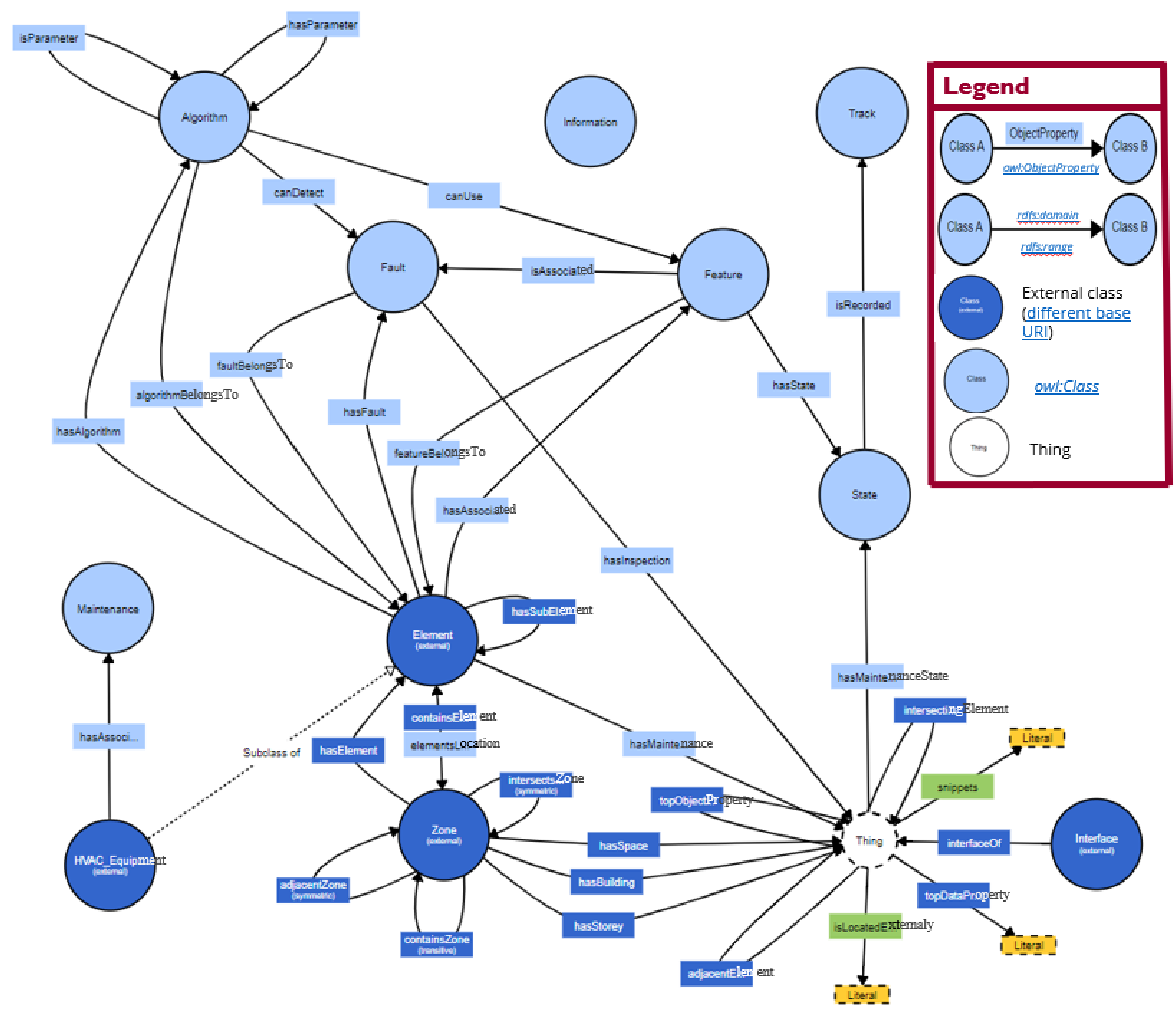


Figure 4.2: The AFDDOnto Lightweight Model

AFFDDOnto includes multiple sub-classes, each associated with a respective superclass in a multi-level structure. The subclasses have all axioms from their parent classes and an additional attribute or restriction that distinguishes them from other subclasses at the same level. However, instances can exist under multiple branches and are not limited to a tree structure. In addition, the definitions of the AFDDOnto axioms are included in Appendix E. The defined relationships enable the enforcement of the defined subject-object relationships in AFDDOnto.

The AFDDOnto is developed over an iterative cycle. The ontology has undergone four main development stages to have the 4th iteration as its most current version. As shown in Table 4, the axioms have been subjected to rigorous revisions to meet the requirements of competency questions. This involves metrics such as logical and declarative axioms, classes, sub-classes, and axioms related to object properties available in AFDDOnto. Metrics such as axioms indicate the number of logical and non-logical axioms, which include subclass relationships, property restrictions, or disjoint-ness assertions.

In contrast, the logical axiom metric only indicates the relationships and constraints based on logical principles in each iterative cycle of the AFDDonto. Further class and object property axiom breakdowns used in AFDDonto are indicated in the below table. The iteration cycle in each stage of ontology development illustrates how the proposed ontology has undergone rigorous changes to construct axioms that can answer the competencies used in this research work. As can be seen, the axiom count until the third iteration has been reduced to include absolutely necessary axioms, and in the fourth iteration, which is the final iteration, it is based on expert feedback that is validated.

Table 4: AFDDOntology Development Stages

AFDDOnt		1st Iteration	2nd Iteration	3rd Iteration	4th Iteration
Metrics					
	No. of Axioms	6,959	3,342	1,402	1,411
	Logical Axiom Count	4,540	1,243	825	830
	Declaration Axiom Count	1,425	917	429	433
	No. of Classes	144	274	256	256
	Object Property Count	26	36	29	31
	Annotation Property Count	14	23	5	5
Class Axioms					
	SubClassOf	155	416	250	250
	DisjointClasses	3	0	9	9
Object Property Axioms					
	SubObjectPropertyOf	2	16	7	7
	InverseObjectProperties	9	6	1	1
	DisjointObjectProperties	0	0	2	2
	TransitiveObjectProperty	0	6	1	1
	SymmetricObjectProperty	0	0	2	2
	ObjectPropertyDomain	4	26	25	27
	ObjectPropertyRange	3	28	24	26

SubPropertyChainOf	0	0	2	2

The AFDDOnto's modular design allows revising, updating, or extending classes, relationships, and axioms to cover other use cases beyond the proposed ontology's scope to cover other built environment domain use cases.

4.3 Knowledge *Model Evaluation*

Verification of AFDDOnto involves the use of several metrics, including Competency, Consistency, Completeness, and Clarity. The Competency metric examines the scope of the ontologies in relation to predefined requirements, while the Consistency metric verifies the ontology's hierarchal information, relationships, and restrictions. Completeness is another metric used to test if the defined competencies enable the ontology to retrieve required individuals. Finally, Clarity examines the ease with which users can interact with the ontology model. Table 5 provides a detailed list of the tools, resources, and approaches used for these metrics in this study.

Table 5: FRP Case Study Metrics and Approaches Implemented

Metrics	Approach	Tool or Resource
Competency	Competency Questions	Case Study
Consistency	Automated Consistency Checking	HermiT Reasoner and Pellet Reasoner
Completeness	SPARQL query	Blazegraph
Clarity	Questionnaire	Semi-structured survey

The proposed AFDDOnto is developed to enable users to retrieve the information needed to answer the competencies defined based on literature analysis, identifying the information needed to facilitate the development of AFDD models as follows: location, type, parameters, history, and features. The case study is used to test and validate these competencies.

Initially, the AFDDOnto is checked with "HermiT," an automated consistency checking tool as a theorem prover that assesses the ontology by inferences. As a result of the consistency test, the ontology iterations underwent multiple checks until no inconsistencies were found. A task-based evaluation was done as the next step in the evaluation process to verify the competency and completeness metrics. This is where the ontology is evaluated based on how well it can be used for the tasks that have been identified. In this study, SPARQL queries were written to retrieve information about the defined competencies and check if the queries were complete. The competencies were constructed using the identified characteristics for the FRP case study. The SPARQL queries were able to retrieve information that was intended and, hence, indicate the completeness of the ontology against the competencies defined.

In a criteria-based evaluation, several criteria, such as Competency, Completeness, Conciseness, and Consistency, are chosen and grouped based on previous validation outcomes. The final metric in the evaluation process of the proposed knowledge model is using a questionnaire in the form of a survey that utilized domain experts to validate the applicability of the AFDDonto against all metrics and specifically the clarity in terms of capturing the concepts needed for AFDD model development.

Additionally, the ontology is verified and validated, ensuring the accuracy and comprehensiveness of the taxonomy's categories and hierarchical relationships. The ontology evaluation approaches include Task-based evaluation, Automated Consistency Checking, Data-driven, and criteria-based evaluation, which consists of one or more metrics, as shown in Figure 4.3.

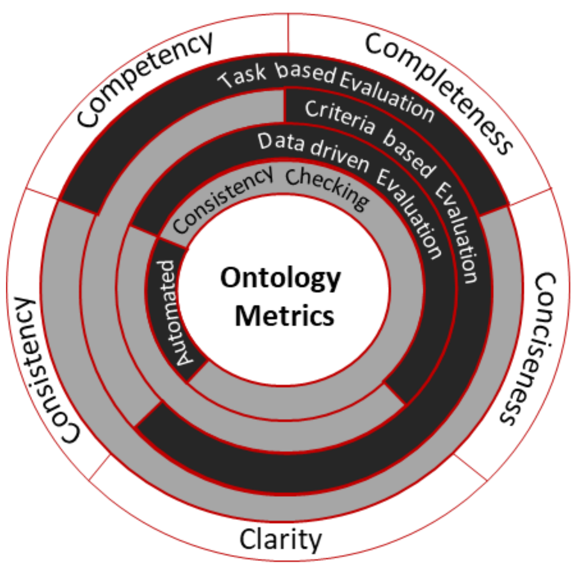


Figure 4.3: Metrics Used for Evaluation of AFDDOntology Adapted from [179]

4.4 AFDDOnto Validation

In this study, the Data-driven Evaluation is not performed as there are no known similar existing ontologies to be tested against. In addition, the AFDDOnto is tested for criteria-based evaluation using an online ontology evaluation tool named OOPS! (Ontology Pitfall Scanner!)

[182] which classifies pitfalls based on the Structural, Functional and Usability-Profiling dimensions resulting from an empirical analysis of over 693 ontologies. The tool was used to make sure no critical or important pitfalls are present and unresolved. The AFDDOnto is verified against ontology metrics, namely competency, completeness, and consistency metrics, which are used to demonstrate the suitability of an ontology for the competency questions defined.

In addition to validation using an online tool, a semi-structured survey is designed and conducted over three months to cover the suggested AFDDOnto principles required to assert and validate the SPARQL-constructed competency questions to satisfy the clarity metric of the ontology. The survey was conducted in one-on-one sessions with the experts. It consists of three sections with three closed-ended questions, four open-ended questions, two explanations, and one reference section. This necessitated the use of a semi-structured survey to accommodate any unforeseen changes that might arise based on the experts' feedback. The survey can be found in Appendix G.

The domain expertise of the respondents included HVAC FDD and Facility management familiar with BIM, BMS, and Machine Learning subjects that verification of the concepts developed for the competencies identified for the proposed AFDDOnto. The respondents were introduced to the main overarching concepts covered by the AFDDOnto and were then provided with the competencies being considered.

The experts were then asked to assess if the concepts were clear and covered the necessary information needed to reach the desired conclusion based on the available evidence. Additionally, three sets of concepts within the proposed AFDDOnto were analyzed and validated separately to understand and revise the relationship and structure between the concepts. Maintenance and Information concepts formed the first group and were investigated for the application of AFDD in HVAC. While the second group, Feature and State regarding HVAC Elements that are intended to enable tracking changes were analyzed. Finally, the third group, Track concepts, were investigated when considering with Feature AND/ State concepts of the proposed AFDDOnto.

A semi-structured survey was conducted to validate AFDDOnto, with the participation of seven experts. The limited number of interviewees was due to the requirement of interdisciplinary domain knowledge in BIM, Ontology Engineering, FDD of HVAC, and Applied AI. This reduced the pool of qualified interviewees. Although the small sample size is a limitation of the study, the survey analysis indicates a high level of consistency as the revisions were based on a systematic approach of triangulations which ensured experts' opinions were taken into consideration and necessary changes were made at each step. The survey results were further investigated through quantitative and qualitative analysis to refine the ontology for Clarity. In discussion 1, 57% of respondents indicated the overlap of Maintenance and Information concepts, suggesting potential shared elements between the two and, hence, the applicability of the Information concept at the Meta-data level. This means that the Maintenance concept is to be used exclusively by HVAC, and the Information plays a more general role in considering information pertaining to the building envelope. The two sources are expected to provide supplementary details for proactive maintenance planning. Hence, both concepts are well within coverage for AFDD of HVAC.

Furthermore, in discussion two, when considering the HVAC system (Element), 100% of respondents (7 out of 7) acknowledged an overlap between Feature and State concepts when considering the Element concept. Hence, the hasState object property was modified to capture the relationship between these two concepts. All respondents agreed on the overlap between Feature and State concepts but stressed that they must be captured separately to separate Features that can be fed for AFDD model development and State, representing a subset of state representative features.

In the third discussion, 86% of the participants agreed that the term 'Track' can be used in the context of keeping a record of historical HVAC system issues. This can be done by utilizing the State concept, where the AFDD model's features can make use of the 'hasState' property. In order to enhance wider adoption beyond AFDD, there was an emphasis on the need for further use cases, such as predictive maintenance. Respondents agreed that certain concepts within the AFDDOnto model overlapped, particularly between Feature and State, and that it was necessary to have separate concepts to capture them.

The ontology is revised based on the feedback received from the experts to include all the relevant axioms required by the ontology to answer the competencies directed at AFDD of HVAC, resulting in the fourth version of AFDDOnto proposed in this thesis. The survey indicate that the proposed knowledge model can answer competency questions and cover the necessary axioms and concepts. However, additional competencies require additional relationships to be defined between the existing concepts available in the AFDDOnto. The fourth version of the AFDDOnto, after testing against metrics and expert opinions, was published and is accessible through Github[4].

[4] https://github.com/arashhosseiniarash/AFDDOntology

4.5 BIM-BAS Integration and Automation

The limited availability of sensory and contextually related data is a challenge in developing AFDD models for commercial buildings. To address this issue, a research methodology has been proposed that involves generating dynamic BIM features and incorporating them into the AFDD model. These contextual features are added to compensate for the lack of sufficient sensory data. This approach facilitates leveraging BIM throughout the stages of AFDD model development and deployment.

The proposed solution integrates four main concepts: BIM, BMS, the AFDD model, and the "AFDDOnto" ontology, as shown in Figure 4.4, AFDDOnto is developed based on the outcomes of chaoter 2 and chapter 3. This approach has enabled the BIM to be extended with AFDDOnto to capture building and HVAC data along with their relationships, serving as a knowledge repository for the facility.

The bi-directional information flow between AFDD and BIM enables the incorporation of analytics and, subsequently, provision for updating them. By utilizing static and dynamic features from BIM and BMS, the AFDD model gains access to a new set of features that can enhance fault detection and diagnostics or compensate for not having access to important features. Additionally, storing AFDD analytics back into the knowledge model maintains an up-to-date representation of the facility, accessible for examination and querying by facility managers. The combined strength of BIM and machine learning techniques facilitates the effectiveness of AFDD for HVAC systems and compensates when relevant data is unavailable.

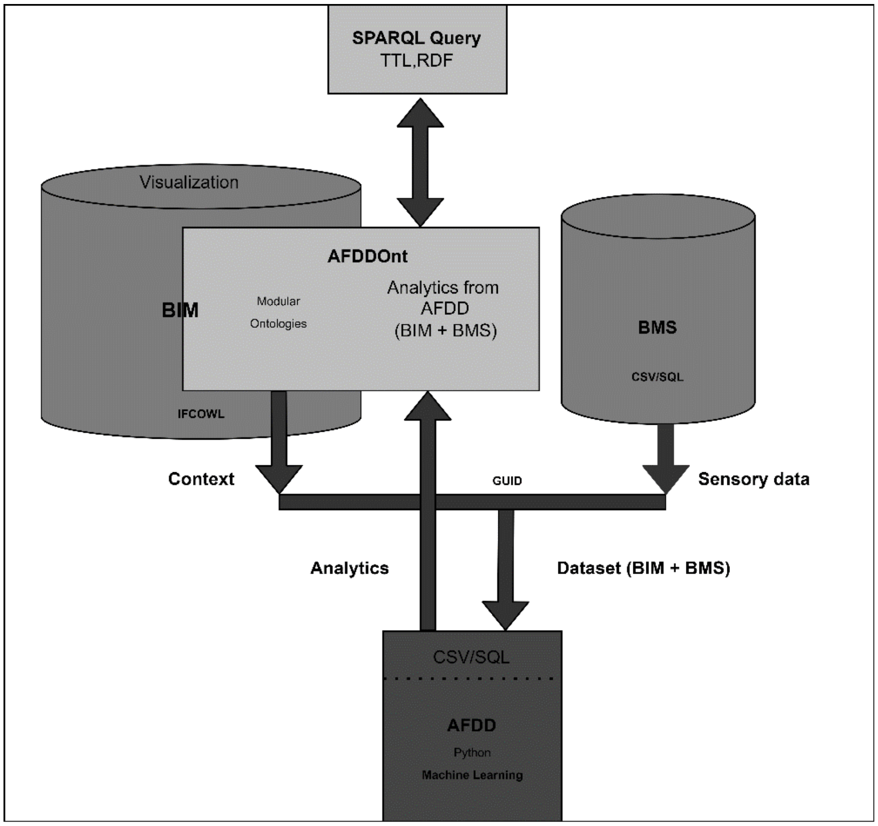


Figure 4.4: BIM and BMS Integration (Use case of AFDD of HVAC)

The proposed DT is created by integrating BIM with BMS through four steps, i.e., (i) Schema Conversion, (ii) Feature Engineering, (iii) Machine Learning for AFDD, and (iv) Knowledge Model Management. These steps are shown in Figure 4.5. As previously mentioned, AFDDOnto is developed to integrate BIM with BMS, drawing from knowledge extracted in the form of association rules, taxonomies, concepts, and relationships derived from the literature analysis of AFDD for HVAC systems and BIM. Furthermore, a case study is used to implement and demonstrate the proposed methodology in the subsequent section of this thesis to generate new dynamic BIM features and update the proposed model.

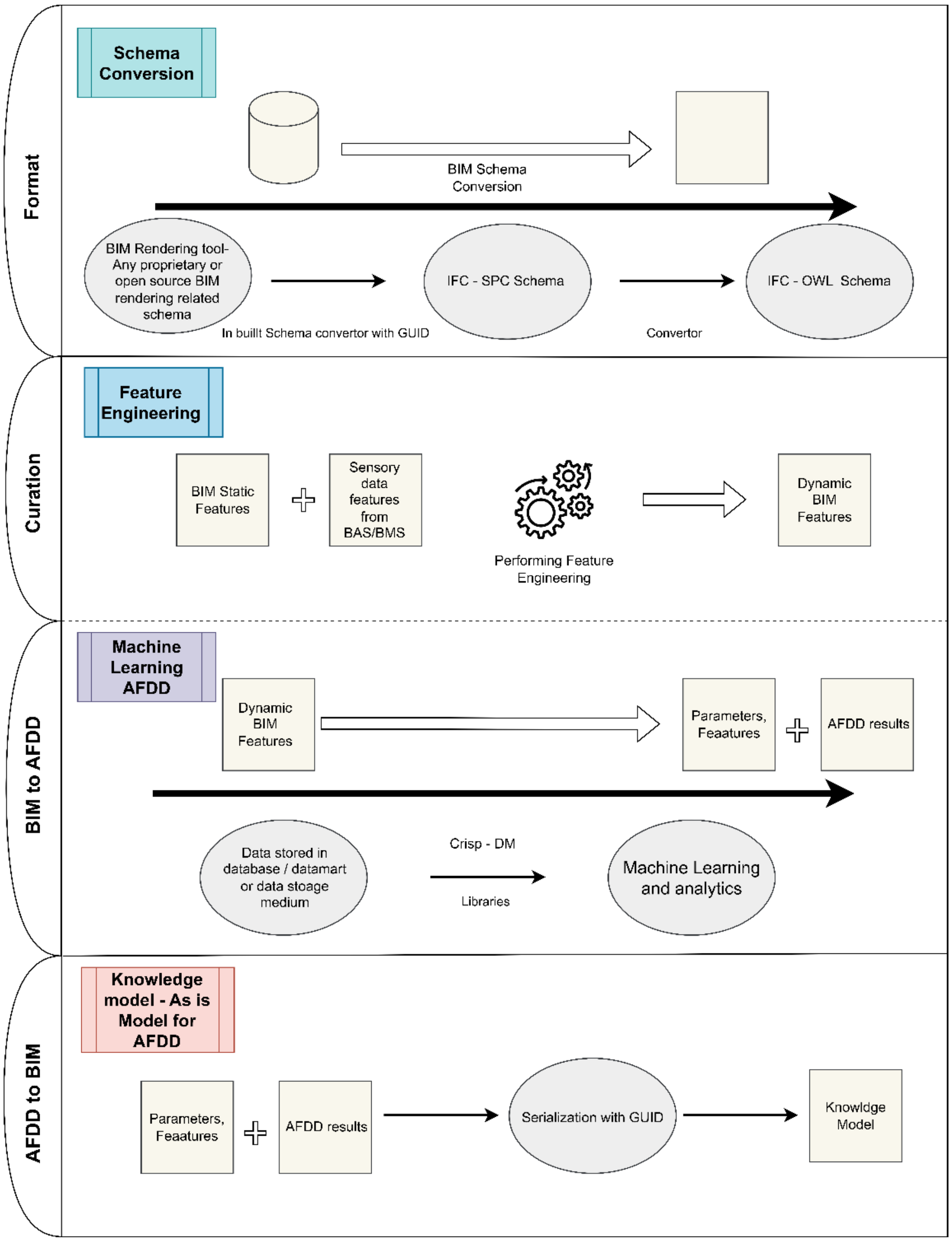


Figure 4.5: Proposed Methodology for BIM-based AFDD of HVAC

Schema Conversions: The first step in developing the proposed BIM-based knowledge model involves schema conversion, aiming to integrate isolated information from BIM and BAS/BMS using an accessible schema. This process facilitates centralized access to the integrated dataset. IFC is typically stored in EXPRESS Schema, but it can also be stored in IFCOWL to ensure compatibility with ontologies. However, as IFCOWL tends to have a larger file size and lacks all the necessary concepts for AFDD, the methodology proposed utilizes AFDDOnto, which maintains compatibility with IFCOWL and incorporates AFDD-related concepts. The modular ontology utilizes existing BOT ontology concepts, hence allowing for the reuse of BOT and IFC-SPC schema alignments for the proposed BIM-based ontology. The conversion and alignment procedure involves initially converting the IFC-SPC schema to IFCOWL and mapping it through BOT modules to AFDDOnto. The resulting ontology contains the necessary concepts for HVAC AFDD, maintenance, and tracking states, extracted from machine learning models developed using BIM and BMS features, as shown in Figure 4.6.

The suggested IFC MVD (Model View Definition) for exporting IFC data includes two options. Firstly, the IFC4 Reference View, which is the broadest MVD; and secondly, the IFC4 Design Transfer View, a subset of the IFC4 Reference View used for transferring BIM data to IFC format. The resulting IFC data is converted to the target schema, IFCOWL, using conversion tools such as IFCtoLBD [183]. The mapping between EXPRESS schema and IFCOWL utilizes the BOT ontology, transforming IFC entities such as Site, Building, Story, Space, and Elements into the destination knowledge model. Each BIM entity retains a GUID (Globally Unique Identifier) during this conversion for reference purposes.

On the other hand, BAS/BMS sensory data is imported for analytics, which may vary in format, requiring compatibility with specific analytics tools. In this study, the dataset was available in CSV format, and TensorFlow and Keras libraries were utilized for model development. The parameters, configurations, and results are imported into the BIM-based knowledge model, where relationships between concepts are captured in the form of axioms, allowing the assertion of individuals. The proposed knowledge model can be accessed using SPARQL for making queries.

Additional custom property sets can be incorporated into the IFC to supplement the knowledge model, containing product information, installation dates, change orders, maintenance work orders, or related data. These details are typically found in manuals, work orders, or product webpages and can be stored by BIM managers in COBie files. In BIM authoring tools such as Revit, ‘P-sets’, or ‘Property Sets’, are utilized to denote a collection of properties linked to an object in a BIM (model). These properties define specific attributes.

On the other hand, 'Entities' in the BIM model refer to individual elements with identifiable characteristics. Within ontology vocabulary, the 'P-set' and 'entities' are treated as individuals, while 'Defined type' and 'Enumeration type' are classes or concepts.

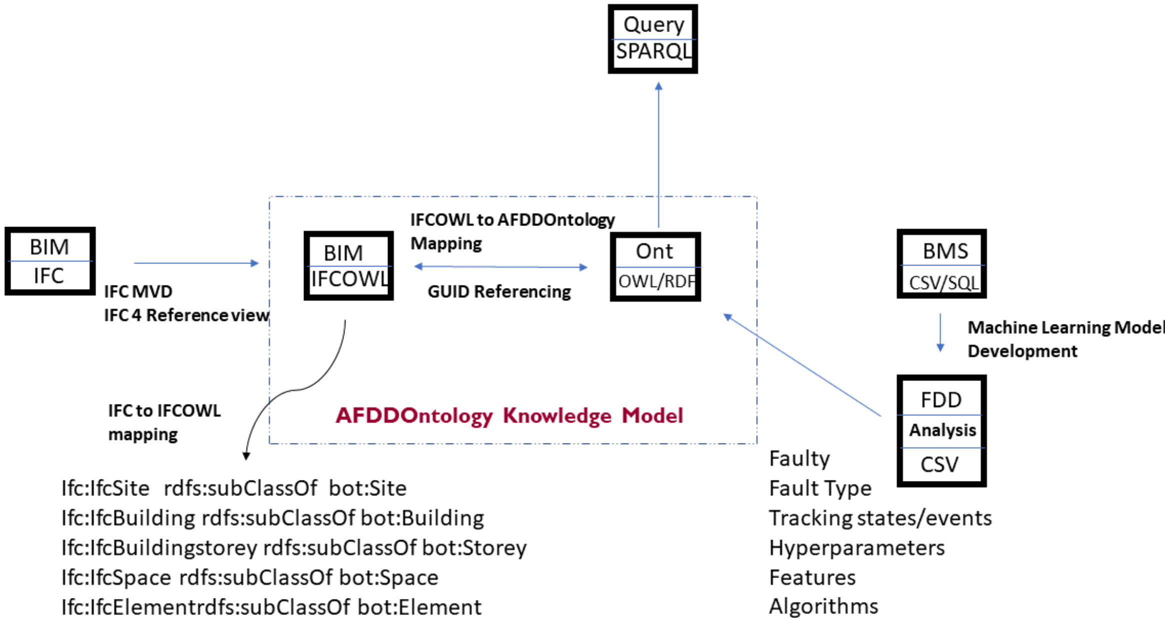


Figure 4.6: Schema Conversion and Mapping Used by AFDDOnto to Integrate BIM and AFDD Analytics

Feature Engineering: In practice, the availability of the features for AFDD model development depends on the number and type of sensors installed in the facility, including the HVAC and spaces. The purpose of the Feature engineering step is to use dynamic BIM and BMS data to enable the introduction of a new set of *Features of Interest (FOI)* that can be created as *calculated measures* (referring to a custom or derived measure that is calculated based on other existing measures or data elements within a dataset).

The goal of the feature engineering step is to improve the dataset by adding features that capture the dynamism of the facility that can be used for AFDD. The Feature Engineering step involves the usual data preparation activities during the pre-processing stages of model development. In addition, it requires identifying and selecting key variables or attributes that can indicate the system's state or condition, which can aid in diagnosing HVAC faults. This can be accomplished by utilizing one or more related features and performing arithmetic operations to introduce a new set of features that capture the building asset's dynamism.

An example of a calculated feature that uses arithmetic operations is the 'Temperature Differential' for an HVAC system. This calculated feature can be created by subtracting the outlet temperature from the inlet temperature. The generated calculated feature can be used as a metric to show the cooling efficiency of HVAC and can be further monitored in real-time to evaluate the performance of the system.

The Features of Interest (FOIs) identified in this study can capture the dynamism and context of HVAC and its associated spaces by using the state of operation of the HVAC system and spatial information such as the location of the HVAC sub-system and air conditioning function. The duration and frequency of data capture and storage are determined based on the type of HVAC fault under consideration. In cases where specific information is unavailable, the knowledge model incorporates concepts to define the frequency of knowledge capture.

The features utilized in this study are sourced from both BIM and BMS. The AFDDOnto axioms are employed to form connections for the relevant features needed for a given task, considering the connections between fault types, system types, and specific features. In the case study under investigation, static BIM features and BMS features are used to introduce a category of FOI that consists of multiple features for each of the building spaces to capture facilities' dynamism and compensate for limited sensory data.

The constraints associated with the type of sensory data that can be used for the AFDD model in addition to the needs of the algorithm being used, depend on factors such as the building type, HVAC type, the variations in occupancy and window states, as well as the number and type of sensors available (the features), and the placement of sensors in the building and HVAC. To assess the importance of the identified features, the study leverages statistical correlation analysis and machine learning algorithms to determine feature importance. These analytical techniques enable feature analysis by allowing identified features relative significance with those of existing features that the fault detection and diagnosis process can use.

The introduced features can be incorporated into the model by uncovering underlying relationships and patterns, contributing to AFDD model development, particularly in situations where relevant features may be unavailable. Sensory data can be categorized based on their function and be grouped and undergo statistical correlation analysis and feature importance analysis to determine which of the generated features can improve the AFDD process or compensate for the lack of availability of certain categories of features.

Dataset and Machine Learning Algorithm for AFDD: The key engineering steps for AFDD model development include firstly the data collection from BIM and BMS/BAS sources. The second step is feature engineering, which enables the generation of BIM-based features, i.e., performing feature transformation to transform static BIM features into dynamic ones. This leads to the creation of a dataset that in this study is referred to as the ‘curated dataset’. Different machine learning algorithms can be then developed over this dataset, for detection (and/or diagnosis) of various faults. Development of such algorithms typically follows the six steps of the

CRISP-DM (Cross-Industry Standard Process for Data Mining), i.e., problem understanding; data understanding; data preparation; modeling; evaluation; and deployment [184]. The study aims to integrate data from BIM and BMS to illustrate the importance of utilizing contextual information and dynamic BIM features. In this step of the proposed methodology, the curated dataset containing features from both the BMS and BIM data sources is utilized for the development of the AFDD model.

The initial BIM features are static and need to be transformed into dynamic BIM features using calculated measures in the feature engineering step. The dynamic features created capture the temporal characteristics and behavior of the HVAC system that can be used to replace or to be added to the dataset in events where certain sensory data are unavailable or that can provide for additional fault-type diagnostics of HVAC. These temporal characteristics can include time intervals, temporal dependencies, or other external factors such as seasonality. These additionally introduced features represent the state of the system and their applicable spaces that are not typically available in BMS/BAS and are added to the curated dataset in the form of new features that can be used for AFDD of HVAC.

The proposed methodology enables the use of various algorithmic techniques tailored to meet specific requirements and characteristics of HVAC levels and faults without being restricted to any particular data-driven algorithm. The effectiveness of an automated fault detection and diagnostics model is assessed in this study by measuring its sensitivity to the introduced set of features.

This sensitivity measure is based on the accuracy of diagnostic results for the faults being investigated. This measure is then used to compare and select the most suitable AFDD model. By comparing algorithms, the methodology enables the asset manager with decision-making capability for informed algorithm selection and comparison. While this study does not directly contribute to developing new machine-learning algorithms, it plays a crucial role in AFDD for HVAC systems by allowing the introduction of new features to help with fault detection, fault diagnostics, and limited sensory data availability.

BIM-based Knowledge Model: The proposed BIM-based model includes ten provisions in terms of concepts or classes to store analytics obtained from AFDD, as shown in Figure 4.7. The Track and State concepts enable the knowledge model to be updated at predetermined intervals to represent the present state of the facility. By enriching BIM with analytics instead of the entire AFDD data, the knowledge model retains only relevant information needed for AFDD use cases, avoiding redundancy. This is achieved by mapping BIM elements and associated features using

Uniform Resource Identifiers (URIs) that maintain compatibility with IFC, including source GUID and custom parameters that are added to the BIM model for AFDD purposes.

The link between BIM and BMS sources is established and maintained by mapping relevant concepts essential for AFDD, ensuring seamless integration. The BIM-based knowledge model requires periodic updates with new analytics based on the HVAC system's maintenance schedule, following predefined intervals recommended by ASHRAE, such as monthly, three-monthly, semi-annually, or annually [189], and additionally whenever AFDD models detect faults. The knowledge model includes this concept under the maintenance class, providing information about frequency, type of maintenance action, and diagnostics pertaining to HVAC.

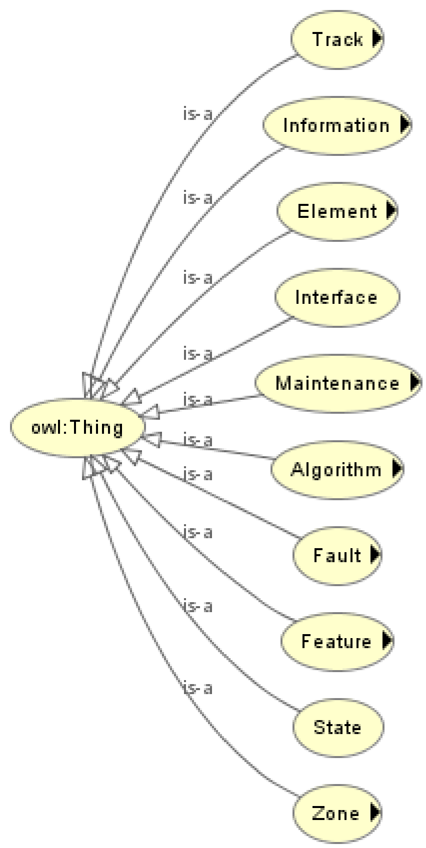


Figure 4.7: Concepts Used in BIM-Based Knowledge Model (AFDDOnto)

Access to the BIM-based knowledge model is facilitated through the SPARQL query language. It allows users to construct queries and retrieve information about features, configurations, analytics, spatial details, maintenance information, images, and links to external files. The semantic model in OWL format can be stored in RDF and TTL formats, making it accessible online. The application of the BIM-based knowledge model for the proposed methodology is demonstrated in the case study section, and further details are provided in the discussion section of this thesis.

Chapter 5: Implementation of the Proposed Integration Solution [5]

5.1 HVAC AFDD Case Study

The ORNL's Flexible Research Platform (FRP) [186] is selected as a case study to demonstrate the proposed solution's analysis, testing, and evaluation. The FRP is a 3,200 ft^2 facility designed to emulate a 1980s-era office building equipped with a single packaged RTU (Roof-Top Unit) connected to a multi-zone VAV (Variable Air Volume) sub-system [187] , as depicted Figure 5.1. The test facility is designed as a controlled and fully equipped lab experiment, where faults are intentionally introduced individually and targeted. This controlled environment allows for precise investigations and analysis, as all sensory data necessary for fault detection and diagnosis are readily available.

The case example used in this study is modeled using two of the most common algorithms used in AFDD (according to the literature analysis carried out in chapter 2 [52]), namely ANN and SVM. The ANN models work based on processing the input data by means of one or more interconnected neuron layers which perform transformation to data [192] [193].

The SVM classifiers work by determining the optimal hyperplane to maximize the margin between a set of classes of data in the feature spaces to enable classification [194] [195]. Common parameters that are fine-tuned when training an ANN include the Activation function (ReLU and softmax), Hidden drop-out ratio, stopping tolerance, random seed, hidden layer size, epoch, and the optimizer (Adam). For SVM, the most common parameters include Loss Function, C value, and Kernel type. In this study, both model families were trained and tested using Python libraries such as TensorFlow, Keras, and scikit-learn.

The case study model, which utilized both BIM and BMS data as input for model development, is used to test the knowledge model using different case-representative algorithms for analytics. Initially, AFDD models are developed using SVM algorithms, with one focusing on fault detection (FD) achieving 96% accuracy and the other targeting diagnostics (FDD) achieving 97% accuracy. Additionally, two models are constructed using ANN, where the Fault Detection model achieved 98% accuracy and the fault detection and diagnostics model achieved 99% accuracy. The high accuracy of the models is due to the controlled environment nature of the test facility, in which the faults are introduced individually for a period of one day.

---

[5] The material of this chapter is published in form of the following publications:

1. Hosseini Gourabpasi A, Nik-Bakht M. BIM-based automated fault detection and diagnostics of HVAC systems in commercial buildings; Journal of Building Engineering, Volume 87, 15 June 2024, 109022. https://doi.org/10.1016/j.jobe.2024.109022

The model incorporated a curated dataset that includes BIM and BMS features. In the case study of this thesis, some groups of features were intentionally removed to simulate limited sensory data availability in commercial buildings. This enabled the study to assess the effectiveness of dynamic BIM features.

The accuracy of the model for fault detection and diagnostics was used as a sensitivity test to determine the most suitable AFDD algorithm that is compatible with the dynamic BIM features in the curated dataset. This approach enables the development of an AFDD model in various ways, such as improving model accuracy, accommodating additional fault types, and addressing the unavailability of sensory data.

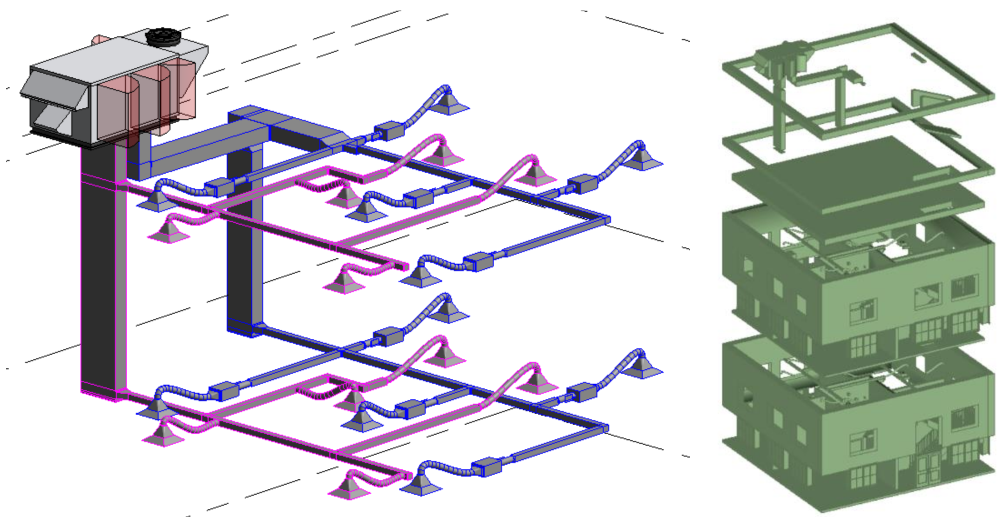


Figure 5.1: The FRP BIM Model

The BIM model for the case study is created by author of this thesis using related documents pertaining to building, HVAC assets, experiment setup and documents that could provide additional context as a source of BIM [196] [197] [198] [199] [200] [201]. The BIM model is authored using Autodesk's Revit to represent architectural and mechanical models of the test facility, including the RTU system.

The BMS data is stored in CSV format, representing the sensory data originating from the BMS system of the HVAC. The dataset is used to create machine learning models, namely SVM models and ANN models for fault detection and diagnostics of seven fault types with a data collection interval of one minute, which can be accessed through Kaggle[1]. The AFDDOnto integrates data from both BIM and AFDD models, including the custom properties that can be created in BIM for specific use cases such as FDD. The methodology for utilizing BIM and BMS

[1] https://www.kaggle.com/datasets/claytonmiller/lbnl-automated-fault-detection-for-buildings-data

data to populate the AFDDOnto involves four main modules, namely format exchange, dataset curation, BIM to AFDD, and AFDD to BIM, which are based on the author's proposed methodology for using BIM and BMS data using AFDDOnto [198].

In the first module, the BIM model is required to be converted to IFCOWL to be compatible with AFDDOnto. The second module is utilized to add BIM-based features to the BMS dataset using feature engineering. The third module involves the development of AFDD models for the FDD of HVAC, where the results and configurations are stored externally. The last module enables the integration of the extracted knowledge into AFDDOnto by retaining links through GUID.

The content of the AFDDOnto can be accessed by the facility manager in three ways namely through query, knowledge graph models, and as well as retaining links to external models. To query the knowledge model SPARQL queries are constructed based on the 13 competencies identified by literature analysis to retrieve the required data. The competencies listed in Appendix F are used to populate the AFDDOnto with the required information.

Additionally, the knowledge model can be visualized through a knowledge graph and also retains the links to external files such as the BIM model. The result of this integration results in the inclusion of data from the BIM and as well as the data AFDD model developed using the BMS dataset. The AFDDOnto serves as a framework for organizing and managing the data and insights derived from the case study, facilitating future analysis, comparisons, and further improvements in the HVAC fault detection and diagnosis model using the relevant concepts capturing the relationship between HVAC as an asset and the building.

The lab facility provides an ideal setting for investigating the proposed methodology by allowing the research to explore how the AFDD may differ in real-world scenarios. In practice, there could be limitations or challenges in obtaining comprehensive sensory data or additional uncertainties due to weather, environment, occupants, and HVAC assets. Therefore, it becomes crucial to explore the potential of utilizing BIM to compensate for any potential shortcomings encountered in practical settings where access to desired sensory data for an effective AFDD model is not feasible.

The objective of incorporating BIM into the study is to evaluate whether BIM can serve as a valuable resource to bridge the gap between the controlled lab experiment and real-world applications. The study aims to investigate the potential of BIM to overcome practical limitations and improve the overall performance and reliability of HVAC fault detection and diagnosis techniques in real-world scenarios.

The proposed methodology generates dynamic BIM features incorporating BIM and BMS features to develop fault detection and diagnostics models for an RTU system connected to ten VAV sub-systems. The floor plan for the two-story building is shown in Figure 5.2, which consists of 10 spaces used as rooms and two additional spaces on each floor used for staircases.

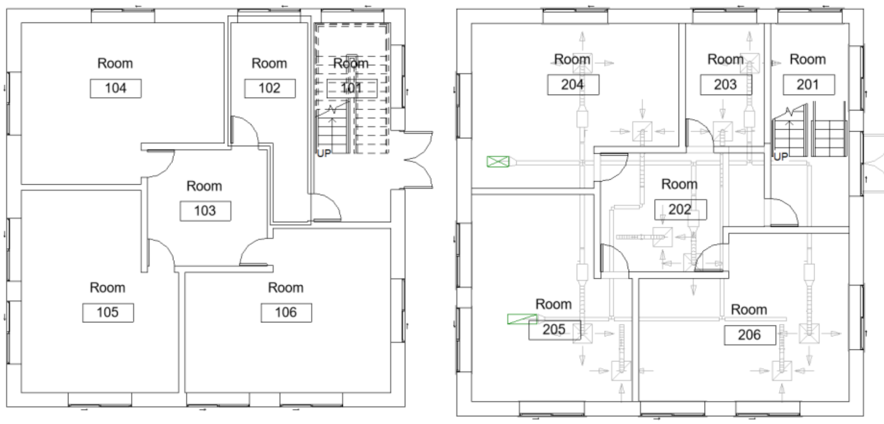


Figure 5.2: Floor Plan of the FRP Facility

5.2 AFDD of HVAC Using BIM and BMS Integration

The conversion process is shown in Figure 5.3. The BIM (model) used in this study is at LOD 300 as per the BIMFORUM standard [199]. At this level of development, the 3D model contains detailed geometric representations of components such as ductwork, equipment, and accessories. While the model provides accurate representations of equipment dimensions, specific details such as insulation, hangers, and supports are not required [199].

The built-in converter of Revit is used to export the model to the IFC-SPC schema. IFC4 Reference view is selected, which is the most comprehensive IFC output. To maintain compatibility between the BIM-based knowledge model and the model created using the BIM authoring tool, the output file includes the GUID of the entities, including the custom entities created in BIM. The IFC model is separately imported to an IFC viewer called "usBIM.viewer+" [200] for analysis.

In this study, the process of converting IFC to IFC OWL is automated using the recommended "IFCtoRDF" converter [71] by buildingSMART. This converter is capable of extracting and mapping the necessary information from the source to the target ontology, making it easier to identify and map the entities required for the AFDD of HVAC. As the IFC file output is now made available in the SPC schema, it needs to be converted to a schema applicable to ontology. Hence, the model is converted to OWL schema using the IFC to LBD converter [183], which maps the entities in both schemas and populates the schema with BIM data. The IFCOWL is imported to the BIM-based ontology. The "Protégé" ontology editor and framework for building intelligent systems [180] is used for knowledge model management of the BIM and BMS integrated data.

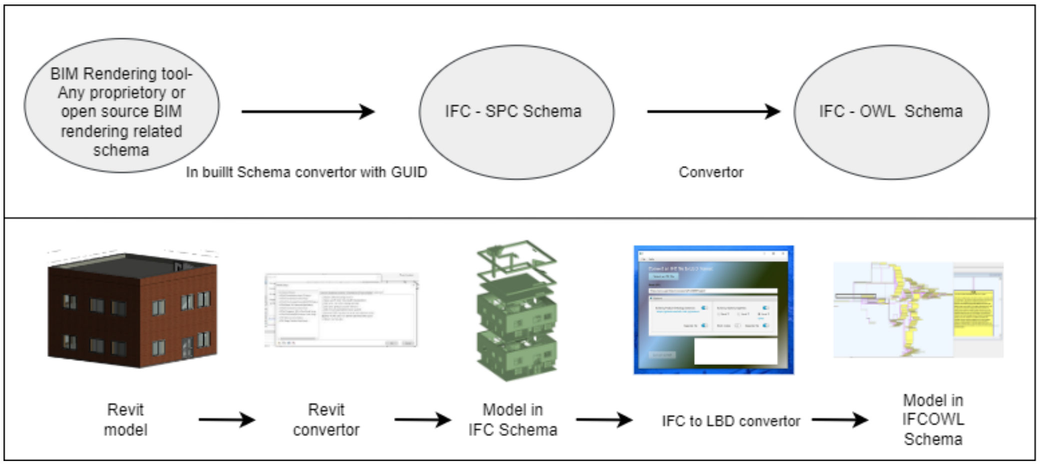


Figure 5.3: Model Conversion for the FRP Case Study

The FRP dataset used in this study consists of 68 features extracted from BMS, representing sensory data collected by various sensors in both time-series and state-representative formats. To analyze the dataset effectively, the BMS features were grouped into 11 categories based on their similarity, regardless of the sensor's location. These categories include *VAV temperature, room temperature, humidity, status, supply and return temperature, metered data, setpoint, flow rate, circuits 1 and 2*, related features.

To address potential scenarios where certain BMS features required for effective AFDD may be unavailable, 80 static features were sourced from BIM. The BIM features were further divided into *Live* and *Static* BIM features. The static features are directly imported, while the dynamic BIM features are the FOI, which act as calculated features comprising both BMS time-series data and BIM data. These dynamic BIM features provide crucial contextual information necessary for AFDD of HVAC systems.

The static BIM features encompass spatial data, including features such as *distance to HVAC from VAV* (the distance at which the VAV is located from the RTU), *room area, window area, door area, opening area* (indicates the area of the sum of the opening, including the door and windows for each space), *next to the shaft* (indicates if the staircase shaft surrounds the walls), *and exterior room* (signifies if the room is situated inside the floor or is facing outdoors) features. Additionally, custom properties are defined to offer supplementary contextual details about HVAC components, faults, and historical information, further enhancing the overall AFDD process.

The FRP dataset is combined with BIM data to create a new curated dataset as a result of model conversion explained above to have the dataset ready for AFDD model development. The GUIDs assigned to the static features of the FRP architectural and mechanical model serve as identifiers to associate the BIM elements with BMS time-series data. This curated dataset contains both BIM and BMS features and acts as a pool of candidate features. To identify the most important features, a feature importance analysis is conducted.

A Random Forest classifier is used to train the dataset, and feature importance is measured as the mean and standard deviation of impurity decrease within each tree. Other feature importance techniques, such as XGBoost Classification Feature Importance (based on gain), Random Classification, and Decision Tree Classification (normalized total reduction of criterion brought by the feature), were also assessed. After evaluating various classifiers, the Random Forest Classifier was selected because it effectively highlights the importance of BIM features in the dataset.

This approach identifies key features that play a significant role in the AFDD process, particularly when certain BMS features are limited or unavailable. In this study, the top feature categories were only considered to limit the number of features. In practice, the number of features can be modified based on the algorithm in use or is limited to the actual availability of BMS features and BIM data for the use case of AFDD. In this study, the following feature categories are investigated, i.e., *Air humidity, Discharge Temp, Suction Temp, Air temp* and *cooling setpoint,* and "*spaceAirConditioned*"; which were selected by performing feature analysis to identify the most impactful features.

The categories that have been selected are determined through feature analysis of the curated dataset. Most of these categories come from the BMS, except for the *spaceAirConditioned* category. The main reason is that *spaceAirConditioned* is the only Live BIM feature among other static features, indicating the importance of such features. Nevertheless, if the dataset is accessible for an extended period and maintenance information is available for the fault types under consideration, the maintenance information will be incorporated to form BIM features. It is anticipated that the use of the proposed methodology in existing buildings will signify the role of BIM features.

The *spaceAirConditioned* category, which represents the 12 spaces used as rooms in the FRP case study, is translated to 10 calculated features that depend on the states of two circuits and two compressors and air conditioning of ten zones. The remaining two spaces are staircases not directly connected to VAV (Variable air volume) and act as a shaft in the building connecting the two stories.

Further, the calculated measure takes into account the HVAC schedule pertaining to the time of operation and modes, The state of the HVAC, i.e., cooling, heating, or fan state, and the spaces having access to VAV considered as being air-conditioned or not, to define the dynamic feature. Each HVAC mode is activated once the thresholds are met. The thresholds of the HVAC schedule are defined based on the occupancy in the facility, which is predefined. The details are provided below.

In the interest of brevity, this study only looks at the system-level function of the defined FOI (*spaceAirConditioned*). The FOI category uses time-series data from BMS to denote the HVAC in the air conditioning state, denoted in binary form (True/False) in the dataset. Figure 5.4 represents the HVAC schedule for both Occupied (Left) and un-Occupied Mode (Right).

The thresholds for both occupancy modes are different as shown. The RTU has a supply air temperature of 12.7 °C. In the occupied mode where occupancy is simulated, once the 24°C threshold is met, the HVAC begins to air condition (Cool) the building spaces; below the defined threshold, only the fan is activated until the space temperature drops below 21°C at which heating is turned on.

To further explain, in this case based on AFDD's taxonomy, the RTU, is the 'HVAC system' and is associated with the cooling 'function'. The fan is the 'HVAC equipment', and the VAV represents the 'sub-system' level of the HVAC and has the heating 'function'. The "dynamic BIM" feature, is presented in the form of a calculated measure which is shown in Figure 5.5, thus enables us to combine the streams of real-time data to be enriched with contextual data from BIM and create a synthetic feature that can be added to the dataset to be used for AFDD. This dynamic BIM feature is intended to compensate for missing sensory data (Context-aware) in this case study spatial information that can be added for AFDD model development.

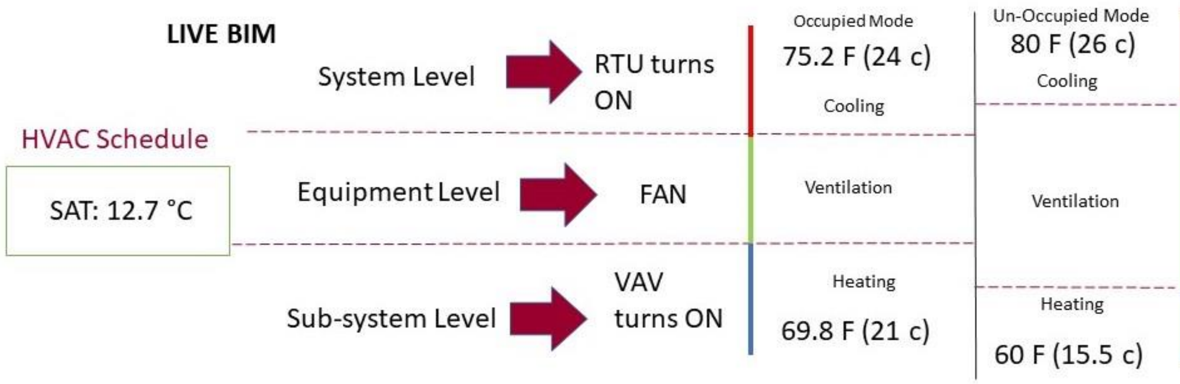


Figure 5.4: BIM Based Calculated Feature Generation for FRP Case Study

For the calculated measure (*spaceAirConditioned*) to be True i.e., to indicate the HVAC system is in an air conditioning state; the following conditions must be met. (As illustrated in Figure 5.5). At least one True value must be present in each row i.e., in the case of following features 1- *Compressor 1 and Circuit 1*, 2- *Compressor 2 and Circuit 2*, or 3- *both Compressors 1 and 2 with Circuits 1 and 2 at least one must hold True*. Please note since Rooms 11 and 12 don't have a separate VAV, so the air conditioning mode in those rooms remains as False.

BIM Feature Air conditioned (True/False) @ Building and HVAC system Level

Static BIM	Room 1	Room 2	Room 3	Room 4	Room 5	Room 6	Room 7	Room 8	Room 9	Room 10	Room 11	Room 12
	True	True	True	True	True	True	True	True	True	True	False	False
When Compressor 1 and circuit 1 one features have values	True	True	True	True	True	True	True	True	True	True	False	False
When Compressor 2 and circuit 2 two features have values	True	True	True	True	True	True	True	True	True	True	False	False
When circuits 1&2 and Comp 1&2 are not working	False	False	False	False	False	False	False	False	False	False	False	False

LIVE BIM

Figure 5.5: Calculated Measure Defined in FRP Case Study for Creating Dynamic BIM Feature

The curated dataset, which combines BMS and BIM features, is imported into the Anaconda distribution program [201], acting as a virtual environment specifically set up to develop AFDD models. ANN modes are initially developed using the newly created dataset, and further, this study uses SVM to test and compare the models. Analytics such as FDD accuracies and parameters used for model development are fed to the knowledge model, and sensitivities of the models for the given features are measured.

The developed models are divided into fault detection models using ANN and fault diagnostics models using ANN. Similarly, SVM models are developed for comparative analysis. The highest accuracy achieved using Berkeley Lab FDD dataset [202] for the AFDD models developed are 0.98 and 0.96 for fault detection utilizing ANN and SVM models, respectively. For fault diagnostics of HVAC faults, the accuracies are 0.99 and 0.97 for ANN and SVM models. The reason for such high accuracies can be attributed to the usage of new HVAC equipment and controlled nature of the facility, individual introduction of faults and a set of diverse sensors installed in the facility with the main objective of AFDD applications.

To understand the importance of the type of sensory data and its availability on the model. The AFDD models were developed by focusing on the top 5 important features, namely *Terminal: Room Air Humidity, RTU: Circuit 2 Discharge Temperature, RTU: Circuit 2 Suction Temperature, Terminal: Room 102 Air Temperature, Terminal: Room Air Temperature Cooling Setpoint*. Several models were created in which one or more combinations of BMS features denoted above were dropped and replaced with the BIM features to simulate scenarios where facilities have limited sensory data available to examine the significance of dynamic BIM features in the absence of common and impactful features.

Chapter 6: Results and Discussion

The results of the study are presented by considering the flow of data, i.e., from BIM to AFDD and, conversely, AFDD to BIM, as shown in Figure 6.1, to illustrate the bi-directionality that can enable realization of the Digital Twin of the facility in terms of the flow of data. The proposed methodology applies to AFDD of HVAC at the system, equipment, and sub-system levels.

The generated dynamic BIM FOI category is made possible by integrating the BIM model of the facility that enables the utilization of the spatial relationships of room spaces and HVAC components with the real-time stream of data being captured by BAS/BMS. The dynamic BIM, utilizes dynamic features generated using sensory and contextual information from BIM and BMS.

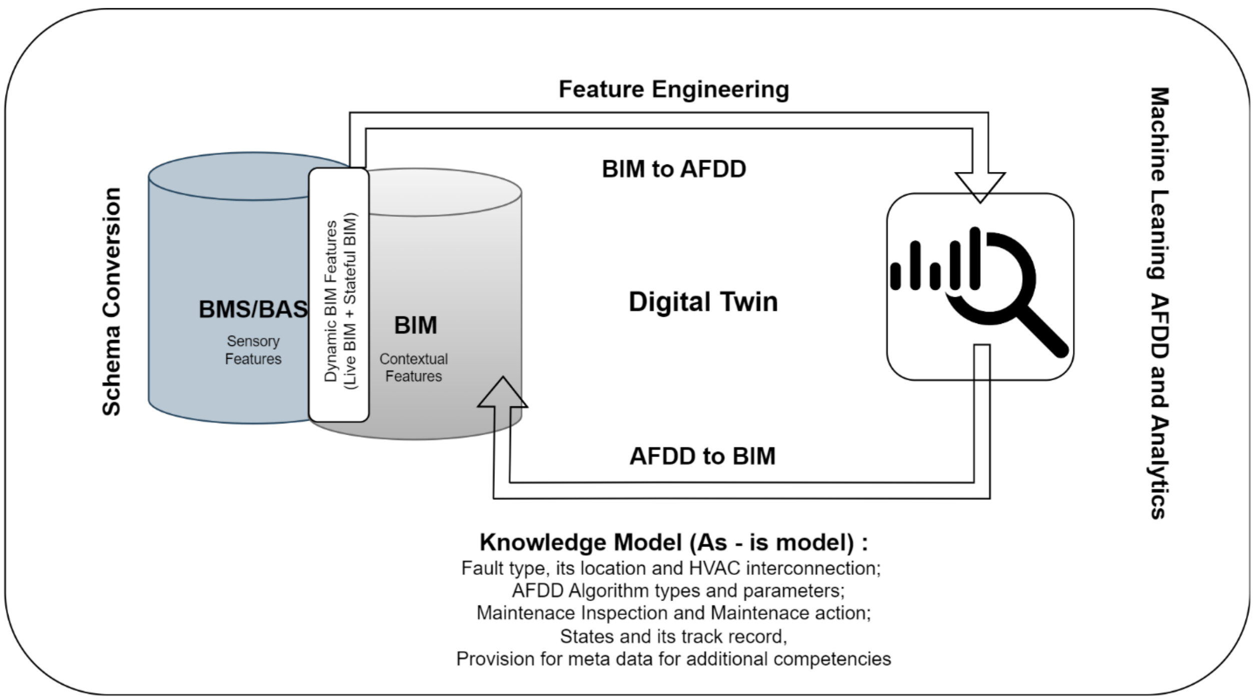


Figure 6.1: Bi-directional Flow of Data Between BIM and BAS/BMS

Initially, in the BIM to AFDD flow, the BIM and BMS data are used to introduce additional features that take into account spatial/contextual information about the HVAC and the thermal zones it serves; by doing so, the generated features are exported to the AFDD model to perform Fault detection and diagnostics using various algorithms. The introduced features are intended to either compensate for the lack of specific sensory data availability or to improve the model.

The results of the AFDD models from the case study Indicate that when features belonging to a specific category are dropped from the curated dataset and replaced with dynamic BIM features; it can lead to an improvement or similar accuracy in fault detection. For instance, when the Room air humidity category, consisting of 12 features for each zone, is unavailable, adding dynamic BIM features can enhance the detection accuracy.

Moreover, in cases where the most significant features are not available, integrating BIM features has shown a slight increase in performance for ANN models but not for SVM models. The findings suggest that BIM features have a more significant impact on ANN models compared to SVM models.

The impact of replacing the top 5 most important features with BIM sensory information from the curated dataset demonstrates the significance of BIM features. Although BIM features can partially compensate for the absence of the most influential features, the models generally did not surpass their performance. The behavior of BIM features in the AFDD model depends on the classification model used. BIM features improved accuracy only when specific BMS features, such as *Terminal: Room Air Temperature Cooling Setpoint* and *Terminal: Room Air Humidity*, were omitted.

It is essential to understand the interplay between BIM and BMS features and their influence on the overall performance of the AFDD model. Notably, for ANN models, replacing dynamic BMS data with BIM features improved the fault detection and diagnosis accuracy by up to 6%, considering the facilities' fully controlled environment that significantly reduces the need for BIM features.

In the flow direction from AFDD to BIM, the parameters and configurations alongside analytics obtained using the AFDD model are imported into the BIM-based knowledge model that includes provision for AFDD knowledge extracted from the model. This knowledge capture serves as a repository for storing AFDD analytics and updating the model to reflect the facility's current state. This process allows for the development of AFDD models using the same configuration for similar facilities and facilitates comparisons between different AFDD models.

The AFDD model's parameters, features, and results, including accuracy, are transferred and stored in CSV format for each fault detection and diagnostics model. Acknowledging that the case study takes place in a fully controlled environment reduces the significance of dynamic BIM features regarding contextual information.

However, this controlled setup allows testing of the AFDD model under limited sensory availability. In real-world scenarios, sensor failures or biases may lead to incomplete data capture, and in such situations, utilizing BIM features becomes valuable, as they can partially compensate for the lack of dynamic data. Despite the controlled environment downplaying the importance of

BIM features in terms of accuracy, their inclusion remains relevant in practical applications, contributing to a more comprehensive and accurate fault detection and diagnosis process.

Although this study does not contribute directly to the development of AFDD algorithms, it was necessary to reproduce existing AFDD models to demonstrate the introduction and usage of dynamic BIM features, particularly in the absence of sensory data commonly used for AFDD. Moving forward, the main contribution of this study will be presented, which is the access to the developed knowledge model and its application for BIM-based AFDD model development. Additionally, the visualization of BIM-based AFDD results in HVAC systems using knowledge models will be discussed.

6.1 Accessing Knowledge in BIM-Based AFDD of HVAC Systems

In the FRP case study, Blazegraph [203] database and Protégé knowledge management editor are used to support SPARQL queries. Blazegraph serves as a scalable and high-performance triplestore for storing RDF data, facilitating efficient and flexible querying of the knowledge model. On the other hand, Protege is employed as an ontology editor and knowledge representation tool, enabling the development and management of the BIM-based ontology.

The knowledge model allows for adequate access and retrieval of information pertaining to features, configurations, analytics, spatial data, maintenance details, images, and links to external files related to the facility. Further, Protégé is used to import the AFDD analytics. This involves importing the analytics and configurations of the AFDD models to the BIM-based ontology, i.e., the parameter's features, which are specific settings, configurations, fault types, and model accuracy. While this step was done semi-manually for the case study, this step can be automated using scripts for full automation of the process.

The AFDD analytics are imported into AFDDOnto using BIM and BMS data. *Figure 6.2*, depicts a snippet of the ABOX version of the AFDDOnto for the 4 AFDD models developed, which is represented in the form of a knowledge graph. This represents how the AFDDOnto concepts of Zone, Information, Element, Algorithm, and Feature populate AFDDOnto. The ABOX model of the AFDDonto with respective instances is stored in the Blazegraph [203] database to facilitate queries, which can be accessed using SPARQL queries.

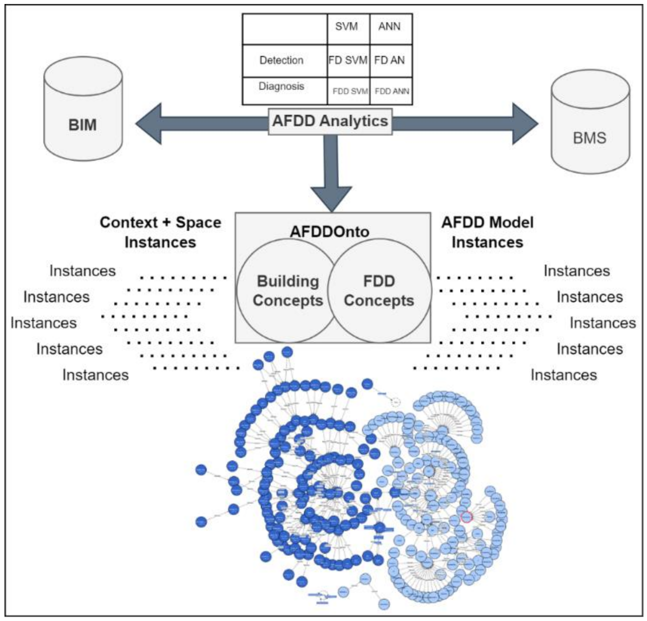


Figure 6.2: Conceptual Excerpt of the ABOX Version of the AFDDOnto Populated with FRP Case Study Instances

Knowledge graphs can be effectively visualized by loading the ontology into Protégé and utilizing the "Graph View" tab; users can visually explore the interconnected nodes and edges representing classes, individuals, and properties within the knowledge graph. This interactive and intuitive graphical representation aids in understanding the hierarchical structure and relationships within the graph, facilitating analysis.

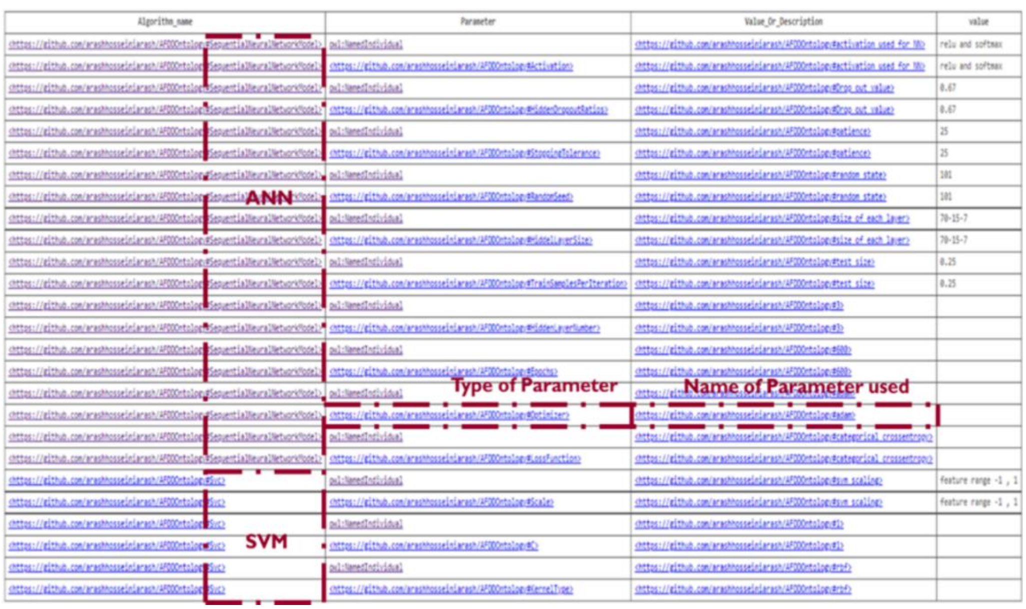
On the other hand, accessing knowledge graphs is possible using Blazegraph's SPARQL endpoint. HVAC operators can execute SPARQL queries to retrieve specific information from the knowledge graph. This enables complex queries (HVAC equipment, spaces and occupant related), relationship analysis (Spatial, associated faults), and valuable insights from the data (Status and historical), making such semantic databases a suitable choice for managing and querying large-scale knowledge graphs in various applications, including the semantic web, data integration, and knowledge representation.

Figure 6.3. depicts a sample query for the AFDDOnto, which contains the IRI namespaces and the query construct to access the AFDDOnto knowledge using the SPARQL query. Based on the competency identified, the user may be interested to know about the parameters used for each algorithm during model development. The SPARQL query retrieves parameters and associated values for SVM and ANN models used in the case study.

Competency Question in Natural Language: What parameters are used for Algorithm "ANN" and "SVM"?

Competency Question Construct using SPARQL:

```
prefix : <https://github.com/arashhosseiniarash/AFDDOnto#>
prefix owl: <http://www.w3.org/2002/07/owl#>
prefix rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
prefix xml: <http://www.w3.org/XML/1998/namespace>
prefix xsd: <http://www.w3.org/2001/XMLSchema#>
prefix rdfs: <http://www.w3.org/2000/01/rdf-schema#>
Prefix bot: <https://w3id.org/bot#>
Prefix brick:<https://brickschema.org/schema/Brick#>
SELECT  ?Algorithm_name ?Parameter ?Value_Or_Description ?value
WHERE { ?Algorithm_name :hasParameter ?Value_Or_Description .
        ?Value_Or_Description a ?Parameter .
OPTIONAL
{?Algorithm_name :hasParameter ?Value_Or_Description .
        ?Value_Or_Description a ?Parameter .
 ?Value_Or_Description rdfs:comment ?value .}
      }
ORDER BY ASC (?Algorithm_name)
```



Algorithm_name	Parameter	Value_Or_Description	value
<https://github.com/arashhosseiniarash/AFDDOntology#SequentialNeuralNetworkModel>	owl:NamedIndividual	<https://github.com/arashhosseiniarash/AFDDOntology#activation used for NN>	relu and softmax
<https://github.com/arashhosseiniarash/AFDDOntology#SequentialNeuralNetworkModel>	<https://github.com/arashhosseiniarash/AFDDOntology#Activation>	<https://github.com/arashhosseiniarash/AFDDOntology#activation used for NN>	relu and softmax
<https://github.com/arashhosseiniarash/AFDDOntology#SequentialNeuralNetworkModel>	owl:NamedIndividual	<https://github.com/arashhosseiniarash/AFDDOntology#Drop out value>	0.67
<https://github.com/arashhosseiniarash/AFDDOntology#SequentialNeuralNetworkModel>	<https://github.com/arashhosseiniarash/AFDDOntology#HiddenDropoutRatios>	<https://github.com/arashhosseiniarash/AFDDOntology#Drop out value>	0.67
<https://github.com/arashhosseiniarash/AFDDOntology#SequentialNeuralNetworkModel>	owl:NamedIndividual	<https://github.com/arashhosseiniarash/AFDDOntology#patience>	25
<https://github.com/arashhosseiniarash/AFDDOntology#SequentialNeuralNetworkModel>	<https://github.com/arashhosseiniarash/AFDDOntology#StoppingTolerance>	<https://github.com/arashhosseiniarash/AFDDOntology#patience>	25
<https://github.com/arashhosseiniarash/AFDDOntology#SequentialNeuralNetworkModel>	owl:NamedIndividual	<https://github.com/arashhosseiniarash/AFDDOntology#random state>	101
<https://github.com/arashhosseiniarash/AFDDOntology#SequentialNeuralNetworkModel>	<https://github.com/arashhosseiniarash/AFDDOntology#RandomSeed>	<https://github.com/arashhosseiniarash/AFDDOntology#random state>	101
<https://github.com/arashhosseiniarash/AFDDOntology#SequentialNeuralNetworkModel>	owl:NamedIndividual	<https://github.com/arashhosseiniarash/AFDDOntology#size of each layer>	70-15-7
<https://github.com/arashhosseiniarash/AFDDOntology#SequentialNeuralNetworkModel>	<https://github.com/arashhosseiniarash/AFDDOntology#HiddenLayerSize>	<https://github.com/arashhosseiniarash/AFDDOntology#size of each layer>	70-15-7
<https://github.com/arashhosseiniarash/AFDDOntology#SequentialNeuralNetworkModel>	owl:NamedIndividual	<https://github.com/arashhosseiniarash/AFDDOntology#test size>	0.25
<https://github.com/arashhosseiniarash/AFDDOntology#SequentialNeuralNetworkModel>	<https://github.com/arashhosseiniarash/AFDDOntology#TrainSamplesPerIteration>	<https://github.com/arashhosseiniarash/AFDDOntology#test size>	0.25
<https://github.com/arashhosseiniarash/AFDDOntology#SequentialNeuralNetworkModel>	owl:NamedIndividual	<https://github.com/arashhosseiniarash/AFDDOntology#3>	
<https://github.com/arashhosseiniarash/AFDDOntology#SequentialNeuralNetworkModel>	<https://github.com/arashhosseiniarash/AFDDOntology#HiddenLayerNumber>	<https://github.com/arashhosseiniarash/AFDDOntology#3>	
<https://github.com/arashhosseiniarash/AFDDOntology#SequentialNeuralNetworkModel>	owl:NamedIndividual	<https://github.com/arashhosseiniarash/AFDDOntology#600>	
<https://github.com/arashhosseiniarash/AFDDOntology#SequentialNeuralNetworkModel>	<https://github.com/arashhosseiniarash/AFDDOntology#Epochs>	<https://github.com/arashhosseiniarash/AFDDOntology#600>	
<https://github.com/arashhosseiniarash/AFDDOntology#SequentialNeuralNetworkModel>	owl:NamedIndividual	[illegible]	
<https://github.com/arashhosseiniarash/AFDDOntology#SequentialNeuralNetworkModel>	<https://github.com/arashhosseiniarash/AFDDOntology#Optimizer>	<https://github.com/arashhosseiniarash/AFDDOntology#adam>	
<https://github.com/arashhosseiniarash/AFDDOntology#SequentialNeuralNetworkModel>	owl:NamedIndividual	<https://github.com/arashhosseiniarash/AFDDOntology#categorical crossentropy>	
<https://github.com/arashhosseiniarash/AFDDOntology#SequentialNeuralNetworkModel>	<https://github.com/arashhosseiniarash/AFDDOntology#LossFunction>	<https://github.com/arashhosseiniarash/AFDDOntology#categorical crossentropy>	
<https://github.com/arashhosseiniarash/AFDDOntology#Svc>	owl:NamedIndividual	<https://github.com/arashhosseiniarash/AFDDOntology#svm scaling>	feature range -1 , 1
<https://github.com/arashhosseiniarash/AFDDOntology#Svc>	<https://github.com/arashhosseiniarash/AFDDOntology#Scale>	<https://github.com/arashhosseiniarash/AFDDOntology#svm scaling>	feature range -1 , 1
<https://github.com/arashhosseiniarash/AFDDOntology#Svc>	owl:NamedIndividual	<https://github.com/arashhosseiniarash/AFDDOntology#1>	
<https://github.com/arashhosseiniarash/AFDDOntology#Svc>	<https://github.com/arashhosseiniarash/AFDDOntology#C>	<https://github.com/arashhosseiniarash/AFDDOntology#1>	
<https://github.com/arashhosseiniarash/AFDDOntology#Svc>	owl:NamedIndividual	<https://github.com/arashhosseiniarash/AFDDOntology#rbf>	
<https://github.com/arashhosseiniarash/AFDDOntology#Svc>	<https://github.com/arashhosseiniarash/AFDDOntology#KernelType>	<https://github.com/arashhosseiniarash/AFDDOntology#rbf>	

Figure 6.3: Competency Question Constructed using SPARQL for Accessing AFDDOnto

Additional related information can be retrieved by combining queries which gives the type of algorithm used for AFDD, the respective parameters used for model development, and their respective values. Such information can be used to reproduce models or improve them. Further, the knowledge capture allows for an informed comparison among AFDD algorithms developed by allowing the operator to identify and compare the type, name and value of the parameters used for model development. The data residing in AFDDOnto can be stored in OWL, RDF, and TTL file formats. Most databases that enable SPARQL endpoints permit retrieval, access, and modification of AFDDOnto knowledge.

The import of analytics using the proposed methodology results in populating the knowledge model in the form of ABOX ontology, which means the knowledge graph is populated with assertations and individuals as opposed to TBOX, which constitutes the knowledge model itself, i.e., formal naming, definition of categories and properties [204]. This entitles the flow of data from AFDD to BIM. The ontology is populated with analytics and configurations from the AFDD model, including data extracted from the BIM.

AFDDOnto incorporates a maintenance concept within its ontology framework, allowing access to maintenance-related information and specifying the frequency of updates for stateful BIM under the Maintenance concept, which uses the ASHRAE's concepts [189] for the maintenance of HVAC. However, due to the nature of the BMS dataset, which contains about a month of data related to HVAC and is much shorter than the required maintenance schedule needed for stateful BIM demonstration.

The limited dataset makes it impossible to test the tracking of maintenance information, which would enable the generation of dynamic BIM features. The minimum duration needed for such a dataset is typically about three months to represent maintenance data of the RTU and VAV systems, but it is expected to show its significance when maintenance data is recorded for much longer durations to cover multiple maintenance. It must be noted that the features that are exported from BIM maintain their link to the BIM objects using the GUIDs. As features are inserted into the knowledge model, the BMS and BIM features in the model can be exported using the protégé export tool to a CSV file that constitutes the URI in the knowledge model for referencing and queries.

The knowledge model can be queried for information about features, maintenance, and AFDD configurations and further used for model comparison between models such as ANN and SVM, as done in this study. In the event of the availability of the dataset for a longer interval of time, the knowledge model can be periodically updated by analytics to represent the as-is state of the facility in the form of a Digital Twin (DT) that can present live state of the asset and building and as well the stateful information.

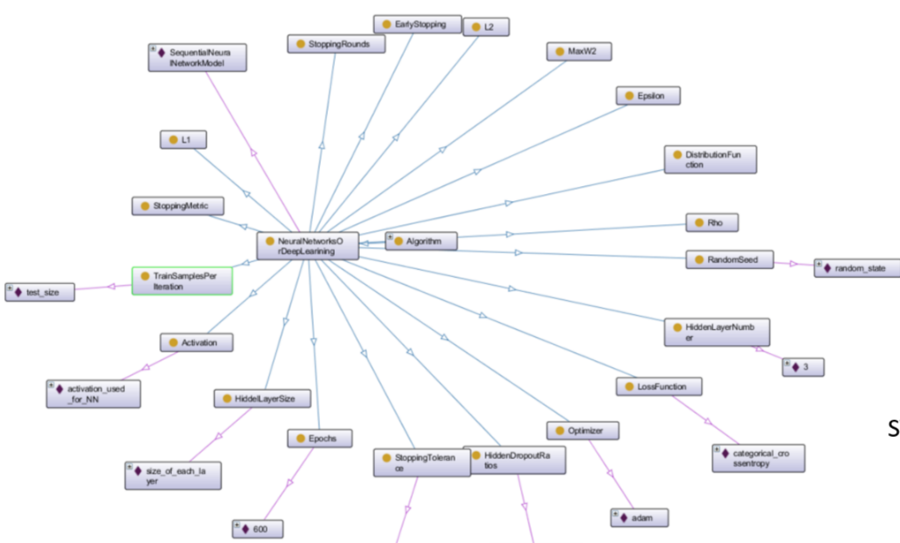
6.2 Visualizing BIM-Based AFDD Results in HVAC Systems with Knowledge Models

Currently, BIM does not incorporate analytics, and once the AFDD model is run, the model's configuration and analytics are not saved. This limits the effectiveness of comparisons and does not allow for the reproduction of the same models. However, using the proposed methodology, such information can be brought into the knowledge model. Knowledge graphs are preferred for visualizations at more abstract levels to identify the relationships between the classes.

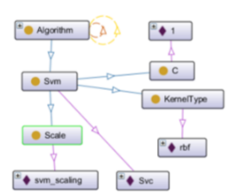
There are multiple tools for visualization using knowledge graphs; WebVOWL [205], for example, can be used to represent the graph online. Through the knowledge graph visualization, stakeholders can visually navigate and interact with the BIM and AFDD data, gaining a holistic view of the building's digital representation and its associated analytics. This visual representation enhances comprehension, facilitates data exploration, and supports decision-making processes related to HVAC system'' fault detection, diagnosis, and maintenance strategies.

For example, as shown in *Figure 6.4*, when the AFDD analytics were executed separately using the ANN and SVM model, both the model's analytics were imported into the knowledge model; the knowledge graph below shows RTU-1 for ANN and RTU-2 for the SVM model created where assertions are visualized using tools such as OWLViz and OntoGraf that can be found in Protégé and also can be listed.

The information provides the facility manager with a clear understanding of each parameter used for AFDD models, along with corresponding values and additional details, such as the accuracy of the model. This information can be utilized to develop similar models for similar facilities, taking into account HVAC systems, layouts, and analysis.



ANN configuration for 1-RTU



SVM configuration for 2-RTU

Figure 6.4: An Excerpt of the ABOX Knowledge Graph for Inserting ANN and SVM Model Configuration (Visualization through knowledge graphs)

This knowledge can be accessed using the BIM-based ontology through SPARQL queries, which enables retrieving information from knowledge graphs. This ability enables the asset manager to access the analytics, such as faults associated with the HVAC system, spaces, and the configuration to detect them. If access to data is needed, which in terms of knowledge models is known as individuals, then queries need to be constructed using SPARQL to retrieve the necessary data.

For example, a SPARQL query is written to allow the asset manager to identify the specific HVAC system in different rooms. The query output shown in Figure 6.5, provides the GUID reference that can be tracked in the IFC model. When coupled with maintenance information, such queries can provide additional information regarding the product's maintenance status, replacement, or life expectancy when BIM information is available.

```
prefix : <https://github.com/arashhosseiniarash/AFDDOntology#>
prefix owl: <http://www.w3.org/2002/07/owl#>
prefix rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
prefix xml: <http://www.w3.org/XML/1998/namespace>
prefix xsd: <http://www.w3.org/2001/XMLSchema#>
prefix rdfs: <http://www.w3.org/2000/01/rdf-schema#>
Prefix bot: <https://w3id.org/bot#>
Prefix brick:<https://brickschema.org/schema/Brick#>
SELECT DISTINCT ?space ?ELement_GUID ?Description
WHERE {
?space bot:containsElement ?ELement_GUID .
?ELement_GUID rdfs:comment ?Description
}
```

(a) SPARQL query construct to retrieve spaces that have HVAC equipment located in them

space	ELement_GUID	Description
Room-103	1Wp7aofZfE$Ravn2W39qi	"Return Diffuser"
Room-103	1Wp7aofZfE$Ravn2W39qi	"103"^^<http://www.w3.org/2001/XMLSchema#integer>
Room-204	1Wp7aofZfE$Ravn2W39qG	"204"^^<http://www.w3.org/2001/XMLSchema#integer>
Room-204	1Wp7aofZfE$Ravn2W39qG	"VE_VAV Box_Round_MEPcontent"
Room-202	1Wp7aofZfE$Ravn2W39qK	"VE_VAV Box_Round_MEPcontent"
Room-202	1Wp7aofZfE$Ravn2W39qK	"202"^^<http://www.w3.org/2001/XMLSchema#integer>
Room-105	1Wp7aofZfE$Ravn2W39qj	"Return Diffuser"
Room-105	1Wp7aofZfE$Ravn2W39qj	"105"^^<http://www.w3.org/2001/XMLSchema#integer>
Room-102	1Wp7aofZfE$Ravn2W39qT	"102"^^<http://www.w3.org/2001/XMLSchema#integer>
Room-102	1Wp7aofZfE$Ravn2W39qT	"VE_VAV Box_Round_MEPcontent"
Room-106	1Wp7aofZfE$Ravn2W39ql	"VE_VAV Box_Round_MEPcontent"
Room-106	1Wp7aofZfE$Ravn2W39ql	"106"^^<http://www.w3.org/2001/XMLSchema#integer>
Room-206	1Wp7aofZfE$Ravn2W39qV	"Return Diffuser"
Room-206	1Wp7aofZfE$Ravn2W39qV	"206"^^<http://www.w3.org/2001/XMLSchema#integer>
Room-205	1Wp7aofZfE$Ravn2W39qx	"205"^^<http://www.w3.org/2001/XMLSchema#integer>
Room-205	1Wp7aofZfE$Ravn2W39qx	"Return Diffuser"
Room-105	1Wp7aofZfE$Ravn2W39qs	"105"^^<http://www.w3.org/2001/XMLSchema#integer>
Room-105	1Wp7aofZfE$Ravn2W39qs	"VE_VAV Box_Round_MEPcontent"
Room-202	1Wp7aofZfE$Ravn2W39q_	"202"^^<http://www.w3.org/2001/XMLSchema#integer>
Room-202	1Wp7aofZfE$Ravn2W39q_	"Return Diffuser"
Room-206	1Wp7aofZfE$Ravn2W39qN	"206"^^<http://www.w3.org/2001/XMLSchema#integer>

(b) The retrieval of SPARQL query for both algorithms used in RTU case study

Figure 6.5: An Excerpt of the Results of Query Retrieval Showing Spaces in the Knowledge Model and Associated GUIDs in BIM

In addition to visualization using the knowledge graphs, the knowledge models allow for storing files and images or providing access to external files such as the BIM model. As an example, from the case study, a PNG file was used to capture a snippet image from the Revit model of the RTU. This image was then linked to the knowledge model and is displayed in *Figure 6.6*.

In practice, maintenance slips, product images, and manuals can be linked to models for referencing. However, IFC files, which can be utilized by IFC viewers, can also be used to link faulty entities and their associated metadata. In this research, access to files and images was performed manually, as automated conversion from IFCOWL to IFC was not available at the time of the study.

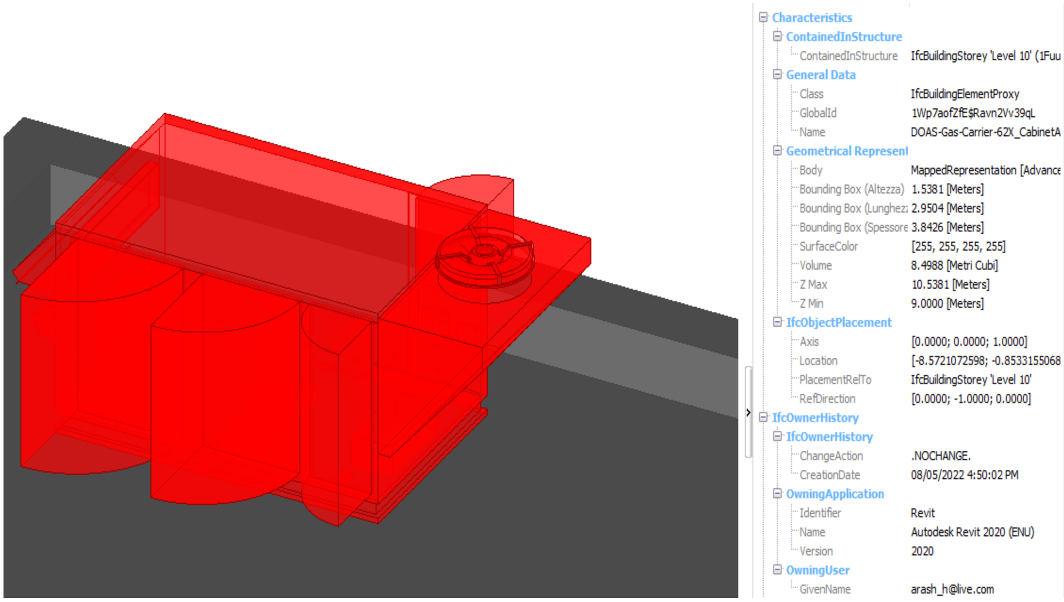


Figure 6.6: A Snippet of an Image Stored in the BIM-based Knowledge Model (Visualization of fault at system level through BIM)

6.3 Discussion

The proposed integrated model aims to improve decision-making capabilities by providing access to context-aware data and a comprehensive knowledge model, as well as spatial allocation of the output analytics. This interaction between the users and the knowledge model aims to optimize maintenance processes and elevate overall operational efficiency. Also, such an integrated solution bridges the gap between traditionally isolated systems, namely BIM and BMS.

As a semantic model of the building, BIM integrates the information and attributes of the building's physical and conceptual elements within a 3D/spatial representation. BIM is composed of various disciplines, i.e., architectural, structural, and MEP, as a single and coordinated model. Hence, BIM fosters accessibility and flow of information needed for HVAC and its maintenance and extends to various use cases applicable to the entire lifecycle of the building.

The proposed BIM-based automated FDD solution for HVAC systems is developed using the following three main components namely AFDDOnto, BIM, and BMS/BAS. The AFDDOnto leverages existing BIM and BMS as essential data repositories, which serve as the primary inputs for the proposed solution. The knowledge model incorporates the necessary axioms to capture essential concepts required for AFDD of HVAC systems, including features, fault types, algorithms, elements, and zones to effectively capture interrelationships.

Additionally, by adopting an ontology-based integration strategy and developing AFDDOnto, which is compatible with IFCOWL, interoperability with IFC, a standard in the industry, is maintained. This integrated model was made possible by following the proposed methodology, which includes schema conversion, feature engineering, machine learning for AFDD, and knowledge model management. These steps are crucial for creating the semantic model.

The applicability of the proposed solution was demonstrated through a case study that tested and validated AFDDOnto and the integration method. This case study showcased how dynamic BIM features, combined with sensory data from BMS, were used to enhance the dataset for AFDD model development. This highlighted the significance of the proposed solution, particularly in scenarios with limited sensory data availability

The proposed solution illustrates how BIM contextual information, such as spatial data including geometry and connectivity, can be effectively leveraged. The impact of BIM features, especially dynamic BIM, is anticipated to be more significant in real-world HVAC scenarios compared to controlled environments like the FRP case study. Real HVAC assets and their built environments involve multiple uncertainties not typically captured by BMS systems, such as environmental conditions, occupancy, and the state of building elements like doors and windows

Furthermore, the role of knowledge graphs and databases in enabling access to previously isolated resources of BIM and BMS is presented through AFDDOnto. This integrated model is designed to maintain connectivity with BIM and BMS using links maintained through URIs. The results illustrate a bi-directional flow of data and knowledge between BIM and BMS for AFDD. This enhances the BMS with enriched contextual information and transforms BIM from a static model to a dynamic one. This dynamic model can evolve into a Digital Twin of the facility by capturing FDD analytics and retaining connectivity to live data streams in BMS.

Chapter 7: Summary and Conclusion

AFDD is crucial for HVAC systems because it ensures efficient operation, reduces energy consumption, and minimizes maintenance costs. These methods are widely being adopted by both industry and academia for the fault detection and diagnosis of HVAC systems. They are generally easier to develop, faster, and often deployed for real-time FDD. Moreover, they improve over time as the model is trained on more data, thereby playing a crucial role in the field.

However, AFDD methods face significant challenges, particularly those that rely solely on sensory data, which is often limited in both number and type in existing buildings. Additionally, the lack of contextual information significantly hinders HVAC modelers and facility managers in effectively diagnosing and managing faults. Addressing these challenges is essential to fully realize the potential of AFDD in improving HVAC system performance and reliability.

The study findings suggest that the present shortcomings associated with AFDD of HVAC can be addressed by utilizing BIM in its dynamic form. This thesis resolved these challenges by defining the goal of developing a BIM-based semantic knowledge methodology applicable to the AFDD of HVAC systems in commercial buildings. This methodology aimed to incorporate dynamic features into the AFDD dataset generated using BIM and BMS data, addressing the challenges of limited sensory data availability and the lack of contextual information.

The developed solution enabled the flow of enriched data from BMS through BIM, compensating for the limited sensory data available for AFDD model development. Additionally, the solution addressed the issue of knowledge loss during AFDD model development by capturing this knowledge in the developed ontology, AFDDOnto. The developed ontology reused IFCOWL to ensure compatibility with the well-established IFC schema, capturing the relationships between HVAC systems and building spaces, and storing model parameters and deployment results. Furthermore, the developed solution enables visualization and access to this knowledge through well-established databases and knowledge graphs, while maintaining connectivity to BIM.

The proposed solution was developed and tested using a methodology and case study. To demonstrate its effectiveness, a case-representative facility was considered in this research, where both BIM data and BMS data were available. The generation of dynamic features illustrated the solution's effectiveness in scenarios with a limited number of sensors and demonstrated the application of BIM in generating context-aware features that can be supplemented to the AFDD dataset for model development. Although the model improvement may not be significant due to the controlled environment and new HVAC systems in the test facility, it is expected that in practice, as uncertainties related to building usage, occupant behavior, environmental factors, and the state of HVAC systems increase, the significance and impact of the model will become more pronounced.

7.1 Research Contributions

The novelty of this research is extracting additional contextual features to amend the BMS data for AFDD of HVAC. This is achieved by adopting both bottom-up (data-driven) and top-down (knowledge-based) approaches to AI, namely machine learning and ontology engineering, respectively. This dual approach captures knowledge in the form of AFDD analytics and updates the BIM-based ontology of the facility to archive and represent the status of the HVAC and its subsystems.

The major contribution of this research work is as follows:

- Utilizing BIM to support AFDD of HVAC by generating dynamic BIM features (context-aware) using BIM and BMS
- Capturing knowledge pertaining to AFDD of HVAC (Model development and deployment) In a BIM-based ontology (AFDDOnto)
- Creating a bi-directional flow of data and information between BIM and BAS/BMS for AFDD of HVAC

The developed integrated solution enables data-driven FDD methods to benefit from the dynamic features generated using BIM and BMS, utilizing the contextual information that resides in BIM. These features can compensate for buildings where limited sensor types are deployed. Additionally, BIM adds context in the form of building semantics, such as spatial relationships, to facilitate AFDD of HVAC systems.

Additionally, the developed BIM-based knowledge model streamlines the AFDD processes by providing users with access to configurations and points of reference (knowledge history) for each of the models created. It also enables the capture of BIM-based AFDD analytics by storing configurations used for AFDD model development, along with model evaluation results. Ultimately, the semantic model allows access to spatial inferences such as HVAC, location, state, and usage.

Furthermore, the developed methodology enables importing the AFDD analytics to the knowledge model without duplicating the data existing in its primary sources. Subsequently, the bidirectional nature of the developed method bridges the previously isolated BIM and BMS data sources by transforming BIM into a dynamic form and allowing BMS to utilize contextual information available in BIM.

Moreover, this unification enables seamless access to information for developing AFDD models using the historical knowledge captured by AFDDOnto, as well as accessing data using SPARQL queries to answer commonly asked questions by its users, which is typically a tedious process.

AFDDOnto is designed to streamline the storage, access, and retrieval of data, enabling the reuse of the ontology in future projects. By integrating concepts from both BIM and BMS/BAS, AFDDOnto enhances access to spatial inferences related to HVAC, location, state, and usage. The incorporation of dynamic BIM features into the AFDD model is expected to provide significant benefits to users such as AFDD model developers, building owners, facility managers, and asset managers, particularly when dealing with a limited set of sensors and uncertainties in building and HVAC systems.

The advantage of using the developed solution compared to existing AFDD models is that it enables the generation of context-aware features from a confined set of existing features. This can facilitate AFDD of HVAC systems and compensate for the lack of a diverse set of features. Additionally, it addresses the inherent limitations of such AFDD models, which often lack contextual information about the building and facility.

7.2 Limitations

The limitations of this research are discussed from two perspectives: research methodology limitations and case study limitations.

From the research methodology perspective, the extensive domain of HVAC maintenance means that the developed BIM-based knowledge model may not capture every aspect of user requirement, necessitating additional competencies for other use cases. Consequently, additional axioms in the form of concepts and relationships may be required. Furthermore, BMS and BAS schemas were not investigated as they were beyond the scope of this study, but maintaining direct connectivity with these systems could provide significant benefits. The developed ontology could also benefit from further refinement through testing on other BIM applications to ensure broader applicability beyond the AFDD of HVAC. Additionally, the conversion procedure from IFC to IFCOWL was not automated during the research, which prevented updates to the IFC file, a common BIM format widely adopted by AEC/FM (Architecture, Engineering, Construction, and Facility Management).

From the case study perspective, several limitations were identified that affected the demonstration of the application of dynamic BIM for AFDD of HVAC, specifically the impact of generated features and stateful BIM features.

However, it is important to note that the selection of this case study served a specific purpose. The aim was to illustrate the impact of adding BIM features to scenarios where a limited set of sensory data is available for the AFDD of HVAC. Thus, the case study enabled a focused analysis of purely data-driven solutions for various sensory data related to HVAC and its environment. Hence, it should be acknowledged that the magnitude at which the BIM features could affect the AFDD may not have been fully demonstrated under these specific conditions. The limitations of the case study are as follows;

Dataset: The duration needed for stateful BIM creation limited the ability to demonstrate dynamism through the used dataset to only Dynamic BIM (live BIM). The BIM feature's impact was limited by the controlled environment of the case study, which prevented a demonstration of the full potential of dynamic BIM such as and its impact on AFDD. This study can immensely benefit by testing the methodology on the dataset of actual buildings to determine and quantify the impact of dynamic BIM on AFDD for real commercial buildings. Ideally, a large dataset covering the entire life span of HVAC equipment is necessary.

Excessive instrumentation and monitoring: In the FRP case study, the energy management control system sets room temperature, scheduling, and other variables. The FRP data includes time-series data such as temperature, humidity, air supply, return airflow, and state representative sensors such as the state of the compressor, condenser, supply fan state (On or Off), and VAV reheating energy sensors that collect data in 1-, 15-, and 60-min frequencies. It also includes the FRP roof's dedicated weather station measured air temperature, humidity, solar radiation (direct normal, diffuse, and global), wind speed, and direction. Presently, most buildings do not possess such a high level of instrumentation and monitoring. As shown in this study, the spatial and dynamic BIM features extracted from BIM could be helpful to compensate for or improve the AFDD development in the absence of features needed for AFDD.

Operational constrain: The RTU-based HVAC system used in the FRP case study defines restrictions in the form of limits for refrigerant temperature, pressure, air temperature, air RH, airflow, and sensor data streams: status, commands, and control signals. The limit is set to increase the accuracy of FDD, whereas, in reality, the data can be erratic, corrupted, missing, and contain anomalies that are not necessarily faults. The BIM in dynamic form can be provisioned with building-specific context and baselines applicable to the facility, including access to maintenance and HVAC-related information.

Controlled environment and building envelope: The building used in the study imposes system environmental constraints such as having no blinds and using the insulated ground beneath the floor to emulate no sensible or hidden internal loads. Further, to limit the effect of uncertainty and lack of contextual information, the following simulated or predetermined parameters are introduced: Simulated occupancy, Specific RTU discharge temperature, and blocking outside air or exhaust air while maintaining static pressure.

All this indicates that environment and building parameters can significantly differ from a controlled environment for example, the occupants' use of windows and blinds adds uncertainty to the model, and having access to the as-is model of the dynamic BIM, can provide better assistance in decision making by proving access to the present state of the DT or having access to stateful information. Moreover, having access to dynamic BIM will enhance the applicability of AFDD models in actual buildings, as opposed to the high accuracies typically achievable only in controlled test facilities with limited environmental variations.

Fault imposition (experimental and simulated) method: Faulty and nonfaulty scenarios for the preliminary dataset, including HVAC systems, data types/facilities, and fault types, last for a predefined duration of one day, as the faults are introduced (Simulated mechanically) for the testing. In reality, the faults can be superimposed, simultaneous, intermittent, or continuous. When complexities related to faults occur, BIM information and visualization can be utilized to make informed decisions.

Additionally, in the developed methodology, BIM can serve as a point of reference for comparing different models, whether they are experimental or simulated. This allows for better assessment and validation of AFDD models in real-world scenarios.

Maintenance: The HVAC system and all its components in the FRP case study have been new and in perfect condition. Hence, no maintenance was carried out on the system, which reduces the uncertainty that the condition of the HVAC can have on FDD. In reality, the HVAC components can be repaired or replaced, which impacts the impact of the remaining life of the equipment and its condition on FDD. In such cases, the developed dynamic BIM can provide useful insight by tracking the maintenance changes in the model.

All the above stated indicates that due to the complexities of HVAC, most AFDD models show higher accuracy in test setups and not necessarily in real buildings. Hence, AFDD models may benefit from contextual information to compensate for the uncertainties associated with HVAC, building, occupants, and usage.

7.3 Future Work

The future direction of the present research effort is illustrated using Figure 7.1, which shows the application of the developed model for the DT version of the solution. It is expected that the BIM-based DT not only provides an ever-updated reality capture of the system but to be enriched with data, analytics, and essential knowledge for various use cases.

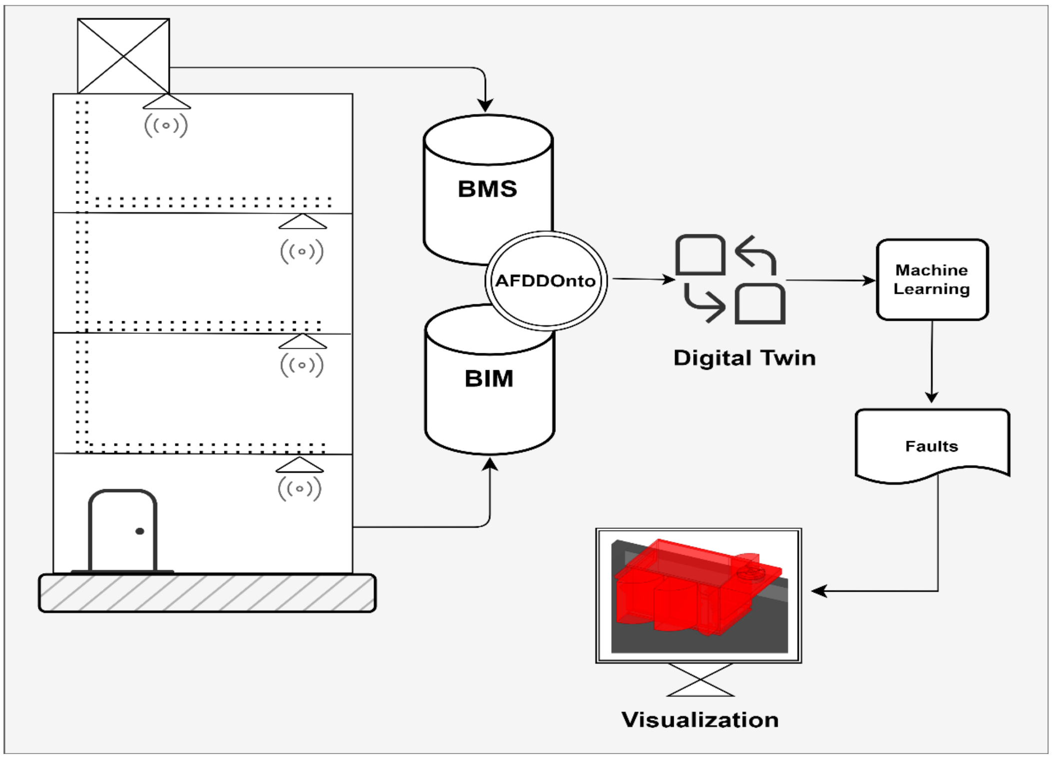


Figure 7.1: Application of BIM-based Knowledge Model for HVAC FDD in Digital Twin Form

The developed BIM-based DT can facilitate real-time monitoring and control of HVAC systems by being connected to BMS and BAS systems. This integration enables the facility manager to visualize the building and HVAC and further optimize HVAC system performance based on factors such as occupancy patterns, outdoor weather conditions, and sunlight orientation.

Also, having access to stateful information, such as HVAC usage data and equipment life expectancy, can be further utilized for tasks such as HVAC upkeep by connecting the facility data with manufacturer data, such as product catalogs, using linked data technologies. By leveraging this comprehensive data, the BIM-based DT can contribute to the overall efficiency and sustainability of HVAC systems, leading to energy savings and improved comfort levels.

The concept of dynamic BIM is relatively new, and while the conceptual aspects are well discussed and developed in the literature, the practical solutions and use cases in action are limited. The lack of application is mostly due to the absence of integration tools (between BIM, BMS, and AFDD algorithms, and the developed AFDDOnto in this thesis is one example of such tools and technologies.

Hence, there are currently no established industry standards or best practices for the development of BIM-based AFDD in HVAC systems. Products of this study enable other future technology developments to support such solutions in practice, to eventually contribute to future standard developments in this domain. The usage of the developed ontology-based knowledge model utilizing a semantic web framework is designed to improve compatibility, interoperability, and shareability among various stakeholders to facilitate further use cases.

Future work can investigate BMS/BAS systems and relevant schemas to identify how the developed model can be extended beyond AFDD of HVAC to provide support for HVAC maintenance and, ultimately, facility management. The combination of additional use cases can facilitate the generation of a digital twin of the facility.

Researchers could explore the integration of BMS/ BAS with the developed model to enhance its capabilities. By examining relevant schemas, it may be possible to extend the model's functionality to cover a broader range of HVAC maintenance tasks. This, in turn, could lead to more efficient and effective facility management practices.

Furthermore, incorporating additional use cases could enable the creation of a comprehensive digital twin of the facility. A digital twin would provide a dynamic, real-time representation of the physical building, allowing for improved monitoring, analysis, and optimization of various building systems. This approach could significantly enhance the overall management and maintenance of the facility, leading to better performance, reduced costs, and increased sustainability.

References

[1] "Major Energy Retrofit Guidelines for Commercial and Institutional Buildings – Non-food Retail," p. 59.

[2] S. C. Government of Canada, "Households and the Environment: Energy Use: Analysis." Accessed: Aug. 05, 2020. [Online]. Available: https://www150.statcan.gc.ca/n1/pub/11-526-s/2013002/part-partie1-eng.htm

[3] D. Chakraborty and H. Elzarka, "Early detection of faults in HVAC systems using an XGBoost model with a dynamic threshold," *Energy and Buildings*, vol. 185, pp. 326–344, Feb. 2019, doi: 10.1016/j.enbuild.2018.12.032.

[4] A. Beghi, R. Brignoli, L. Cecchinato, G. Menegazzo, and M. Rampazzo, "A data-driven approach for fault diagnosis in HVAC chiller systems," in *2015 IEEE Conference on Control Applications (CCA)*, Sydney, Australia: IEEE, Sep. 2015, pp. 966–971. doi: 10.1109/CCA.2015.7320737.

[5] R. Yan, Z. Ma, Y. Zhao, and G. Kokogiannakis, "A decision tree based data-driven diagnostic strategy for air handling units," *Energy and Buildings*, vol. 133, pp. 37–45, Dec. 2016, doi: 10.1016/j.enbuild.2016.09.039.

[6] S. Sun, G. Li, H. Chen, Q. Huang, S. Shi, and W. Hu, "A hybrid ICA-BPNN-based FDD strategy for refrigerant charge faults in variable refrigerant flow system," *Applied Thermal Engineering*, vol. 127, pp. 718–728, Dec. 2017, doi: 10.1016/j.applthermaleng.2017.08.047.

[7] Y. Chen and J. Wen, "A whole building fault detection using weather based pattern matching and feature based PCA method," in *2017 IEEE International Conference on Big Data (Big Data)*, Boston, MA: IEEE, Dec. 2017, pp. 4050–4057. doi: 10.1109/BigData.2017.8258421.

[8] S. Baldi, F. Zhang, T. Le Quang, P. Endel, and O. Holub, "Passive versus active learning in operation and adaptive maintenance of Heating, Ventilation, and Air Conditioning," *Applied Energy*, vol. 252, p. 113478, Oct. 2019, doi: 10.1016/j.apenergy.2019.113478.

[9] S. Frank, X. Jin, D. Studer, and A. Farthing, "Assessing barriers and research challenges for automated fault detection and diagnosis technology for small commercial buildings in the United States," *Renewable and Sustainable Energy Reviews*, vol. 98, pp. 489–499, Dec. 2018, doi: 10.1016/j.rser.2018.08.046.

[10] J. Winkler, J. Munk, and W. Hunt, "Barriers to Broader Utilization of Fault Detection Technologies for Improving Residential HVAC Equipment Efficiency," NREL/TP-5500-82024, 1862663, MainId:82797, Apr. 2022. doi: 10.2172/1862663.

[11] Y. Fan, X. Cui, H. Han, and H. Lu, "Chiller fault diagnosis with field sensors using the technology of imbalanced data," *Applied Thermal Engineering*, vol. 159, p. 113933, Aug. 2019, doi: 10.1016/j.applthermaleng.2019.113933.

[12] M. S. Mirnaghi and F. Haghighat, "Fault detection and diagnosis of large-scale HVAC systems in buildings using data-driven methods: A comprehensive review," *Energy and Buildings*, vol. 229, p. 110492, Dec. 2020, doi: 10.1016/j.enbuild.2020.110492.

[13] Q. Lu, X. Xie, A. K. Parlikad, and J. M. Schooling, "Digital twin-enabled anomaly detection for built asset monitoring in operation and maintenance," *Automation in Construction*, vol. 118, p. 103277, Oct. 2020, doi: 10.1016/j.autcon.2020.103277.

[14] B. Dong, Z. O'Neill, and Z. Li, "A BIM-enabled information infrastructure for building energy Fault Detection and Diagnostics," *Automation in Construction*, vol. 44, pp. 197–211, Aug. 2014, doi: 10.1016/j.autcon.2014.04.007.

[15] A. Andriamamonjy, D. Saelens, and R. Klein, "An auto-deployed model-based fault detection and diagnosis approach for Air Handling Units using BIM and Modelica,"

Automation in Construction, vol. 96, pp. 508–526, Dec. 2018, doi: 10.1016/j.autcon.2018.09.016.
[16] X. Xie, J. Merino, N. Moretti, P. Pauwels, J. Y. Chang, and A. Parlikad, "Digital twin enabled fault detection and diagnosis process for building HVAC systems," *Automation in Construction*, vol. 146, p. 104695, Feb. 2023, doi: 10.1016/j.autcon.2022.104695.
[17] X. Yang and S. Ergan, "BIM for FM: Information Requirements to Support HVAC-Related Corrective Maintenance," *J. Archit. Eng.*, vol. 23, no. 4, p. 04017023, Dec. 2017, doi: 10.1061/(ASCE)AE.1943-5568.0000272.
[18] A. Zabin, V. A. González, Y. Zou, and R. Amor, "Applications of machine learning to BIM: A systematic literature review," *Advanced Engineering Informatics*, vol. 51, p. 101474, Jan. 2022, doi: 10.1016/j.aei.2021.101474.
[19] "A Reference Architecture for Data-Driven Smart Buildings Using Brick and LBD Ontologies."
[20] H. H. Hosamo, P. R. Svennevig, K. Svidt, D. Han, and H. K. Nielsen, "A Digital Twin predictive maintenance framework of air handling units based on automatic fault detection and diagnostics," *Energy and Buildings*, vol. 261, p. 111988, Apr. 2022, doi: 10.1016/j.enbuild.2022.111988.
[21] Z. Zhang, H. Han, X. Cui, and Y. Fan, "Novel application of multi-model ensemble learning for fault diagnosis in refrigeration systems," *Applied Thermal Engineering*, vol. 164, p. 114516, Jan. 2020, doi: 10.1016/j.applthermaleng.2019.114516.
[22] S. Deshmukh, L. Glicksman, and L. Norford, "Case study results: fault detection in air-handling units in buildings," *Advances in Building Energy Research*, pp. 1–17, Nov. 2018, doi: 10.1080/17512549.2018.1545143.
[23] Z. Shi and W. O'Brien, "Development and implementation of automated fault detection and diagnostics for building systems: A review," *Automation in Construction*, vol. 104, pp. 215–229, Aug. 2019, doi: 10.1016/j.autcon.2019.04.002.
[24] A. Afram and F. Janabi-Sharifi, "Review of modeling methods for HVAC systems," *Applied Thermal Engineering*, vol. 67, no. 1–2, pp. 507–519, Jun. 2014, doi: 10.1016/j.applthermaleng.2014.03.055.
[25] W. Kim and S. Katipamula, "A review of fault detection and diagnostics methods for building systems," *Science and Technology for the Built Environment*, vol. 24, no. 1, pp. 3–21, Jan. 2018, doi: 10.1080/23744731.2017.1318008.
[26] Y. Zhao, T. Li, X. Zhang, and C. Zhang, "Artificial intelligence-based fault detection and diagnosis methods for building energy systems: Advantages, challenges and the future," *Renewable and Sustainable Energy Reviews*, vol. 109, pp. 85–101, Jul. 2019, doi: 10.1016/j.rser.2019.04.021.
[27] A. P. Rogers, F. Guo, and B. P. Rasmussen, "A review of fault detection and diagnosis methods for residential air conditioning systems," *Building and Environment*, vol. 161, p. 106236, Aug. 2019, doi: 10.1016/j.buildenv.2019.106236.
[28] Y. Zhao, C. Zhang, Y. Zhang, Z. Wang, and J. Li, "A review of data mining technologies in building energy systems: Load prediction, pattern identification, fault detection and diagnosis," *Energy and Built Environment*, vol. 1, no. 2, pp. 149–164, Apr. 2020, doi: 10.1016/j.enbenv.2019.11.003.
[29] Y. Li and Z. O'Neill, "A critical review of fault modeling of HVAC systems in buildings," *Build. Simul.*, vol. 11, no. 5, pp. 953–975, Oct. 2018, doi: 10.1007/s12273-018-0458-4.

[30] M. Bang, S. S. Engelsgaard, E. K. Alexandersen, M. Riber Skydt, H. R. Shaker, and M. Jradi, “Novel real-time model-based fault detection method for automatic identification of abnormal energy performance in building ventilation units,” *Energy and Buildings*, vol. 183, pp. 238–251, Jan. 2019, doi: 10.1016/j.enbuild.2018.11.006.

[31] H. Han, X. Cui, Y. Fan, and H. Qing, “Least squares support vector machine (LS-SVM)-based chiller fault diagnosis using fault indicative features,” *Applied Thermal Engineering*, vol. 154, pp. 540–547, May 2019, doi: 10.1016/j.applthermaleng.2019.03.111.

[32] A. Ranade, G. Provan, A. El-Din Mady, and D. O’Sullivan, “A computationally efficient method for fault diagnosis of fan-coil unit terminals in building Heating Ventilation and Air Conditioning systems,” *Journal of Building Engineering*, vol. 27, p. 100955, Jan. 2020, doi: 10.1016/j.jobe.2019.100955.

[33] N/A, “2013 ASHRAE Handbook - Fundamentals (SI Edition)”, [Online]. Available: https://app-knovel-com.lib-ezproxy.concordia.ca/hotlink/toc/id:kpASHRAEC1/ashrae-handbook-fundamentals/ashrae-handbook-fundamentals

[34] G. Serale, M. Fiorentini, A. Capozzoli, D. Bernardini, and A. Bemporad, “Model Predictive Control (MPC) for Enhancing Building and HVAC System Energy Efficiency: Problem Formulation, Applications and Opportunities,” *Energies*, vol. 11, no. 3, p. 631, Mar. 2018, doi: 10.3390/en11030631.

[35] A. Beghi, L. Cecchinato, F. Peterle, M. Rampazzo, and F. Simmini, “Model-based fault detection and diagnosis for centrifugal chillers,” in *2016 3rd Conference on Control and Fault-Tolerant Systems (SysTol)*, Barcelona, Spain: IEEE, Sep. 2016, pp. 158–163. doi: 10.1109/SYSTOL.2016.7739744.

[36] A. Beghi, R. Brignoli, L. Cecchinato, G. Menegazzo, M. Rampazzo, and F. Simmini, “Data-driven Fault Detection and Diagnosis for HVAC water chillers,” *Control Engineering Practice*, vol. 53, pp. 79–91, Aug. 2016, doi: 10.1016/j.conengprac.2016.04.018.

[37] K. Yan, J. Huang, W. Shen, and Z. Ji, “Unsupervised learning for fault detection and diagnosis of air handling units,” *Energy and Buildings*, vol. 210, p. 109689, Mar. 2020, doi: 10.1016/j.enbuild.2019.109689.

[38] G. Li and Y. Hu, “An enhanced PCA-based chiller sensor fault detection method using ensemble empirical mode decomposition based denoising,” *Energy and Buildings*, vol. 183, pp. 311–324, Jan. 2019, doi: 10.1016/j.enbuild.2018.10.013.

[39] D. A. T. Tran, Y. Chen, M. Q. Chau, and B. Ning, “A robust online fault detection and diagnosis strategy of centrifugal chiller systems for building energy efficiency,” *Energy and Buildings*, vol. 108, pp. 441–453, Dec. 2015, doi: 10.1016/j.enbuild.2015.09.044.

[40] Z. Li *et al.*, “An efficient online wkNN diagnostic strategy for variable refrigerant flow system based on coupled feature selection method,” *Energy and Buildings*, vol. 183, pp. 222–237, Jan. 2019, doi: 10.1016/j.enbuild.2018.11.020.

[41] D. Bigaud, A. Charki, A. Caucheteux, F. Titikpina, and T. Tiplica, “Detection of Faults and Drifts in the Energy Performance of a Building Using Bayesian Networks,” *Journal of Dynamic Systems, Measurement, and Control*, vol. 141, no. 10, p. 101011, Oct. 2019, doi: 10.1115/1.4043922.

[42] S. Katipamula and M. Brambley, “Review Article: Methods for Fault Detection, Diagnostics, and Prognostics for Building Systems—A Review, Part I,” *HVAC&R Res.*, vol. 11, no. 1, pp. 3–25, Jan. 2005, doi: 10.1080/10789669.2005.10391123.

[43] D. Li, G. Hu, and C. J. Spanos, "A data-driven strategy for detection and diagnosis of building chiller faults using linear discriminant analysis," *Energy and Buildings*, vol. 128, pp. 519–529, Sep. 2016, doi: 10.1016/j.enbuild.2016.07.014.

[44] S. Miyata, J. Lim, Y. Akashi, Y. Kuwahara, and K. Tanaka, "Fault detection and diagnosis for heat source system using convolutional neural network with imaged faulty behavior data," *Science and Technology for the Built Environment*, vol. 26, no. 1, pp. 52–60, Jan. 2020, doi: 10.1080/23744731.2019.1651619.

[45] R. Chiosa, M. S. Piscitelli, C. Fan, and A. Capozzoli, "Towards a self-tuned data analytics-based process for an automatic context-aware detection and diagnosis of anomalies in building energy consumption timeseries," *Energy and Buildings*, vol. 270, p. 112302, Sep. 2022, doi: 10.1016/j.enbuild.2022.112302.

[46] F. Xiao, Y. Zhao, J. Wen, and S. Wang, "Bayesian network based FDD strategy for variable air volume terminals," *Automation in Construction*, vol. 41, pp. 106–118, May 2014, doi: 10.1016/j.autcon.2013.10.019.

[47] H. Wang, Y. Chen, C. W. H. Chan, J. Qin, and J. Wang, "Online model-based fault detection and diagnosis strategy for VAV air handling units," *Energy and Buildings*, vol. 55, pp. 252–263, Dec. 2012, doi: 10.1016/j.enbuild.2012.08.016.

[48] A. Rosato, F. Guarino, V. Filomena, S. Sibilio, and L. Maffei, "Experimental Calibration and Validation of a Simulation Model for Fault Detection of HVAC Systems and Application to a Case Study," *Energies*, vol. 13, no. 15, p. 3948, Aug. 2020, doi: 10.3390/en13153948.

[49] M. Bonvini, M. D. Sohn, J. Granderson, M. Wetter, and M. A. Piette, "Robust on-line fault detection diagnosis for HVAC components based on nonlinear state estimation techniques," *Applied Energy*, vol. 124, pp. 156–166, Jul. 2014, doi: 10.1016/j.apenergy.2014.03.009.

[50] K. Yan, X. Chen, X. Zhou, Z. Yan, and J. Ma, "Physical Model Informed Fault Detection and Diagnosis of Air Handling Units Based on Transformer Generative Adversarial Network," *IEEE Trans. Ind. Inf.*, vol. 19, no. 2, pp. 2192–2199, Feb. 2023, doi: 10.1109/TII.2022.3193733.

[51] B. Gunay, B. W. Hobson, D. Darwazeh, and J. Bursill, "Estimating energy savings from HVAC controls fault correction through inverse greybox model-based virtual metering," *Energy and Buildings*, vol. 282, p. 112806, Mar. 2023, doi: 10.1016/j.enbuild.2023.112806.

[52] A. Hosseini Gourabpasi and M. Nik-Bakht, "Knowledge Discovery by Analyzing the State of the Art of Data-Driven Fault Detection and Diagnostics of Building HVAC," *CivilEng*, vol. 2, no. 4, pp. 986–1008, Nov. 2021, doi: 10.3390/civileng2040053.

[53] K. Verbert, R. Babuška, and B. De Schutter, "Combining knowledge and historical data for system-level fault diagnosis of HVAC systems," *Engineering Applications of Artificial Intelligence*, vol. 59, pp. 260–273, Mar. 2017, doi: 10.1016/j.engappai.2016.12.021.

[54] K. Haruehansapong, W. Roungprom, M. Kliangkhlao, K. Yeranee, and B. Sahoh, "Deep Learning-Driven Automated Fault Detection and Diagnostics Based on a Contextual Environment: A Case Study of HVAC System," *Buildings*, vol. 13, no. 1, p. 27, Dec. 2022, doi: 10.3390/buildings13010027.

[55] M. C. Comstock, J. E. Braun, and E. A. Groll, "A survey of common faults for chillers / Discussion," *ASHRAE Transactions*, vol. 108, p. 819, 2002.

[56] K. W. Roth, D. Westphalen, P. Llana, and M. Feng, "The Energy Impact of Faults in U.S. Commercial Buildings," p. 9, 2004.

[57] S. M. Frank, J. Kim, J. Cai, and J. E. Braun, “Common Faults and Their Prioritization in Small Commercial Buildings: February 2017 - December 2017,” NREL/SR--5500-70136, 1457127, Jun. 2018. doi: 10.2172/1457127.
[58] K. W. Roth, D. Westphalen, M. Y. Feng, P. Llana, and L. Quartararo, “Energy Impact of Commercial Building Controls and Performance Diagnostics: Market Characterization, Energy Impact of Building Faults and Energy Savings Potential,” p. 413.
[59] Wei Zhang, Hongzhi Liao, and Na Zhao, “Research on the FP Growth Algorithm about Association Rule Mining,” in *2008 International Seminar on Business and Information Management*, Wuhan: IEEE, Dec. 2008, pp. 315–318. doi: 10.1109/ISBIM.2008.177.
[60] R. Agrawal, T. Imieliński, and A. Swami, “Mining association rules between sets of items in large databases,” *SIGMOD Rec.*, vol. 22, no. 2, pp. 207–216, Jun. 1993, doi: 10.1145/170036.170072.
[61] “Semantic Web - W3C.” Accessed: Nov. 30, 2020. [Online]. Available: https://www.w3.org/standards/semanticweb/
[62] M. H. Rasmussen, M. Lefrançois, P. Pauwels, C. A. Hviid, and J. Karlshøj, “Managing interrelated project information in AEC Knowledge Graphs,” *Automation in Construction*, vol. 108, p. 102956, Dec. 2019, doi: 10.1016/j.autcon.2019.102956.
[63] G. T. Fierro, “Design of an Effective Ontology and Query Processor Enabling Portable Building Applications,” 2019.
[64] B. Butzin, F. Golatowski, and D. Timmermann, “A survey on information modeling and ontologies in building automation,” in *IECON 2017 - 43rd Annual Conference of the IEEE Industrial Electronics Society*, Beijing: IEEE, Oct. 2017, pp. 8615–8621. doi: 10.1109/IECON.2017.8217514.
[65] L. Z. Ghomari and A. R. Ghomari, “Ontology versus terminology, from the perspective of ontologists,” *IJWS*, vol. 1, no. 4, p. 315, 2012, doi: 10.1504/IJWS.2012.052531.
[66] Y. Kalfoglou and M. Schorlemmer, “Ontology mapping: the state of the art,” *The Knowledge Engineering Review*, vol. 18, no. 1, pp. 1–31, Jan. 2003, doi: 10.1017/S0269888903000651.
[67] T. A. El-Diraby and S. M. Gill, “A taxonomy for construction terms in privatized-infrastructure finance: supporting semantic exchange of project risk information,” *Construction Management and Economics*, vol. 24, no. 3, pp. 271–285, Mar. 2006, doi: 10.1080/01446190500434971.
[68] “Practical Knowledge Modelling: Ontology Development 101,” Udemy. Accessed: Oct. 28, 2021. [Online]. Available: https://www.udemy.com/course/practical-knowledge-modelling/
[69] P. Delgoshaei, M. A. Austin, and D. Veronica, “Semantic Models and Rule-based Reasoning for Fault Detection and Diagnostics: Applications in Heating, Ventilating and Air Conditioning Systems,” p. 6, 2017.
[70] “RDF-Based Semantics - OWL.” Accessed: Nov. 02, 2021. [Online]. Available: https://www.w3.org/2007/OWL/wiki/RDF-Based_Semantics
[71] G. F. Schneider, “Automated Ontology Matching in the Architecture, Engineering and Construction Domain - A Case Study,” p. 15.
[72] J. Euzenat, A. Mocan, and F. Scharffe, “Ontology Alignments,” in *Ontology Management*, vol. 7, M. Hepp, P. Leenheer, A. Moor, and Y. Sure, Eds., in Computing for Human Experience, vol. 7. , Boston, MA: Springer US, 2008, pp. 177–206. doi: 10.1007/978-0-387-69900-4_6.

[73] "Data - W3C." Accessed: Sep. 19, 2019. [Online]. Available: https://www.w3.org/standards/semanticweb/data

[74] L. Liu and M. T. Özsu, Eds., *Encyclopedia of Database Systems*. New York, NY: Springer New York, 2018. doi: 10.1007/978-1-4614-8265-9.

[75] B. Ben Mahria, I. Chaker, and A. Zahi, "A novel approach for learning ontology from relational database: from the construction to the evaluation," *J Big Data*, vol. 8, no. 1, p. 25, Dec. 2021, doi: 10.1186/s40537-021-00412-2.

[76] "Standard 224-2023 -- Standard for the Application of Building Information Modeling (ANSI Approved)." Accessed: Mar. 09, 2024. [Online]. Available: https://www.techstreet.com/ashrae/standards/ashrae-224-2023?product_id=2576553

[77] K. Kim, H. Kim, W. Kim, C. Kim, J. Kim, and J. Yu, "Integration of ifc objects and facility management work information using Semantic Web," *Automation in Construction*, vol. 87, pp. 173–187, Mar. 2018, doi: 10.1016/j.autcon.2017.12.019.

[78] F. Abdoul-Wali, "Fault Detection and Localization using IFC: A Case Study of BIM-based Visualisation of BAS-related Faults," *CIBSE Technical Symposium*, p. 17, 2019.

[79] Y.-C. Lee, C. M. Eastman, and W. Solihin, "An ontology-based approach for developing data exchange requirements and model views of building information modeling," *Advanced Engineering Informatics*, vol. 30, no. 3, pp. 354–367, Aug. 2016, doi: 10.1016/j.aei.2016.04.008.

[80] A. Motamedi, A. Hammad, and Y. Asen, "Knowledge-assisted BIM-based visual analytics for failure root cause detection in facilities management," *Automation in Construction*, vol. 43, pp. 73–83, Jul. 2014, doi: 10.1016/j.autcon.2014.03.012.

[81] B. Dong, Z. O'Neill, and Z. Li, "A BIM-enabled information infrastructure for building energy Fault Detection and Diagnostics," *Automation in Construction*, vol. 44, pp. 197–211, Aug. 2014, doi: 10.1016/j.autcon.2014.04.007.

[82] G. F. Schneider, Y. Kalantari, G. Kontes, S. Steiger, and D. Rovas, "AN ONTOLOGY-BASED TOOL FOR AUTOMATED CONFIGURATION AND DEPLOYMENT OF TECHNICAL BUILDING MANAGEMENT SERVICES," presented at the Proceedings of the Central European Symposium on Building Physics, 2016, p. 9.

[83] P. Pauwels, S. Zhang, and Y.-C. Lee, "Semantic web technologies in AEC industry: A literature overview," *Automation in Construction*, vol. 73, pp. 145–165, Jan. 2017, doi: 10.1016/j.autcon.2016.10.003.

[84] H. Dibowski, J. Vass, O. Holub, and J. Rojicek, "Automatic setup of fault detection algorithms in building and home automation," in *2016 IEEE 21st International Conference on Emerging Technologies and Factory Automation (ETFA)*, Berlin, Germany: IEEE, Sep. 2016, pp. 1–6. doi: 10.1109/ETFA.2016.7733622.

[85] H. Dibowski, O. Holub, and J. Rojicek, "Knowledge-Based Fault Propagation in Building Automation Systems," in *2016 International Conference on Systems Informatics, Modelling and Simulation (SIMS)*, Riga, Latvia: IEEE, Jun. 2016, pp. 124–132. doi: 10.1109/SIMS.2016.22.

[86] A. Mallak, C. Weber, M. Fathi, A. Behravan, and R. Obermaisser, "A Graph-Based Sensor Fault Detection and Diagnosis for Demand-Controlled Ventilation Systems Extracted from a Semantic Ontology," in *2018 IEEE 22nd International Conference on Intelligent Engineering Systems (INES)*, Las Palmas de Gran Canaria: IEEE, Jun. 2018, pp. 000377–000382. doi: 10.1109/INES.2018.8523895.

[87] B. Balaji *et al.*, "Brick : Metadata schema for portable smart building applications," *Applied Energy*, vol. 226, pp. 1273–1292, Sep. 2018, doi: 10.1016/j.apenergy.2018.02.091.
[88] "HVAC – Project Haystack." Accessed: Nov. 19, 2020. [Online]. Available: https://project-haystack.org/tag/hvac
[89] "SAREF extension for building." Accessed: Nov. 19, 2020. [Online]. Available: https://saref.etsi.org/saref4bldg/v1.1.2/
[90] "Google/digitalbuildings." Accessed: Dec. 01, 2020. [Online]. Available: https://github.com/google/digitalbuildings
[91] P. Pauwels and W. Terkaj, "EXPRESS to OWL for construction industry: Towards a recommendable and usable ifcOWL ontology," *Automation in Construction*, vol. 63, pp. 100–133, Mar. 2016, doi: 10.1016/j.autcon.2015.12.003.
[92] M. H. Rasmussen, P. Pauwels, M. Lefrançois, G. F. Schneider, C. A. Hviid, and J. Karlshøj, "Recent changes in the Building Topology Ontology," 2017, doi: 10.13140/RG.2.2.32365.28647.
[93] "Home - BrickSchema." Accessed: Nov. 19, 2020. [Online]. Available: https://brickschema.org/
[94] M. H. Rasmussen, M. Lefrançois, P. Pauwels, C. A. Hviid, and J. Karlshøj, "Managing interrelated project information in AEC Knowledge Graphs," *Automation in Construction*, vol. 108, p. 102956, Dec. 2019, doi: 10.1016/j.autcon.2019.102956.
[95] "ifcOWL," buildingSMART Technical. Accessed: Jan. 28, 2024. [Online]. Available: https://technical.buildingsmart.org/standards/ifc/ifc-formats/ifcowl/
[96] "Industry Foundation Classes (IFC)," buildingSMART Technical. Accessed: Jan. 28, 2024. [Online]. Available: https://technical.buildingsmart.org/standards/ifc/
[97] S. Tang, D. R. Shelden, C. M. Eastman, P. Pishdad-Bozorgi, and X. Gao, "A review of building information modeling (BIM) and the internet of things (IoT) devices integration: Present status and future trends," *Automation in Construction*, vol. 101, pp. 127–139, May 2019, doi: 10.1016/j.autcon.2019.01.020.
[98] A. Schumann, J. Hayes, P. Pompey, and O. Verscheure, "Adaptable Fault Identification for Smart Buildings," p. 4.
[99] G. Ramanathan, M. Husmann, C. Niedermeier, N. Vicari, K. Garcia, and S. Mayer, "Assisting automated fault detection and diagnostics in building automation through semantic description of functions and process data," in *Proceedings of the 8th ACM International Conference on Systems for Energy-Efficient Buildings, Cities, and Transportation*, Coimbra Portugal: ACM, Nov. 2021, pp. 228–229. doi: 10.1145/3486611.3492230.
[100] H. Dibowski, O. Holub, and J. Rojicek, "Ontology-based automatic setup of virtual sensors in building automation systems," in *2016 8th International Congress on Ultra Modern Telecommunications and Control Systems and Workshops (ICUMT)*, Lisbon, Portugal: IEEE, Oct. 2016, pp. 375–381. doi: 10.1109/ICUMT.2016.7765388.
[101] M. Valinejadshoubi, O. Moselhi, A. Bagchi, and A. Salem, "Development of an IoT and BIM-based automated alert system for thermal comfort monitoring in buildings," *Sustainable Cities and Society*, vol. 66, p. 102602, Mar. 2021, doi: 10.1016/j.scs.2020.102602.
[102] E. A. Pärn and D. J. Edwards, "Conceptualising the FinDD API plug-in: A study of BIM-FM integration," *Automation in Construction*, vol. 80, pp. 11–21, Aug. 2017, doi: 10.1016/j.autcon.2017.03.015.

[103] Z. Riaz, M. Arslan, A. K. Kiani, and S. Azhar, “CoSMoS: A BIM and wireless sensor based integrated solution for worker safety in confined spaces,” *Automation in Construction*, vol. 45, pp. 96–106, Sep. 2014, doi: 10.1016/j.autcon.2014.05.010.
[104] W. Terkaj, G. F. Schneider, and P. Pauwels, “Reusing Domain Ontologies in Linked Building Data: the Case of Building Automation and Control,” p. 12.
[105] K. Kim, H. Kim, W. Kim, C. Kim, J. Kim, and J. Yu, “Integration of ifc objects and facility management work information using Semantic Web,” *Automation in Construction*, vol. 87, pp. 173–187, Mar. 2018, doi: 10.1016/j.autcon.2017.12.019.
[106] R. Eftekharirad, M. Nik-Bakht, and A. Hammad, “Extending IFC for Fire Emergency Real-Time Management Using Sensors and Occupant Information,” p. 9, 2018.
[107] J. V. Moreno, R. Machete, A. P. Falcão, A. B. Gonçalves, and R. Bento, “Dynamic Data Feeding into BIM for Facility Management: A Prototype Application to a University Building,” *Buildings*, vol. 12, no. 5, p. 645, May 2022, doi: 10.3390/buildings12050645.
[108] A. H. Gourabpasi and M. Nik-Bakht, “Stateful BIM(BIM) for Digital twinning of the built environment,” in *Next-Generation Cities Encyclopedia*, vol. Volume 2: Built and Natural Environment, In press.
[109] “GitHub - arashhosseiniarash/AFDDOntology.” Accessed: Jan. 28, 2024. [Online]. Available: https://github.com/arashhosseiniarash/AFDDOntology
[110] R. G. Kreider and J. I. Messner, “The Uses of BIM: Classifying and Selecting BIM Uses. Version 0.9.” The Pennsylvania State University, University Park, PA, USA, Sep. 2013. [Online]. Available: http://bim.psu.edu
[111] T. Han, T. Ma, Z. Fang, Y. Zhang, and C. Han, “A BIM-IoT and intelligent compaction integrated framework for advanced road compaction quality monitoring and management,” *Computers and Electrical Engineering*, vol. 100, p. 107981, May 2022, doi: 10.1016/j.compeleceng.2022.107981.
[112] K. Kang, J. Lin, and J. Zhang, “BIM- and IoT-based monitoring framework for building performance management,” *Journal of Structural Integrity and Maintenance*, vol. 3, no. 4, pp. 254–261, Oct. 2018, doi: 10.1080/24705314.2018.1536318.
[113] K. McGlinn, B. Yuce, H. Wicaksono, S. Howell, and Y. Rezgui, “Usability evaluation of a web-based tool for supporting holistic building energy management,” *Automation in Construction*, vol. 84, pp. 154–165, Dec. 2017, doi: 10.1016/j.autcon.2017.08.033.
[114] A. H. Oti, E. Kurul, F. Cheung, and J. H. M. Tah, “A framework for the utilization of Building Management System data in building information models for building design and operation,” *Automation in Construction*, vol. 72, pp. 195–210, Dec. 2016, doi: 10.1016/j.autcon.2016.08.043.
[115] M. Shahinmoghadam, W. Natephra, and A. Motamedi, “BIM- and IoT-based virtual reality tool for real-time thermal comfort assessment in building enclosures,” *Building and Environment*, vol. 199, p. 107905, Jul. 2021, doi: 10.1016/j.buildenv.2021.107905.
[116] H. Yin, “Building Management System to support building renovation,” p. 6, 2010.
[117] P. Zangeneh and B. McCabe, “Ontology-based knowledge representation for industrial megaprojects analytics using linked data and the semantic web,” *Advanced Engineering Informatics*, vol. 46, p. 101164, Oct. 2020, doi: 10.1016/j.aei.2020.101164.
[118] W. Natephra, A. Motamedi, N. Yabuki, and T. Fukuda, “Integrating 4D thermal information with BIM for building envelope thermal performance analysis and thermal comfort evaluation in naturally ventilated environments,” *Building and Environment*, vol. 124, pp. 194–208, Nov. 2017, doi: 10.1016/j.buildenv.2017.08.004.

[119] M.-Y. Cheng and N.-W. Chang, “Dynamic construction material layout planning optimization model by integrating 4D BIM,” *Engineering with Computers*, vol. 35, no. 2, pp. 703–720, Apr. 2019, doi: 10.1007/s00366-018-0628-0.

[120] Ping-Sun Chan, H.-Y. Chan, and P.-H. Yuen, “BIM-enabled streamlined fault localization with system topology, RFID technology and real-time data acquisition interfaces,” in *2016 IEEE International Conference on Automation Science and Engineering (CASE)*, Fort Worth, TX, USA: IEEE, Aug. 2016, pp. 815–820. doi: 10.1109/COASE.2016.7743486.

[121] M. Hamid, O. Tolba, and A. El Antably, “BIM semantics for digital fabrication: A knowledge-based approach,” *Automation in Construction*, vol. 91, pp. 62–82, Jul. 2018, doi: 10.1016/j.autcon.2018.02.031.

[122] T. Fukuda, K. Yokoi, N. Yabuki, and A. Motamedi, “An indoor thermal environment design system for renovation using augmented reality,” *Journal of Computational Design and Engineering*, vol. 6, no. 2, pp. 179–188, Apr. 2019, doi: 10.1016/j.jcde.2018.05.007.

[123] A. Andriamamonjy, D. Saelens, and R. Klein, “An auto-deployed model-based fault detection and diagnosis approach for Air Handling Units using BIM and Modelica,” *Automation in Construction*, vol. 96, pp. 508–526, Dec. 2018, doi: 10.1016/j.autcon.2018.09.016.

[124] H. Zahid, O. Elmansoury, and R. Yaagoubi, “Dynamic Predicted Mean Vote: An IoT-BIM integrated approach for indoor thermal comfort optimization,” *Automation in Construction*, vol. 129, p. 103805, Sep. 2021, doi: 10.1016/j.autcon.2021.103805.

[125] M.-Y. Cheng, K.-C. Chiu, Y.-M. Hsieh, I.-T. Yang, J.-S. Chou, and Y.-W. Wu, “BIM integrated smart monitoring technique for building fire prevention and disaster relief,” *Automation in Construction*, vol. 84, pp. 14–30, Dec. 2017, doi: 10.1016/j.autcon.2017.08.027.

[126] Y. Zhang and L. Bai, “Rapid structural condition assessment using radio frequency identification (RFID) based wireless strain sensor,” *Automation in Construction*, vol. 54, pp. 1–11, Jun. 2015, doi: 10.1016/j.autcon.2015.02.013.

[127] H.-M. Chen, K.-C. Chang, and T.-H. Lin, “A cloud-based system framework for performing online viewing, storage, and analysis on big data of massive BIMs,” *Automation in Construction*, vol. 71, pp. 34–48, Nov. 2016, doi: 10.1016/j.autcon.2016.03.002.

[128] M. Marzouk and A. Abdelaty, “Monitoring thermal comfort in subways using building information modeling,” *Energy and Buildings*, vol. 84, pp. 252–257, Dec. 2014, doi: 10.1016/j.enbuild.2014.08.006.

[129] B. Balaji *et al.*, “Brick : Metadata schema for portable smart building applications,” *Applied Energy*, vol. 226, pp. 1273–1292, Sep. 2018, doi: 10.1016/j.apenergy.2018.02.091.

[130] T. Bloch, “Connecting research on semantic enrichment of BIM - review of approaches, methods and possible applications,” *ITcon*, vol. 27, pp. 416–440, Apr. 2022, doi: 10.36680/j.itcon.2022.020.

[131] B. Zhong, C. Gan, H. Luo, and X. Xing, “Ontology-based framework for building environmental monitoring and compliance checking under BIM environment,” *Building and Environment*, vol. 141, pp. 127–142, Aug. 2018, doi: 10.1016/j.buildenv.2018.05.046.

[132] E. González, J. D. Piñeiro, J. Toledo, R. Arnay, and L. Acosta, “An approach based on the ifcOWL ontology to support indoor navigation,” *Egyptian Informatics Journal*, vol. 22, no. 1, Art. no. 1, Mar. 2021, doi: 10.1016/j.eij.2020.02.008.

[133] E. Corry, P. Pauwels, S. Hu, M. Keane, and J. O'Donnell, "A performance assessment ontology for the environmental and energy management of buildings," *Automation in Construction*, vol. 57, pp. 249–259, Sep. 2015, doi: 10.1016/j.autcon.2015.05.002.
[134] S. Howell, Y. Rezgui, and T. Beach, "Integrating building and urban semantics to empower smart water solutions," *Automation in Construction*, vol. 81, pp. 434–448, Sep. 2017, doi: 10.1016/j.autcon.2017.02.004.
[135] S. Hu, E. Corry, E. Curry, W. J. N. Turner, and J. O'Donnell, "Building performance optimisation: A hybrid architecture for the integration of contextual information and time-series data," *Automation in Construction*, vol. 70, pp. 51–61, Oct. 2016, doi: 10.1016/j.autcon.2016.05.018.
[136] G. Williams, M. Gheisari, P.-J. Chen, and J. Irizarry, "BIM2MAR: An Efficient BIM Translation to Mobile Augmented Reality Applications," *J. Manage. Eng.*, vol. 31, no. 1, p. A4014009, Jan. 2015, doi: 10.1061/(ASCE)ME.1943-5479.0000315.
[137] Universidad del Bío Bío, A. Bonilla Castro, and R. García Alvarado, "BIM-Integration of solar thermal systems in early housing design," *rdlc*, vol. 16, no. 2, pp. 323–338, Aug. 2017, doi: 10.7764/RDLC.16.2.323.
[138] S.-H. Wang, W.-C. Wang, K.-C. Wang, and S.-Y. Shih, "Applying building information modeling to support fire safety management," *Automation in Construction*, vol. 59, pp. 158–167, Nov. 2015, doi: 10.1016/j.autcon.2015.02.001.
[139] E. Patti *et al.*, "Combining Building Information Modelling and Ambient Data in Interactive Virtual and Augmented Reality Environments," *IT Prof.*, pp. 1–1, 2017, doi: 10.1109/MITP.2017.265104553.
[140] R. Attar, S. Breslav, A. Khan, and G. Kurtenbach, "Sensor-enabled Cubicles for Occupant-centric Capture of Building Performance Data," p. 9.
[141] B. Dave, A. Buda, A. Nurminen, and K. Främling, "A framework for integrating BIM and IoT through open standards," *Automation in Construction*, vol. 95, pp. 35–45, Nov. 2018, doi: 10.1016/j.autcon.2018.07.022.
[142] G. Desogus *et al.*, "Preliminary performance monitoring plan for energy retrofit: A cognitive building: The 'Mandolesi Pavillon' at the University of Cagliari," in *2017 AEIT International Annual Conference*, Cagliari: IEEE, Sep. 2017, pp. 1–6. doi: 10.23919/AEIT.2017.8240529.
[143] A. Fonnet, N. Alves, N. Sousa, M. Guevara, and L. Magalhaes, "Heritage BIM integration with mixed reality for building preventive maintenance," in *2017 24º Encontro Português de Computação Gráfica e Interação (EPCGI)*, Guimaraes: IEEE, Oct. 2017, pp. 1–7. doi: 10.1109/EPCGI.2017.8124304.
[144] A. Genty, "Virtual Reality for the construction industry, The CALLISTO-SARI project, benefits for BOUYGUES CONSTRUCTION.," in *VRIC '15 Proceedings of the 2015 Virtual Reality International Conference Article No. 11*,
[145] H. Hamledari, B. McCabe, S. Davari, and A. Shahi, "Automated Schedule and Progress Updating of IFC-Based 4D BIMs," *J. Comput. Civ. Eng.*, vol. 31, no. 4, Art. no. 4, Jul. 2017, doi: 10.1061/(ASCE)CP.1943-5487.0000660.
[146] T.-W. Kang and H.-S. Choi, "BIM perspective definition metadata for interworking facility management data," *Advanced Engineering Informatics*, vol. 29, no. 4, pp. 958–970, Oct. 2015, doi: 10.1016/j.aei.2015.09.004.

[147] Z. Ma, Z. Wei, and X. Zhang, "Semi-automatic and specification-compliant cost estimation for tendering of building projects based on IFC data of design model," *Automation in Construction*, vol. 30, pp. 126–135, Mar. 2013, doi: 10.1016/j.autcon.2012.11.020.

[148] S. Macit İlal and H. M. Günaydın, "Computer representation of building codes for automated compliance checking," *Automation in Construction*, vol. 82, pp. 43–58, Oct. 2017, doi: 10.1016/j.autcon.2017.06.018.

[149] S. Meža, Ž. Turk, and M. Dolenc, "Component based engineering of a mobile BIM-based augmented reality system," *Automation in Construction*, vol. 42, pp. 1–12, Jun. 2014, doi: 10.1016/j.autcon.2014.02.011.

[150] A. Motamedi, A. Hammad, and Y. Asen, "Knowledge-assisted BIM-based visual analytics for failure root cause detection in facilities management," *Automation in Construction*, vol. 43, pp. 73–83, Jul. 2014, doi: 10.1016/j.autcon.2014.03.012.

[151] M. Pouke, J.-P. Virtanen, M. Badri, and T. Ojala, "Comparison of two workflows for Web-based 3D smart home visualizations," in *2018 IEEE International Conference on Future IoT Technologies (Future IoT)*, Eger: IEEE, Jan. 2018, pp. 1–8. doi: 10.1109/FIOT.2018.8325599.

[152] F. Shalabi and Y. Turkan, "IFC BIM-Based Facility Management Approach to Optimize Data Collection for Corrective Maintenance," *Journal of Performance of Constructed Facilities*, vol. 31, no. 1, p. 04016081, Feb. 2017, doi: 10.1061/(ASCE)CF.1943-5509.0000941.

[153] W. Shen, Q. Hao, and Y. Xue, "A loosely coupled system integration approach for decision support in facility management and maintenance," *Automation in Construction*, vol. 25, pp. 41–48, Aug. 2012, doi: 10.1016/j.autcon.2012.04.003.

[154] Z. Shen and R. R. A. Issa, "QUANTITATIVE EVALUATION OF THE BIM-ASSISTED CONSTRUCTION DETAILED COST ESTIMATES," p. 25.

[155] W. Solihin, C. Eastman, Y.-C. Lee, and D.-H. Yang, "A simplified relational database schema for transformation of BIM data into a query-efficient and spatially enabled database," *Automation in Construction*, vol. 84, pp. 367–383, Dec. 2017, doi: 10.1016/j.autcon.2017.10.002.

[156] H. Ufuk Gökçe and K. Umut Gökçe, "Integrated System Platform for Energy Efficient Building Operations," *J. Comput. Civ. Eng.*, vol. 28, no. 6, p. 05014005, Nov. 2014, doi: 10.1061/(ASCE)CP.1943-5487.0000288.

[157] B. Dong, Z. O'Neill, and Z. Li, "A BIM-enabled information infrastructure for building energy Fault Detection and Diagnostics," *Automation in Construction*, vol. 44, pp. 197–211, Aug. 2014, doi: 10.1016/j.autcon.2014.04.007.

[158] J. Park, J. Chen, and Y. K. Cho, "Self-corrective knowledge-based hybrid tracking system using BIM and multimodal sensors," *Advanced Engineering Informatics*, vol. 32, pp. 126–138, Apr. 2017, doi: 10.1016/j.aei.2017.02.001.

[159] T. Gerrish, K. Ruikar, M. Cook, M. Johnson, M. Phillip, and C. Lowry, "BIM application to building energy performance visualisation and management: Challenges and potential," *Energy and Buildings*, vol. 144, pp. 218–228, Jun. 2017, doi: 10.1016/j.enbuild.2017.03.032.

[160] M. Dibley, H. Li, Y. Rezgui, and J. Miles, "An ontology framework for intelligent sensor-based building monitoring," *Automation in Construction*, vol. 28, pp. 1–14, Dec. 2012, doi: 10.1016/j.autcon.2012.05.018.

[161] Z. Riaz, M. Arslan, A. K. Kiani, and S. Azhar, "CoSMoS: A BIM and wireless sensor based integrated solution for worker safety in confined spaces," *Automation in Construction*, vol. 45, pp. 96–106, Sep. 2014, doi: 10.1016/j.autcon.2014.05.010.

[162] C. Z. Li, F. Xue, X. Li, J. Hong, and G. Q. Shen, “An Internet of Things-enabled BIM platform for on-site assembly services in prefabricated construction,” *Automation in Construction*, vol. 89, pp. 146–161, May 2018, doi: 10.1016/j.autcon.2018.01.001.
[163] M. Arslan, Z. Riaz, A. K. Kiani, and S. Azhar, “REAL-TIME ENVIRONMENTAL MONITORING, VISUALIZATION AND NOTIFICATION SYSTEM FOR CONSTRUCTION H&S MANAGEMENT,” p. 20.
[164] J.-H. Woo, M. A. Peterson, and B. Gleason, “Developing a Virtual Campus Model in an Interactive Game-Engine Environment for Building Energy Benchmarking,” *J. Comput. Civ. Eng.*, vol. 30, no. 5, p. C4016005, Sep. 2016, doi: 10.1061/(ASCE)CP.1943-5487.0000600.
[165] S. F. Beck, J. Abualdenien, I. H. Hijazi, A. Borrmann, and T. H. Kolbe, “Analyzing Contextual Linking of Heterogeneous Information Models from the Domains BIM and UIM,” *IJGI*, vol. 10, no. 12, Art. no. 12, Nov. 2021, doi: 10.3390/ijgi10120807.
[166] M. Niknam and S. Karshenas, “A shared ontology approach to semantic representation of BIM data,” *Automation in Construction*, vol. 80, pp. 22–36, Aug. 2017, doi: 10.1016/j.autcon.2017.03.013.
[167] H. Alavi and N. Forcada, “User-Centric BIM-Based Framework for HVAC Root-Cause Detection,” *Energies*, vol. 15, no. 10, p. 3674, May 2022, doi: 10.3390/en15103674.
[168] Y. Cao, S. N. Kamaruzzaman, and N. M. Aziz, “Building Information Modeling (BIM) Capabilities in the Operation and Maintenance Phase of Green Buildings: A Systematic Review,” *Buildings*, vol. 12, no. 6, p. 830, Jun. 2022, doi: 10.3390/buildings12060830.
[169] X. Gao and P. Pishdad-Bozorgi, “BIM-enabled facilities operation and maintenance: A review,” *Advanced Engineering Informatics*, vol. 39, pp. 227–247, Jan. 2019, doi: 10.1016/j.aei.2019.01.005.
[170] J. C. P. Cheng, W. Chen, K. Chen, and Q. Wang, “Data-driven predictive maintenance planning framework for MEP components based on BIM and IoT using machine learning algorithms,” *Automation in Construction*, vol. 112, p. 103087, Apr. 2020, doi: 10.1016/j.autcon.2020.103087.
[171] P. Schönfelder, A. Aziz, B. Faltin, and M. König, “Automating the retrospective generation of As-is BIM models using machine learning,” *Automation in Construction*, vol. 152, p. 104937, Aug. 2023, doi: 10.1016/j.autcon.2023.104937.
[172] T. Bloch and R. Sacks, “Comparing machine learning and rule-based inferencing for semantic enrichment of BIM models,” *Automation in Construction*, vol. 91, pp. 256–272, Jul. 2018, doi: 10.1016/j.autcon.2018.03.018.
[173] G. F. Schneider, Y. Kalantari, G. Kontes, S. Steiger, and D. Rovas, “AN ONTOLOGY-BASED TOOL FOR AUTOMATED CONFIGURATION AND DEPLOYMENT OF TECHNICAL BUILDING MANAGEMENT SERVICES,” p. 9.
[174] I. Ha, H. Kim, S. Park, and H. Kim, “Image retrieval using BIM and features from pretrained VGG network for indoor localization,” *Building and Environment*, vol. 140, pp. 23–31, Aug. 2018, doi: 10.1016/j.buildenv.2018.05.026.
[175] H. Hassanpour, P. Mhaskar, J. M. House, and T. I. Salsbury, “A hybrid modeling approach integrating first-principles knowledge with statistical methods for fault detection in HVAC systems,” *Computers & Chemical Engineering*, vol. 142, p. 107022, Nov. 2020, doi: 10.1016/j.compchemeng.2020.107022.
[176] T. Hong, Z. Wang, X. Luo, and W. Zhang, “State-of-the-art on research and applications of machine learning in the building life cycle,” *Energy and Buildings*, vol. 212, p. 109831, Apr. 2020, doi: 10.1016/j.enbuild.2020.109831.

[177] F. Shalabi and Y. Turkan, “BIM–energy simulation approach for detecting building spaces with faults and problematic behavior,” *ITcon*, vol. 25, pp. 342–360, Jun. 2020, doi: 10.36680/j.itcon.2020.020.
[178] M. Fernandez, A. Gomez-Pearez, and N. Juristo, “Methontology: From Ontological Art Towards Ontological Engineering,” p. 9, 1997.
[179] J. France-Mensah and W. J. O’Brien, “A shared ontology for integrated highway planning,” *Advanced Engineering Informatics*, vol. 41, p. 100929, Aug. 2019, doi: 10.1016/j.aei.2019.100929.
[180] “protégé.” Accessed: Jul. 09, 2023. [Online]. Available: https://protege.stanford.edu/
[181] G. F. Schneider, “Automated Ontology Matching in the Architecture, Engineering and Construction Domain - A Case Study,” p. 15, 2019.
[182] “OOPS! - OntOlogy Pitfall Scanner!” Accessed: Oct. 05, 2023. [Online]. Available: https://oops.linkeddata.es/
[183] J. Oraskari *et al.*, *IFCtoLBD*. (Feb. 2023). Python. Accessed: Jul. 09, 2023. [Online]. Available: https://github.com/jyrkioraskari/IFCtoLBD
[184] C. Schröer, F. Kruse, and J. M. Gómez, “A Systematic Literature Review on Applying CRISP-DM Process Model,” *Procedia Computer Science*, vol. 181, pp. 526–534, 2021, doi: 10.1016/j.procs.2021.01.199.
[185] P. Forner, D. Galbreath, D. Halel, W. Lupson, R. Granderson, and T. Offord, “ANSI/ASHRAE/ACCA Standard 1 80-201 8,” p. 42.
[186] “Flexible Research Platforms | ORNL.” Accessed: Jul. 17, 2023. [Online]. Available: https://www.ornl.gov/content/flexible-research-platforms
[187] “LBNL Automated Fault Detection for Buildings Data.” Accessed: Jul. 09, 2023. [Online]. Available: https://www.kaggle.com/datasets/claytonmiller/lbnl-automated-fault-detection-for-buildings-data
[188] A. Afram, F. Janabi-Sharifi, A. S. Fung, and K. Raahemifar, “Artificial neural network (ANN) based model predictive control (MPC) and optimization of HVAC systems: A state of the art review and case study of a residential HVAC system,” *Energy and Buildings*, vol. 141, pp. 96–113, Apr. 2017, doi: 10.1016/j.enbuild.2017.02.012.
[189] A. Afram and F. Janabi-Sharifi, “Review of modeling methods for HVAC systems,” *Applied Thermal Engineering*, vol. 67, no. 1–2, pp. 507–519, Jun. 2014, doi: 10.1016/j.applthermaleng.2014.03.055.
[190] V. Martinez-Viol, E. M. Urbano, K. Kampouropoulos, M. Delgado-Prieto, and L. Romeral, “Support vector machine based novelty detection and FDD framework applied to building AHU systems,” in *2020 25th IEEE International Conference on Emerging Technologies and Factory Automation (ETFA)*, Vienna, Austria: IEEE, Sep. 2020, pp. 1749–1754. doi: 10.1109/ETFA46521.2020.9212088.
[191] H. Han, X. Cui, Y. Fan, and H. Qing, “Least squares support vector machine (LS-SVM)-based chiller fault diagnosis using fault indicative features,” *Applied Thermal Engineering*, vol. 154, pp. 540–547, May 2019, doi: 10.1016/j.applthermaleng.2019.03.111.
[192] D. Goldwasser, B. Ball, A. Farthing, S. Frank, and P. Im, “Advances in Calibration of Building Energy Models to Time Series Data: Preprint,” *Renewable Energy*, p. 11, 2018.
[193] J. Granderson, G. Lin, A. Harding, P. Im, and Y. Chen, “Building fault detection data to aid diagnostic algorithm creation and performance testing,” *Sci Data*, vol. 7, no. 1, p. 65, Dec. 2020, doi: 10.1038/s41597-020-0398-6.

[194] J. Lee, P. Im, J. D. Munk, M. Malhotra, M. Kim, and Y. Song, “Comparison Evaluations of VRF and RTU Systems Performance on Flexible Research Platform,” *Advances in Civil Engineering*, vol. 2018, pp. 1–16, 2018, doi: 10.1155/2018/7867128.

[195] A. Casillas, G. Lin, and J. Granderson, “Curation of Ground-Truth Validated Benchmarking Datasets for Fault Detection & Diagnostics Tools,” 2020, doi: 10.20357/B7NG6Z.

[196] P. Im, J. R. New, and J. Joe, “Empirical Validation of Building Energy Modeling using Flexible Research Platform,” presented at the Building Simulation 2019, Rome, Italy, Jun. 2022, pp. 4515–4521. doi: 10.26868/25222708.2019.210263.

[197] P. Im, J. D. Munk, and A. C. Gehl, “Evaluation of Variable Refrigerant Flow Systems Performance and the Enhanced Control Algorithm on Oak Ridge National Laboratory s Flexible Research Platform,” ORNL/TM--2015/225, 1186004, Jun. 2015. doi: 10.2172/1186004.

[198] A. Hosseini Gourabpasi and M. Nik-Bakht, “BIM-based automated fault detection and diagnostics of HVAC systems in commercial buildings,” *Journal of Building Engineering*, p. 109022, Mar. 2024, doi: 10.1016/j.jobe.2024.109022.

[199] “Level of Development Specification,” 2015, [Online]. Available: https://bim-international.com/wp-content/uploads/2016/03/LOD-Specification-2015.pdf

[200] “IFC Viewer | usBIM | ACCA software.” Accessed: Jul. 09, 2023. [Online]. Available: https://www.accasoftware.com/en/ifc-viewer

[201] “Anaconda | The World’s Most Popular Data Science Platform,” Anaconda. Accessed: Jul. 20, 2023. [Online]. Available: https://www.anaconda.com/

[202] “Roof top unit (RTU),” Fault Detection and Diagnostics. Accessed: Feb. 17, 2024. [Online]. Available: https://faultdetection.lbl.gov/dataset/rtu/

[203] “Blazegraph Database.” Accessed: Jul. 24, 2023. [Online]. Available: https://blazegraph.com/

[204] B. Ben Mahria, I. Chaker, and A. Zahi, “A novel approach for learning ontology from relational database: from the construction to the evaluation,” *J Big Data*, vol. 8, no. 1, p. 25, Dec. 2021, doi: 10.1186/s40537-021-00412-2.

[205] “WebVOWL - Web-based Visualization of Ontologies.” Accessed: Jul. 20, 2023. [Online]. Available: http://vowl.visualdataweb.org/webvowl.html

[206] J. Ikonen *et al.*, “USE OF EMBEDDED RFID TAGS IN CONCRETE ELEMENT SUPPLY CHAINS,” p. 29.

[207] H. Li, M. Lu, G. Chan, and M. Skitmore, “Proactive training system for safe and efficient precast installation,” *Automation in Construction*, vol. 49, pp. 163–174, Jan. 2015, doi: 10.1016/j.autcon.2014.10.010.

[208] R. Y. Zhong *et al.*, “Prefabricated construction enabled by the Internet-of-Things,” *Automation in Construction*, vol. 76, pp. 59–70, Apr. 2017, doi: 10.1016/j.autcon.2017.01.006.

[209] F. Petrushevski, “Personalized lighting control based on a space model,” in *Proceedings of the 2012 ACM Conference on Ubiquitous Computing - UbiComp ’12*, Pittsburgh, Pennsylvania: ACM Press, 2012, p. 568. doi: 10.1145/2370216.2370311.

[210] L. Oppermann, M. Shekow, and D. Bicer, “Mobile cross-media visualisations made from building information modelling data,” in *Proceedings of the 18th International Conference on Human-Computer Interaction with Mobile Devices and Services Adjunct - MobileHCI ’16*, Florence, Italy: ACM Press, 2016, pp. 823–830. doi: 10.1145/2957265.2961852.

[211] J. Teizer *et al.*, "Internet of Things (IoT) for Integrating Environmental and Localization Data in Building Information Modeling (BIM)," presented at the 34th International Symposium on Automation and Robotics in Construction, Taipei, Taiwan, Jul. 2017. doi: 10.22260/ISARC2017/0084.
[212] K. Kim, Y. Cho, and S. Zhang, "Integrating work sequences and temporary structures into safety planning: Automated scaffolding-related safety hazard identification and prevention in BIM," *Automation in Construction*, vol. 70, pp. 128–142, Oct. 2016, doi: 10.1016/j.autcon.2016.06.012.
[213] X. Yuan, C. J. Anumba, and M. K. Parfitt, "Cyber-physical systems for temporary structure monitoring," *Automation in Construction*, vol. 66, pp. 1–14, Jun. 2016, doi: 10.1016/j.autcon.2016.02.005.
[214] E. A. Pärn and D. J. Edwards, "Conceptualising the FinDD API plug-in: A study of BIM-FM integration," *Automation in Construction*, vol. 80, pp. 11–21, Aug. 2017, doi: 10.1016/j.autcon.2017.03.015.
[215] G. Xu, M. Li, C.-H. Chen, and Y. Wei, "Cloud asset-enabled integrated IoT platform for lean prefabricated construction," *Automation in Construction*, vol. 93, pp. 123–134, Sep. 2018, doi: 10.1016/j.autcon.2018.05.012.
[216] Y. Niu, W. Lu, D. Liu, K. Chen, C. Anumba, and G. G. Huang, "An SCO-Enabled Logistics and Supply Chain–Management System in Construction," *J. Constr. Eng. Manage.*, vol. 143, no. 3, p. 04016103, Mar. 2017, doi: 10.1061/(ASCE)CO.1943-7862.0001232.
[217] C. Z. Li, F. Xue, X. Li, J. Hong, and G. Q. Shen, "An Internet of Things-enabled BIM platform for on-site assembly services in prefabricated construction," *Automation in Construction*, vol. 89, pp. 146–161, May 2018, doi: 10.1016/j.autcon.2018.01.001.
[218] B. Choi, H.-S. Lee, M. Park, Y. K. Cho, and H. Kim, "Framework for Work-Space Planning Using Four-Dimensional BIM in Construction Projects," *J. Constr. Eng. Manage.*, vol. 140, no. 9, p. 04014041, Sep. 2014, doi: 10.1061/(ASCE)CO.1943-7862.0000885.
[219] A. J.-P. Tixier, M. R. Hallowell, B. Rajagopalan, and D. Bowman, "Construction Safety Clash Detection: Identifying Safety Incompatibilities among Fundamental Attributes using Data Mining," *Automation in Construction*, vol. 74, pp. 39–54, Feb. 2017, doi: 10.1016/j.autcon.2016.11.001.
[220] X.-S. Chen, C.-C. Liu, and I.-C. Wu, "A BIM-based visualization and warning system for fire rescue," *Advanced Engineering Informatics*, vol. 37, pp. 42–53, Aug. 2018, doi: 10.1016/j.aei.2018.04.015.
[221] T. Gao, B. Akinci, S. Ergan, and J. Garrett, "An approach to combine progressively captured point clouds for BIM update," *Advanced Engineering Informatics*, vol. 29, no. 4, pp. 1001–1012, Oct. 2015, doi: 10.1016/j.aei.2015.08.005.
[222] F. Troncoso-Pastoriza, J. López-Gómez, and L. Febrero-Garrido, "Generalized Vision-Based Detection, Identification and Pose Estimation of Lamps for BIM Integration," *Sensors*, vol. 18, no. 7, p. 2364, Jul. 2018, doi: 10.3390/s18072364.
[223] F. Troncoso-Pastoriza, P. Eguía-Oller, R. P. Díaz-Redondo, and E. Granada-Álvarez, "Generation of BIM data based on the automatic detection, identification and localization of lamps in buildings," *Sustainable Cities and Society*, vol. 36, pp. 59–70, Jan. 2018, doi: 10.1016/j.scs.2017.10.015.
[224] D. Rebolj, Z. Pučko, N. Č. Babič, M. Bizjak, and D. Mongus, "Point cloud quality requirements for Scan-vs-BIM based automated construction progress monitoring,"

Automation in Construction, vol. 84, pp. 323–334, Dec. 2017, doi: 10.1016/j.autcon.2017.09.021.

[225] K. M. Rashid, J. Louis, and K. K. Fiawoyife, “Wireless electric appliance control for smart buildings using indoor location tracking and BIM-based virtual environments,” *Automation in Construction*, vol. 101, pp. 48–58, May 2019, doi: 10.1016/j.autcon.2019.01.005.

[226] G. N. Sava, S. Pluteanu, V. Tanasiev, R. Patrascu, and H. Necula, “Integration of BIM Solutions and IoT in Smart Houses,” in *2018 IEEE International Conference on Environment and Electrical Engineering and 2018 IEEE Industrial and Commercial Power Systems Europe (EEEIC / I&CPS Europe)*, Palermo: IEEE, Jun. 2018, pp. 1–4. doi: 10.1109/EEEIC.2018.8494628.

[227] C. Z. Li *et al.*, “Integrating RFID and BIM technologies for mitigating risks and improving schedule performance of prefabricated house construction,” *Journal of Cleaner Production*, vol. 165, pp. 1048–1062, Nov. 2017, doi: 10.1016/j.jclepro.2017.07.156.

[228] G. Lee *et al.*, “A BIM- and sensor-based tower crane navigation system for blind lifts,” *Automation in Construction*, vol. 26, pp. 1–10, Oct. 2012, doi: 10.1016/j.autcon.2012.05.002.

[229] M. Alahmad *et al.*, “THE ‘BIM’S 4D+’ DIMENSION: REAL TIME ENERGY MONITORING,” in *2011 IEEE GCC Conference and Exhibition (GCC)*, Dubai, United Arab Emirates: IEEE, Feb. 2011, pp. 589–592. doi: 10.1109/IEEEGCC.2011.5752603.

[230] J. Matthews, P. E. D. Love, S. Heinemann, R. Chandler, C. Rumsey, and O. Olatunj, “Real time progress management: Re-engineering processes for cloud-based BIM in construction,” *Automation in Construction*, vol. 58, pp. 38–47, Oct. 2015, doi: 10.1016/j.autcon.2015.07.004.

[231] R. Bortolini, C. T. Formoso, and D. D. Viana, “Site logistics planning and control for engineer-to-order prefabricated building systems using BIM 4D modeling,” *Automation in Construction*, vol. 98, pp. 248–264, Feb. 2019, doi: 10.1016/j.autcon.2018.11.031.

[232] Z. Alwan, D. Greenwood, and B. Gledson, “Rapid LEED evaluation performed with BIM based sustainability analysis on a virtual construction project,” *Construction Innovation*, vol. 15, no. 2, pp. 134–150, Apr. 2015, doi: 10.1108/CI-01-2014-0002.

[233] P. Coates, Y. Arayici, and Z. Ozturk, “New Concepts of Post Occupancy Evaluation (POE) Utilizing BIM Benchmarking Techniques and Sensing Devices,” in *Sustainability in Energy and Buildings*, vol. 12, N. M’Sirdi, A. Namaane, R. J. Howlett, and L. C. Jain, Eds., Berlin, Heidelberg: Springer Berlin Heidelberg, 2012, pp. 319–329. doi: 10.1007/978-3-642-27509-8_27.

[234] I. Ha, H. Kim, S. Park, and H. Kim, “Image retrieval using BIM and features from pretrained VGG network for indoor localization,” *Building and Environment*, vol. 140, pp. 23–31, Aug. 2018, doi: 10.1016/j.buildenv.2018.05.026.

[235] D. H. Shin and P. S. Dunston, “Evaluation of Augmented Reality in steel column inspection,” *Automation in Construction*, vol. 18, no. 2, pp. 118–129, Mar. 2009, doi: 10.1016/j.autcon.2008.05.007.

[236] M. Bassier, B. Van Genechten, and M. Vergauwen, “Classification of sensor independent point cloud data of building objects using random forests,” *Journal of Building Engineering*, vol. 21, pp. 468–477, Jan. 2019, doi: 10.1016/j.jobe.2018.04.027.

[237] M. Arslan, Z. Riaz, and S. Munawar, “Building Information Modeling (BIM) Enabled Facilities Management Using Hadoop Architecture,” in *2017 Portland International Conference on Management of Engineering and Technology (PICMET)*, Portland, OR: IEEE, Jul. 2017, pp. 1–7. doi: 10.23919/PICMET.2017.8125462.

[238] S. Zhang, J. Teizer, J.-K. Lee, C. M. Eastman, and M. Venugopal, “Building Information Modeling (BIM) and Safety: Automatic Safety Checking of Construction Models and Schedules,” *Automation in Construction*, vol. 29, pp. 183–195, Jan. 2013, doi: 10.1016/j.autcon.2012.05.006.

[239] L. Bottaccioli *et al.*, “Building Energy Modelling and Monitoring by Integration of IoT Devices and Building Information Models,” IEEE, Jul. 2017, pp. 914–922. doi: 10.1109/COMPSAC.2017.75.

[240] J. Lee, N. Kwon, N. Ham, J. Kim, and Y. Ahn, “BIM-Based Digital Fabrication Process for a Free-Form Building Project in South Korea,” *Advances in Civil Engineering*, vol. 2019, pp. 1–18, May 2019, doi: 10.1155/2019/4163625.

[241] J. Lee, Y.-J. Park, C.-H. Choi, and C.-H. Han, “BIM-assisted labor productivity measurement method for structural formwork,” *Automation in Construction*, vol. 84, pp. 121–132, Dec. 2017, doi: 10.1016/j.autcon.2017.08.009.

[242] Z. Pučko, N. Šuman, and D. Rebolj, “Automated continuous construction progress monitoring using multiple workplace real time 3D scans,” *Advanced Engineering Informatics*, vol. 38, pp. 27–40, Oct. 2018, doi: 10.1016/j.aei.2018.06.001.

[243] A. Mirzaei, F. Nasirzadeh, M. Parchami Jalal, and Y. Zamani, “4D-BIM Dynamic Time–Space Conflict Detection and Quantification System for Building Construction Projects,” *Journal of Construction Engineering and Management*, vol. 144, no. 7, p. 04018056, Jul. 2018, doi: 10.1061/(ASCE)CO.1943-7862.0001504.

[244] S. N. Razavi and C. T. Hass, “A Data Fusion Model for Location Estimation in Construction,” presented at the 26th International Symposium on Automation and Robotics in Construction, Austin, TX, USA, Jun. 2009. doi: 10.22260/ISARC2009/0004.

[245] Y. Peng, J.-R. Lin, J.-P. Zhang, and Z.-Z. Hu, “A hybrid data mining approach on BIM-based building operation and maintenance,” *Building and Environment*, vol. 126, pp. 483–495, Dec. 2017, doi: 10.1016/j.buildenv.2017.09.030.

[246] F. H. Abanda and L. Byers, “An investigation of the impact of building orientation on energy consumption in a domestic building using emerging BIM (Building Information Modelling),” *Energy*, vol. 97, pp. 517–527, Feb. 2016, doi: 10.1016/j.energy.2015.12.135.

[247] F. Baek, I. Ha, and H. Kim, “Augmented reality system for facility management using image-based indoor localization,” *Automation in Construction*, vol. 99, pp. 18–26, Mar. 2019, doi: 10.1016/j.autcon.2018.11.034.

[248] V. Getuli, S. M. Ventura, P. Capone, and A. L. C. Ciribini, “BIM-based Code Checking for Construction Health and Safety,” *Procedia Engineering*, vol. 196, pp. 454–461, 2017, doi: 10.1016/j.proeng.2017.07.224.

[249] S. Habibi, “Micro-climatization and real-time digitalization effects on energy efficiency based on user behavior,” *Building and Environment*, vol. 114, pp. 410–428, Mar. 2017, doi: 10.1016/j.buildenv.2016.12.039.

[250] P. R. Zekavat, S. Moon, and L. E. Bernold, “Performance of short and long range wireless communication technologies in construction,” *Automation in Construction*, vol. 47, pp. 50–61, Nov. 2014, doi: 10.1016/j.autcon.2014.07.008.

[251] M. Heidari, E. Allameh, B. de Vries, H. Timmermans, J. Jessurun, and F. Mozaffar, “Smart-BIM virtual prototype implementation,” *Automation in Construction*, vol. 39, pp. 134–144, Apr. 2014, doi: 10.1016/j.autcon.2013.07.004.

Appendices

Appendix A: List of the 82 Studies Considered for Knowledge Discovery Through Association Rule Mining

No.	Author(s)	Title	Year
1.	K. Yan, J. Huang, W. Shen, and Z. Ji	Unsupervised learning for fault detection and diagnosis of air handling units	2020
2.	K. Yan, A. Chong, and Y. Mo	Generative adversarial network for fault detection diagnosis of chillers	2020
3.	A. Ranade, G. Provan, A. El-Din Mady, and D. O'Sullivan	A computationally efficient method for fault diagnosis of fan-coil unit terminals in building Heating Ventilation and Air Conditioning systems	2020
4.	S. Miyata, J. Lim, Y. Akashi, Y. Kuwahara, and K. Tanaka	Fault detection and diagnosis for heat source system using convolutional neural network with imaged faulty behavior data	2020
5.	Z. Zhang, H. Han, X. Cui, and Y. Fan,	Novel application of multi-model ensemble learning for fault diagnosis in refrigeration systems	2020
6.	Y. Fan, X. Cui, H. Han, and H. Lu	Chiller fault detection and diagnosis by knowledge transfer based on adaptive imbalanced processing	2020
7.	A. Montazeri and S. M. Kargar,	Fault detection and diagnosis in air handling using data-driven methods	2020
8.	J. Liu et al.	Data-driven and association rule mining-based fault diagnosis and action mechanism analysis for building chillers	2020
9.	M. Elnour, N. Meskin, and M. Al-Naemi	Sensor data validation and fault diagnosis using Auto-Associative Neural Network for HVAC systems	2020
10.	Z. Li et al.	Machine learning based diagnosis strategy for refrigerant charge amount malfunction of variable refrigerant flow system	2020
11.	Y. Fan, X. Cui, H. Han, and H. Lu	Feasibility and improvement of fault detection and diagnosis based on factory-installed sensors for chillers	2020

12.	K. Yan, Z. Ji, H. Lu, J. Huang, W. Shen, and Y. Xue	Fast and Accurate Classification of Time Series Data Using Extended ELM: Application in Fault Diagnosis of Air Handling Units	2019
13.	A. Motomura et al.	Fault evaluation process in HVAC system for decision making of how to respond to system faults	2019
14.	Z. Li et al.	An efficient online wkNN diagnostic strategy for variable refrigerant flow system based on coupled feature selection method	2019
15.	G. Li and Y. Hu	An enhanced PCA-based chiller sensor fault detection method using ensemble empirical mode decomposition based denoising	2019
16.	D. Li, D. Li, C. Li, L. Li, and L. Gao	A novel data-temporal attention network based strategy for fault diagnosis of chiller sensors	2019
17.	D. Li, Y. Zhou, G. Hu, and C. J. Spanos	Handling Incomplete Sensor Measurements in Fault Detection and Diagnosis for Building HVAC Systems	2019
18.	H. Han, X. Cui, Y. Fan, and H. Qing	Least squares support vector machine (LS-SVM)-based chiller fault diagnosis using fault indicative features	2019
19.	D. Bigaud, A. Charki, A. Caucheteux, F. Titikpina, and T. Tiplica	Detection of Faults and Drifts in the Energy Performance of a Building Using Bayesian Networks	2019
20.	A. Beghi, R. Brignoli, L. Cecchinato, G. Menegazzo, and M. Rampazzo	A data-driven approach for fault diagnosis in HVAC chiller systems	2019
21.	J. Liu, M. Zhang, H. Wang, W. Zhao, and Y. Liu	Sensor Fault Detection and Diagnosis Method for AHU Using 1-D CNN and Clustering Analysis	2019
22.	C. Zhong, K. Yan, Y. Dai, N. Jin, and B. Lou	Energy Efficiency Solutions for Buildings: Automated Fault Diagnosis of Air Handling Units Using Generative Adversarial Networks	2019

23.	C. Yang, W. Shen, B. Gunay, and Z. Shi	Toward Machine Learning-based Prognostics for Heating Ventilation and Air-Conditioning Systems,	2019
24.	L. Gao, D. Li, D. Li, L. Yao, L. Liang, and Y. Gao	A Novel Chiller Sensors Fault Diagnosis Method Based on Virtual Sensors	2019
25.	M. Tahmasebi, K. Eaton, N. Nassif, and R. Talib	Integrated Machine Learning Modeling and Fault Detection Approach for Chilled Water Systems	2019
26.	J. Liu, G. Li, B. Liu, K. Li, and H. Chen	Knowledge discovery of data-driven-based fault diagnostics for building energy systems: A case study of the building variable refrigerant flow system	2019
27.	A. Behravan, M. Abboush, and R. Obermaisser	Deep Learning Application in Mechatronics Systems' Fault Diagnosis, a Case Study of the Demand-Controlled Ventilation and Heating System	2019
28.	H. Zhang, H. Chen, Y. Guo, J. Wang, G. Li, and L. Shen	Sensor fault detection and diagnosis for a water source heat pump air-conditioning system based on PCA and preprocessed by combined clustering	2019
29.	M. Elnour, N. Meskin, and M. Al-Naemi	Sensor Fault Diagnosis of Multi-Zone HVAC Systems Using Auto-Associative Neural Network	2019
30.	Y. Fan, X. Cui, H. Han, and H. Lu	Chiller fault diagnosis with field sensors using the technology of imbalanced data	2019
31.	B. Jin, D. Li, S. Srinivasan, S.-K. Ng, K. Poolla, and A. Sangiovanni-Vincentelli	Detecting and Diagnosing Incipient Building Faults Using Uncertainty Information from Deep Neural Networks	2019
32.	K. Yan and J. Hua	Deep Learning Technology for Chiller Faults Diagnosis	2019
33.	X. J. Luo, K. F. Fong, Y. J. Sun, and M. K. H. Leung	Development of clustering-based sensor fault detection and diagnosis strategy for chilled water system	2019
34.	Y. H. Eom, J. W. Yoo, S. B. Hong, and M. S. Kim	Refrigerant charge fault detection method of air source heat pump system using convolutional neural network for energy saving	2019
35.	K. Yan, C. Zhong, Z. Ji, and J. Huang	Semi-supervised learning for early detection and diagnosis of various air handling unit faults	2018

36.	K. Yan, L. Ma, Y. Dai, W. Shen, Z. Ji, and D. Xie	Cost-sensitive and sequential feature selection for chiller fault detection and diagnosis	2018
37.	Z. Wang, Z. Wang, X. Gu, S. He, and Z. Yan	Feature selection based on Bayesian network for chiller fault diagnosis from the perspective of field applications	2018
38.	C. G. Mattera, J. Quevedo, T. Escobet, H. R. Shaker, and M. Jradi	Fault Detection and Diagnostics in Ventilation Units Using Linear Regression Virtual Sensors	2018
39.	M. Hu et al.	A machine learning bayesian network for refrigerant charge faults of variable refrigerant flow air conditioning system	2018
40.	Y. Guo et al.	Deep learning-based fault diagnosis of variable refrigerant flow air-conditioning system for building energy saving	2018
41.	M. Dey, S. P. Rana, and S. Dudley	Smart building creation in large scale HVAC environments through automated fault detection and diagnosis	2018
42.	M. Dey, S. P. Rana, and S. Dudley	Semi-Supervised Learning Techniques for Automated Fault Detection and Diagnosis of HVAC Systems	2018
43.	F. Simmini, M. Rampazzo, A. Beghi, and F. Peterle	Local Principal Component Analysis for Fault Detection in Air-Condensed Water Chillers	2018
44.	Y. Chen and J. Wen	Development and Field Evaluation of Data-driven Whole Building Fault Detection and Diagnosis Strategy	2018
45.	K. Yan, C. Zhong, Z. Ji, and J. Huang	Evaluating Semi-supervised Learning for Automated Fault Detection and Diagnosis of Air Handling Units	2018
46.	Y. Chen, J. Wen, T. Chen, and O. Pradhan	Bayesian Networks for Whole Building Level Fault Diagnosis and Isolation	2018
47.	G. Li et al.	An improved decision tree-based fault diagnosis method for practical variable refrigerant flow system using virtual sensor-based fault indicators	2018
48.	X. Liu, Y. Li, X. Liu, and J. Shen	Fault diagnosis of chillers using very deep convolutional network	2018

49.	R. Huang et al.	An effective fault diagnosis method for centrifugal chillers using associative classification	2018
50.	Z. Wang, L. Wang, K. Liang, and Y. Tan,	Enhanced chiller fault detection using Bayesian network and principal component analysis	2018
51.	J. Liu, G. Li, H. Chen, J. Wang, Y. Guo, and J. Li	A robust online refrigerant charge fault diagnosis strategy for VRF systems based on virtual sensor technique and PCA-EWMA method	2017
52.	K. Yan, Z. Ji, and W. Shen	Online fault detection methods for chillers combining extended kalman filter and recursive one-class SVM	2017
53.	K. Verbert, R. Babuška, and B. De Schutter	Combining knowledge and historical data for system-level fault diagnosis of HVAC systems	2017
54.	P. M. Van Every, M. Rodriguez, C. B. Jones, A. A. Mammoli, and M. Martínez-Ramón	Advanced detection of HVAC faults using unsupervised SVM novelty detection and Gaussian process models	2017
55.	W. J. N. Turner, A. Staino, and B. Basu	Residential HVAC fault detection using a system identification approach	2017
56.	S. Sun, G. Li, H. Chen, Q. Huang, S. Shi, and W. Hu	A hybrid ICA-BPNN-based FDD strategy for refrigerant charge faults in variable refrigerant flow system	2017
57.	S. Shi et al.	Refrigerant charge fault diagnosis in the VRF system using Bayesian artificial neural network combined with Relief Filter	2017
58.	S. C. Mukhopadhyay, O. A. Postolache, K. P. Jayasundera, and A. K. Swain, Eds.	Sensors for everyday life: environmental and food engineering	2017
59.	K. Mittal, J. P. Wilson, B. P. Baillie, S. Gupta, G. M. Bollas, and P. B. Luh	Supervisory Control for Resilient Chiller Plants Under Condenser Fouling	2017

60.	Y. Guo et al.	Modularized PCA method combined with expert-based multivariate decoupling for FDD in VRF systems including indoor unit faults	2017
61.	Y. Guo et al.	An enhanced PCA method with Savitzky-Golay method for VRF system sensor fault detection and diagnosis	2017
62.	Y. Chen and J. Wen	A whole building fault detection using weather based pattern matching and feature based PCA method	2017
63.	L. Chang, H. Wang, and L. Wang	Cloud-Based parallel implementation of an intelligent classification algorithm for fault detection and diagnosis of HVAC systems	2017
64.	Z. Wang, Z. Wang, S. He, X. Gu, and Z. F. Yan	Fault detection and diagnosis of chillers using Bayesian network merged distance rejection and multi-source non-sensor information	2017
65.	Y. Chen and J. Wen	Whole building system fault detection based on weather pattern matching and PCA method	2017
66.	J. Wang et al.	Liquid flood back detection for scroll compressor in a VRF system under heating mode	2017
67.	S. Shi et al.	An efficient VRF system fault diagnosis strategy for refrigerant charge amount based on PCA and dual neural network model	2017
68.	R. Yan, Z. Ma, Y. Zhao, and G. Kokogiannakis	A decision tree based data-driven diagnostic strategy for air handling units	2016
69.	K. Sun, G. Li, H. Chen, J. Liu, J. Li, and W. Hu	A novel efficient SVM-based fault diagnosis method for multi-split air conditioning system's refrigerant charge fault amount	2016
70.	J. Liu, Y. Hu, H. Chen, J. Wang, G. Li, and W. Hu	A refrigerant charge fault detection method for variable refrigerant flow (VRF) air-conditioning systems	2016
71.	J. Liu, H. Chen, J. Wang, G. Li, H. Li, and W. Hu	Fault diagnosis of refrigerant charge based on PCA and decision tree for variable refrigerant flow systems	2016

72.	G. Li et al.	An improved fault detection method for incipient centrifugal chiller faults using the PCA-R-SVDD algorithm	2016
73.	G. Li et al.	A sensor fault detection and diagnosis strategy for screw chiller system using support vector data description-based D-statistic and DV-contribution plots	2016
74.	D. Li, G. Hu, and C. J. Spanos	A data-driven strategy for detection and diagnosis of building chiller faults using linear discriminant analysis	2016
75.	Y. Hu, G. Li, H. Chen, H. Li, and J. Liu	Sensitivity analysis for PCA-based chiller sensor fault detection	2016
76.	S. He, Z. Wang, Z. Wang, X. Gu, and Z. Yan	Fault detection and diagnosis of chiller using Bayesian network classifier with probabilistic boundary	2016
77.	Y. Gao, S. Liu, F. Li, and Z. Liu	Fault detection and diagnosis method for cooling dehumidifier based on LS-SVM NARX model,	2016
78.	A. Beghi, R. Brignoli, L. Cecchinato, G. Menegazzo, M. Rampazzo, and F. Simmini	Data-driven Fault Detection and Diagnosis for HVAC water chillers	2016
79.	R. Yan, Z. Ma, G. Kokogiannakis, and Y. Zhao	A sensor fault detection strategy for air handling units using cluster analysis	2016
80.	D. A. T. Tran, Y. Chen, H. L. Ao, and H. N. T. Cam	An enhanced chiller FDD strategy based on the combination of the LSSVR-DE model and EWMA control charts	2016
81.	D. A. T. Tran, Y. Chen, and C. Jiang	Comparative investigations on reference models for fault detection and diagnosis in centrifugal chiller systems	2016
82.	C. Audivet Durán and M. E. Sanjuán	On-Line Early Fault Detection of a Centrifugal Chiller Based on Data Driven Approach	2016

Appendix B: Data-Driven FDD Algorithms Based on Machine Learning Approach

The AFDD techniques reviewed and used in the thesis are broadly grouped and categorized into supervised and unsupervised learning. This study also covers more general algorithms, such as Bayesian network (BN) and ARM algorithms, which may not traditionally fit in any of these two broad categories. Most of the reviewed studies implementing AFDD are supervised methods and treat the FDD as essentially a classification problem. Unsupervised methods are mainly adopted in the pre-processing phase or are used for fault detection through clustering.

Figure 9.1, shows the machine learning algorithms for FDD based on learning type. SVM (support vector machine), decision tree, and regression methods are grouped into supervised, and dimensionality reduction techniques, instance-based classification and clustering belong to the unsupervised category. However, ANN/deep learning, ensemble learning, Bayesian networks, and ARM in the literature have used both supervised and unsupervised methods. Bayesian methods are used where event information is required to be included in the models. The events describe the states of discrete or continuous variables, such as a room being occupied or not by its occupants or considering the HVAC operation schedule, respectively. In hybrid methods, the machine learning approach for fault detection and diagnostics are different from one another. For detailed definitions of the algorithms, readers can refer to the author's published journal paper [52].

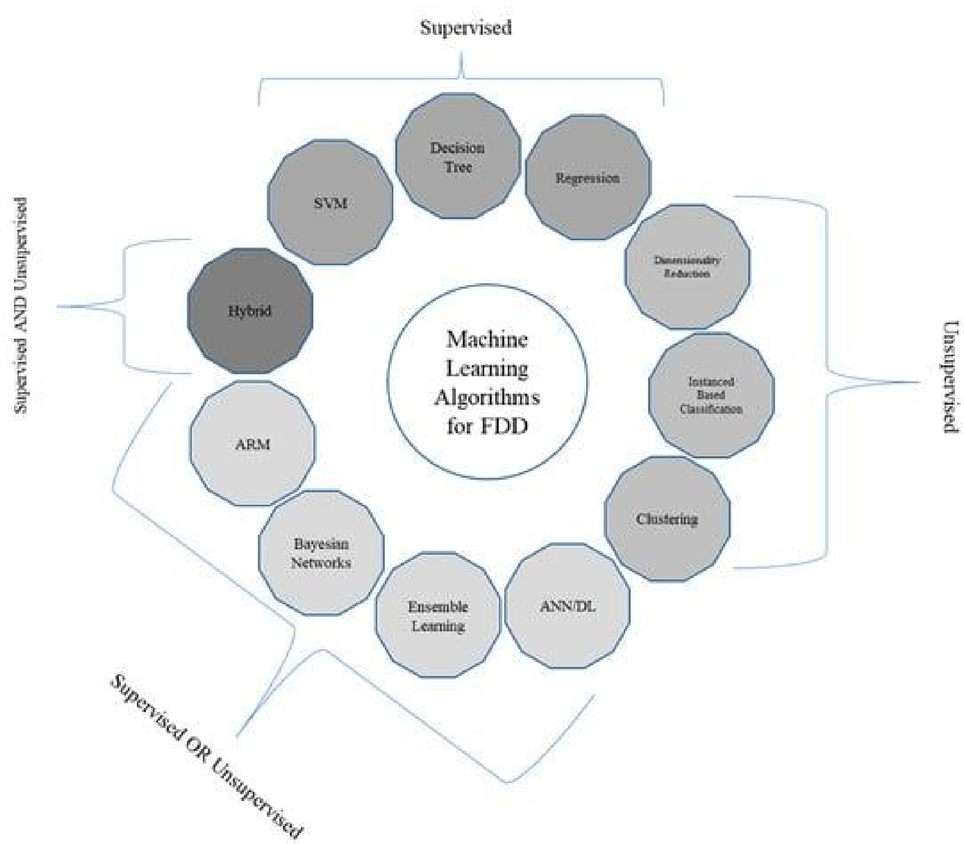


Figure 9.1: Machine learning FDD algorithms based on learning type

Appendix C: List of Potential BIM Uses That Can Benefit from Dynamism

SI No.	BIM use	Title	Application	Variable (Data)
1	Logistic Planning	[206]	Tracking and information management	Location
2	Building System Analysis	[113]	Energy monitoring and control interface	Co2, Motion detector, Temperature, Humidity, Energy consumption
3	Logistic Planning / Space Management and Tracking	[158]	Tracking in dynamic and complex indoor construction sites/ tracking system using BIM and multimodal sensors	Location, Motion
4	Safety	[163]	Improved Health and Safety	Temperature, Humidity
5	Building (Preventative) Maintenance Scheduling/safety	[126]	Structural condition assessment	Contact scanning, Strain
6	Safety	[207]	Construction safety	Location
7	Asset Management	[142]	Performance Monitoring	Temperature, RH, Illuminance, Windows opening, User presence(occupancy), Electrical consumption
8	Logistic Planning	[208]	Visibility and traceability	Location
9	Asset Management	[209]	Personalized lighting control	Quality of service: (system response time + lighting performance), Lighting
10	Building System Analysis	[128]	Monitoring thermal comfort	Temperature, Humidity
11	3D Control and Planning (Digital layout)	[210]	Mobile Cross-Media Visualizations	Image
12	Building (Preventative) Maintenance Scheduling	[150]	Facilities management	(Conditions of the mechanical assets and their visualization codes), Status
13	Space Management and Tracking/ Asset Management/Safety	[211]	Continuous progress monitoring function and workspace safety	Location, Temperature, Humidity, Illumination
14	Phase Planning (4D Modeling)/ Safety	[212]	Safety hazard identification and prevention	Schedule
15	Building System Analysis	[156]	Energy Efficient Building Operations	Temperature, Humidity, CO2, Lux level
16	Engineering Analysis	[164]	Building Energy Benchmarking	Temperature, Humidity, Co2, Energy consumption
17	Engineering Analysis /safety	[213]	Temporary structure monitoring	On/off, Weight, Motion, Location
18	Safety	[161]	Worker safety in confined spaces	Temperature, Oxygen

19	Asset Management	[214]	BIM-FM integration	Barcode, (Location) ,COBie data
20	Phase Planning (4D Modeling)/ Design Reviews	[149]	4D modeling	Monitoring of construction process, Schedule
21	Asset Management	[151]	Smart Home Visualizations	(Case 1) Temperature Humidity, Air quality and Pressure difference (Case 2) Temperature
22	Asset Management/ Construction System Design (Virtual Mockup)	[139]	Interactive AR and VR environment	Building condition, Energy consumption
23	Logistic Planning	[215]	Construction asset tracking	Location
24	Asset Management	[136]	Facility management	Object and model link, Location, Motion, Rotation
25	Disaster Planning	[125]	Fire Prevention and Disaster Relief System	Temperature, Smoke
26	Building System Analysis	[159]	Building energy performance visualization	Temperature, CO2, Humidity, Energy, Lighting
27	Logistic Planning	[216]	Logistics and Supply Chain–Management System	Location
28	Space Management and Tracking	[160]	Sensor-based building monitoring	Person count, Environmental monitor, Temp, Lux, Monitoring
29	Logistic Planning	[217]	On-site assembly services in prefabricated construction	Location
30	Construction System Design (Virtual Mockup) /Design Reviews	[144]	Virtual reality for the construction industry	(Gesture User language), Motion, Sound, Point cloud, Video
31	Asset Management	[155]	Query-efficient and spatially enabled database	Objects (Count)
32	Building (Preventative) Maintenance Scheduling / Disaster Planning	[153]	Decision support in facility management and maintenance	(Asset tracking) location
33	Building System Analysis / Asset Management	[141]	Integrates the built environment data with IoT sensors in a campus space/utilization, user comfort, energy usage monitoring and energy-saving and a reduced carbon footprint	CO2, Humidity, Temperature And other sensors, Campus booking system, (occupancy), Heating and ventilation
34	Phase Planning (4D Modeling)	[218]	Status of work-space occupation and to identify workspace–related problems,	(Construction method DB and material information DB), Schedule, Occupancy, Material information
35	3D coordination/Safety	[219]	Safety clashes	Raw textual injury reports
36	Disaster Planning	[220]	Visualization and warning system for fire rescue	Temp, CO, Visibility

37	Existing Conditions Modeling	[221]	As built	Point cloud
38	Building system analysis/ Code Validation/Safety	[131]	(Building environmental monitoring and compliance checking)	Temp, Humidity, Light, Air quality
39	Disaster Planning	[138]	Fire safety management	Simulation (Escape route Fire equipment information), Condition
40	Engineering analysis	[137]	Solar thermal integration in early design	Orientation, Inclination
41	Engineering analysis /Asset Management	[222]	Detection, Identification and Pose Estimation of Lamps	Point cloud data, Location, Image
42	Engineering analysis/ Asset Management	[223]	Detection, Identification and Pose Estimation of Lamps	Point cloud data, Image
43	Existing Conditions Modeling	[224]	Scan-vs-BIM based automated construction progress monitoring	Point cloud, Image
44	Asset Management	[225]	Appliance control for smart buildings	Position data (9 data points per second), Acceleration data
45	Building System Analysis	[226]	Energy efficiency during operation	Humidity, Luminosity, Temperature
46	Phase Planning	[227]	Improving schedule performance of prefabricated house construction	Schedule, Location
47	Building (Preventative) Maintenance Scheduling	[157]	Fault Detection and Diagnostics	Energy Consumption
48	Safety/ Space Management and Tracking	[228]	Sensor-based tower crane navigation system	Location tracking
49	Asset Management	[229]	Integration of the BIM software with real-time monitoring (Energy)	Energy consumption, Voltage sensor, Current transducer
50	Phase Planning (4D Modeling)	[230]	Progress monitoring and management of the construction of a reinforced concrete (RC) structure	Construction progress (Schedule)
51	Logistic Planning/ Phase Planning (4D Modeling)	[231]	Logistics	Schedule
52	Asset Management	[140]	Visualizing thermal changes in spaces	Light, Current, Co2, Motion, Humidity
53	Cost Estimation	[147]	Cost estimation for tender of building projects	Cost
54	Sustainability (LEED) Evaluation	[232]	Streamline the environmental assessment of buildings	(LEED evaluation), Code
55	Cost Estimation (Quantity Take-Off)	[154]	Detailed Estimate	cost
56	Building System Analysis	[233]	Post occupancy evaluation (energy house example)	Temperature, Humidity

57	Engineering Analysis	[134]	Integrating building and urban semantics to empower smart water solutions	Water consumption
58	Building System Analysis	[118]	Thermal performance analysis and thermal comfort evaluation	Temp, Humidity
59	Asset Management/ Space Management and Tracking	[234]	Image-based indoor localization	Image
60	Asset Management	[152]	Corrective maintenance	Condition, Temperature, CO2, Occupancy, Point cloud
61	Building (Preventative) Maintenance Scheduling	[143]	Building preventive maintenance	image
62	Disaster Planning	[106]	Emergency response	Temperature, Location
63	3D Control and Planning (Digital Layout)/safety	[235]	Steel column inspection	Video, Location, Position(tilt)
64	3D Control and Planning (Digital layout)/ Phase Planning (4D Modeling)/ Cost Estimation	[119]	Construction material layout planning	Schedule, (QTO)
65	Code Validation	[148]	Automated compliance checking	Building code (Text)
66	Existing Conditions Modeling	[236]	Automatically identify structural elements for the purposes of Scan to-BIM	Point cloud
67	Building System Analysis	[135]	Measured and predicted environmental and energy performance	Energy consumption data
68	Asset Management	[237]	(BIM) Enabled Facilities Management	Temperature, Water sensor
69	code validation/ Phase Planning (4D Modeling)	[238]	Automatic Safety Checking of Construction Models and Schedules	Schedule, best practice (text)(code)
70	Building System Analysis	[239]	Thermal Energy Modelling and Simulation,	Air temp, Relative humidity
71	Asset Management	[120]	Streamlined Fault Localization	Video, location
72	Digital Fabrication	[240]	Free-Form Building Project	Drawing (image)
73	Cost Estimation	[241]	Quantity takeoff/ Labor productivity/ visual progress control systems	Cost, Schedule
74	Digital Fabrication	[121]	Digital Fabrication	-
75	Existing Conditions Modeling	[242]	Construction progress monitoring	Point cloud, Location
76	Engineering Analysis	[122]	Thermal environment design system for renovation using augmented reality	Indoor wall temperature, Outdoor temp, Emissivity, Absorptivity, Transmissivity, Camera

77	Asset Management/ Building System Analysis	[133]	Environmental and energy management of buildings	Room temperature, Relative humidity
78	Phase Planning (4D Modeling)	[243]	Dynamic Time–Space Conflict Detection and Quantification	Schedule
79	Logistic Planning	[244]	Automated identification and location estimation of construction materials, equipment, and tools	Location, Tracking
80	Building System Analysis	[114]	Facilities management/Building performance	Energy consumption
81	Asset Management	[245]	Operation and maintenance	HVAC, Electrical supply, Water supply
82	Engineering Analysis	[246]	Impact of building orientation on energy consumption	Orientation
83	Asset Management	[247]	Facility management using indoor localization	Image, Position, Orientation
84	Cost Estimation	[146]	FM	Cost
85	Code Validation/Safety	[248]	Rule-based Code Checking validates	Code
86	Cost Estimation	[116]	Building Management System to support building renovation	Cost, HVAC, Lightening
87	Building System Analysis	[249]	Reduce energy consumption and provide comfort conditions	Humidity, Temperature, Lightening, Ambient noise, User behavior
88	Existing Conditions Modeling	[250]	Collecting automatically live "as-built" data.	(RFID) location
89	Design Reviews	[251]	Virtual prototype	(Predefined task)

Appendix D: List of Unified BIM Uses

Unified BIM use list						
Sl No.	Building Phase	BIM use		Sl No.	Building Phase	BIM use
1	Plan Design Construct Operation	Existing Conditions Modeling		13	Construct	Digital Fabrication
2	Plan Design Construct Operation	Cost Estimation		14	Design Construct	Construction System Design (Virtual Mockup)
3	Plan Design Construct	Phase Planning (4D Modeling)		15	Construct	Logistic Planning
4	Plan	Site Analysis		16	Construct	Safety

	Design			Operation	
5	Plan Design	Programming	17	Construct Operation	Record Modeling
6	Plan Design	Design Reviews	18	Operation	Disaster Planning
7	Design	Code Validation	19	Operation	Space Management and Tracking
8	Design	Sustainability (LEED) Evaluation	20	Operation	Asset Management
9	Design	Engineering Analysis	21	Operation	Building System Analysis
10	Design	Design Authoring	22	Operation	Building (Preventative) Maintenance Scheduling
11	Design Construct	3D Coordination	23	Operation	Commissioning
12	Construct	3D Control and Planning (Digital layout)			

Appendix E: Object Properties and Their Characteristics for the Proposed AFDDOnto

Object Property	Domain	Range	SubProperty Of	Inverse Of	Characteristics	Disjoint With
adjacentZone	Zone	Zone			Symmetric	intersectsZone
algorithmBelongsTo	Algorithm	Element				
canDetect	Algorithm	Fault				
canUse	Algorithm	Feature				

containsZone	Zone	Zone			Transitive	
hasBuilding	Zone	Building	ContainsZone			
hasSpace	Zone	Space	ContainsZone			
hasStorey	Zone	Storey	ContainsZone			
elementsLocation	Element	Zone	ContainsElement			
faultBelongsTo	Fault	Element				
featureBelongsTo	Feature	Element				
hasAlgorithm	Element	Algorithm				
hasAssociated	Element	Feature				
hasElement	Zone	Element	ContainsZone o hasElement subPropertyOf:hasElement			
adjacentElement			hasElement			intersecctingElement
containsElement			hasElement containsZone o containsElement subPropertyOf:containsElement	elements Location		
intersectingElement			hasElement			adjacentElement

hasFault	Element	Fault				
hasInspection	Fault	MaintenanceInspection				
hasMaintenance	Element	MaintenanceAction				
hasParameter	Algorithm	Algorithm				
hasState	Feature	State				
hasSubElement	Element	Element				
interfaceOf	Interface					
intersectsZone	Zone	Zone		intersectsZone	Symmetric	adjacentZone
isAssociatedWith	Feature	Fault				
isParameter	Algorithm	Algorithm				
isRecorded	State	Track				

Appendix F: Natural Language Competencies and SPARQL Construct

prefix : <https://github.com/arashhosseiniarash/AFDDOnto#>
prefix owl: <http://www.w3.org/2002/07/owl#>
prefix rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
prefix xml: <http://www.w3.org/XML/1998/namespace>
prefix xsd: <http://www.w3.org/2001/XMLSchema#>
prefix rdfs: <http://www.w3.org/2000/01/rdf-schema#>
Prefix bot: <https://w3id.org/bot#>
Prefix brick:<https://brickschema.org/schema/Brick#>

1- Query the name of the Algorithms used under NeuralNeteorks and SVM

```
SELECT *
{
  ?NN_Algorithms_used a :NeuralNetworksOrDeepLearining .
  ?SVM_Algorithm_used a :Svm
}
```

2- What are the instances belonging to the RTU class

```
SELECT ?x
WHERE {
 ?x rdf:type brick:RTU
}
```

3- RTU belongs to which class? AND AHU belongs to which class? AND HVAC_Equipment belongs to which class?

```
SELECT ?x
WHERE {
 brick:RTU rdfs:subClassOf ?x
}
AND
SELECT ?x
WHERE {
 brick:AHU rdfs:subClassOf ?x
}
AND
SELECT ?x
WHERE {
 brick:HVAC_Equipment rdfs:subClassOf ?x
}
```

4- List the Fault names belonging to HVAC x

```
SELECT *
{
  ?Fault_name :faultBelongsTo ?HVAC .
  }
```

5- List the parameters used for all the respective algorithm
OR List the parameters used for Svc
OR List the parameters used for SequentialNeuralNetworkModel

```
SELECT *
{
  ?ParametersUsed :isParameter ?AlgorithmUsed .
}
OR
SELECT ?ParametersUsed
{
  ?ParametersUsed :isParameter :Svc.
}
OR
SELECT ?ParametersUsed
{
  ?ParametersUsed :isParameter :SequentialNeuralNetworkModel.
}
```

6- What is the name of the Algorithm(instance) under the class SVM
   OR
   What is the name of the Algorithm (instance) under the class NN

```
?AlgorithmNAme a :Svm
OR
?AlgorithmNAme a :NeuralNetworksOrDeepLearining
```

7- For HVAC x, what Features are needed?

```
SELECT ?HVAC_Name ?Feature_Name
WHERE { ?HVAC_Name :hasAssociated ?Feature_Name }
```

8- For Hvac x, which algorithm is used? And to which class does the algorithm belong.

```
SELECT ?HVAC_Name ?Algorithm_Used ?Class_name
WHERE { ?HVAC_Name :hasAlgorithm ?Algorithm_Used .
      ?Algorithm_Used a ?Class_name
     }
```

9- For Algorithm x, what Parameters are needed?

```
SELECT  ?Algorithm_name ?Parameter ?Value_Or_Description ?value
WHERE { ?Algorithm_name :hasParameter ?Value_Or_Description .
      ?Value_Or_Description a ?Parameter .
OPTIONAL
{?Algorithm_name :hasParameter ?Value_Or_Description .
```

```
       ?Value_Or_Description a ?Parameter .
 ?Value_Or_Description rdfs:comment ?value .}
      }
ORDER BY ASC (?Algorithm_name)
```

10- For HVAC x which faults are detected?

```
SELECT  ?HVAC_Name ?Fault
WHERE { ?HVAC_Name :hasFault ?Fault .
      }
ORDER BY ASC (?HVAC_Name)
```

11-What features are needed for Algorithm x?

```
SELECT  ?Algorithm_Name ?Feature_Name
WHERE { ?Algorithm_Name :canUse ?Feature_Name  .
      }
ORDER BY ASC (?Algorithm_Name)
```

12- What subsystems does HqVAC system x contain? Provide their GUIDs?

```
SELECT ?HVAC_system ?Element_GUID ?HVAC_sub_system ?Additional_information
WHERE {
?HVAC_system bot:hasSubElement ?Element_ID .
  ?Element_GUID a ?HVAC_sub_system .
  ?Element_GUID rdfs:comment ?Additional_information
}
```

13- The HVAC equipment name and GUID at each space

```
SELECT DISTINCT ?space ?ELement_GUID ?Description
WHERE {
?space bot:containsElement ?ELement_GUID .
   ?ELement_GUID rdfs:comment ?Description
}
```

Appendix G: Semi-structured Survey for Clarity Validation of AFDDOnto

The "AFDDOnto" survey captures the main concepts for the building and its HVAC system that most users need to know to make an AFDD model. This survey is meant to determine how clear the model is to its users and to look more closely at some concepts exclusively used in AFDDOnto.

These main concepts are:

1-**Algorithm**: The type of Machine Learning algorithm that was used for AFDD.
2-**Element**: It contains the different types of HVAC Equipment.
3-**Fault**: The different kinds of problems that can happen to the HVAC system.
4-**Feature**: Data being collected from the sensor that can be used for AFDD.
5-**Information**: Supplementary information about the building envelope and HVAC as a physical asset
6-**Interface**: A general term for the relationship between two or more things in the world, where at least one is a building element or zone.
7-**Maintenance**: Information about HVAC system maintenance, including an action plan and inspection plan
8-**State**: This shows how an entity of interest is doing right now.
9-**Track**: Keeping an eye on a State for a certain amount of time.
10-**Zone**: A part of the real world or a virtual world that is both in this world and has a 3D space.
Is there important information that can be used for AFDD that you couldn't find in the AFDDOnto concepts?

Here are the survey discussion topics you can contribute to increasing the ***Clarity*** of the 'AFDDOnto'; ***Clarity*** measures how effectively the 'AFDDOnto' communicates the intended meaning of the defined terms.

Concepts: Maintenance and Information for AFDD
Discussion 1: Do the concepts of **Maintenance** and **Information** overlap in any way that can be used for AFDD?
Maintenance: Information about HVAC system maintenance, including an action plan and inspection plan

Information: Supplementary information about the building envelope and HVAC as a physical asset

If Yes, What are the areas of overlap between **Maintenance** and **Information** concepts? If No, how can the two concepts (**Maintenance** and **Information**) be related? What information from these two can be used for AFDD? If any, note the relationship.

Concepts: Feature and State with respect to Element
Discussion 2: When considering the HVAC system (**Element**), is there any overlap between **Feature** and **State** concepts?
Feature: Data being collected from the sensor that can be used for AFDD.

State: This shows how an entity of interest is doing right now.
Element: It contains the different types of HVAC Equipment.

If Yes? What are the areas of overlap between the **Feature** and **State** concepts? How can the **Feature** and **State** concepts be related individually (**Feature** and **State** separately) or together (In terms of States representing different Features) to the **Element** concept representing HVAC equipment?

Concept: Track with respect to Feature AND/ State
Discussion 3: How can the **Track** concept be used to give the user a history of the HVAC system's problems in either a binary form of faulty or un-faulty or categories (Type of the faults)? Considering the existence of **Feature** and **State** concepts.
Track: Keeping an eye on a **State** for a certain amount of time.
Feature: Data being collected from the sensor that can be used for AFDD.
State: This shows how an entity of interest is doing right now.
In AFDDOnto, the **Feature** concepts have States, and States can further be Tracked.

Is there any overlap between **Track**, **Feature**, and **State** concepts? Can you identify the type of relationship that may exist?