INSTITUTO SUPERIOR DE ENGENHARIA DO PORTO

MESTRADO EM ENGENHARIA INFORMÁTICA

Módulo de Gestão de alarmes na Indústria 4.0

TIAGO FILIPE MOREIRA COELHO

outubro de 2022

POLITÉCNICO
DO PORTO

Instituto Superior de Engenharia do Porto

Alarm Management Module for 4.0 Industry

Tiago Filipe Moreira Coelho

Dissertação para obtenção do Grau de Mestre em Engenharia Informática, Área de Especialização em Sistemas de Informação e Conhecimento

Orientador: Dr. Isabel Praça

Júri:

Presidente:

Vogais:

Porto, 15th October 2022

Dedicatória

Resumo

Com a constante evolução tecnológica, mais particularmente na área da indústria, tornou-se necessária a evolução dos métodos de monitorização de forma a garantir a segurança e o funcionamento devido de todos os domínios em questão. Com o intuito de contribuir para a monitorização de empresas com alto envolvimento tecnológico, o GECAD iniciou o desenvolvimento de um projeto, *Alarm Management Model for 4.0 Industry,* que consiste no desenvolvimento de um sistema capaz de gerir alarmes de várias fontes diferentes, facilitando o acesso a informação sobre os mesmos de forma clara e objetiva. O foco principal do projeto passa por definir, com base nos alarmes detetados, prioridades entre eles de forma individual ou combinada e permitir ao utilizador ter uma experiência simples e eficaz aquando da utilização do sistema desenvolvido.

Palavras-chave: Indústria 4.0, Gestão de Alarmes.

Abstract

With the constant technology evolution, more particularly in the industrial area, it became the necessary the evolution of the monitorization methods in order to guarantee the security and the behavior of all the compartments in the system. With the goal of contributing to the monitorization of this type of companies, GECAD started the development of a new project, Alarm Management Model for 4.0 Industry that consists of developing a system capable of detecting alarms from various sources and manage it. The main focus of the project is about defining, based on the detected alarms, the individual and combined priority of them so it allows the user to have a simpler experience when using the developed system.

Keywords: 4.0 Industry, Alarm Management.

Agradecimentos

Index

Table of Figures

Table of Tables

Notation and Glossary

Acronym List

GECAD	*Research Group on Intelligent Engineering and Computing for Advanced Innovation and Development*
IoT	*Internet of Things*
KPI	*Key Performance Indicator*
ADAS	*Advanced Driver Assistance Systems*
AI	*Artificial Intelligence*
IT	*Information Technology*
W3C	*World Wide Web Consortium*
RDF	*Resource Description Framework*
SSN	*Semantic Sensor Network*
OWL	*Web Ontology Language*
ML	*Machine Learning*
BRMS	*Business Rule Management System*
MVEL	*MVFLEX Expression Language*
DSL	*Domain Specific Languages*
DMN	*Decision Model and Notation*
SaaS	*Software as a Service*
IaaS	*Infrastructure as a Service*
PaaS	*Platform as a Service*
TLS	*Transport Layer Security*
APM	*Application Performance Management*
EDM	*Electrical Discharge Machines*
TCP	*Transmission Control Protocol*

UDP	*User Datagram Protocol*
SSH	*Secure Shell*
VNC	*Virtual Network Computing*

Symbol List

λ_{max}	*Matrix average value divided by the priority vector value*

1 Introduction

Technology is present in our daily life and it does not stop evolving. It is present in many different application domains and the industrial area is no exception. The constant growth of technological development in this area generates an immeasurable number of new challenges, related not only to the management of a huge variety of new devices and all the information associated to them, but also to the interaction between humans and robotic automation. With the introduction of various types of new technologies, it is expected that the number of vulnerabilities increases, enforcing the existence of analysis and supervision tools on the industrial area to guarantee proper functioning of the machines involved and workers and information's safety. It is mandatory to have a good management of all the information provided to have an efficient decision process that can respond to problems with ease and precision. The high demand for this type of technology culminated in many companies investing in this market hoping to find a way to manage devices using intelligent technology, which resulted in the development of Alarm Management Modules.

This section consists in introducing the *Alarm Management Module for 4.0 Industry* project, by showing an overall overview of the project and its goals, approach, and planning.

1.1 Project Context

An Alarm Management Module can be defined as a system used for prioritizing, grouping and classifying event notifications used in supervisory control and data acquisition. Even though the potential of this type of systems is going to be explored in the context of industry, Alarm Management Functions can be applied to many other areas such as IT or medical care.

In the industrial context, the possession of an advanced alarm management system can bring many benefits (B.R. Mehta, 2015), such as:

- Improving operator effectiveness, securing uptime and safety and reducing potential losses.
- Minimizing the number and impact of abnormal situations.
- Reducing time and effort to develop, deploy and maintain alarm systems.
- Identifying alarm system problems and performance, as well as workers workload.

Additionally, not having a proper alarm management system can lead to:

- No alarms, despite of occurrences of certain events.
- Poor prioritization of the alarms.
- Flood of alarms with no relevancy.
- Missing events due to poor configuration.

In normal situations, an excessive presence of alarms can cause fatigue to the workers that are managing the system as humans naturally are only able to do a limited number of things at a time, therefore it is necessary to make sure that there is a way to present the alarms on a rate that the humans behind the process are able to act upon the presented problems.

In practical terms, an Alarm Management Module that is properly configured will reduce the number of alarms shown to the operator, presenting only the ones that forcibly need a decision from a human being. Other way of easing the operator’s decision making is providing the list of the alarms to take care of priority-wise so he knows what to do first.

The project described in this report will have as main goal the analysis of the current state of Alarm Management related work and the usage of that same analysis to produce a unique system, capable of managing alarms using different available technologies. It is developed in association with GECAD, a research center that has as main goal the development of scientific research in the areas of Intelligence in Engineering and Computing Complex Systems (GECAD, 2013). The development of this project is integrated in TMDEI - Thesis - subject at ISEP.

1.2 Problem Description

The technological development comes with a lot of new problems and flaws that when added to human potential mistakes can bring dangerous outcomes. Therefore, it is necessary to create a project where the management, in this case of alarms, will be the most intuitive possible and the least dependent on humans. Organizations are using more often devices that allow the collection of data from different sources and trying to utilize this data as knowledge to create a system capable of making decisions based on the origin of the alarms and their properties. Many variables are considered when making these decisions based on information provided by several systems that cooperate exchanging data and provide:

- Risk Assessments.
- Simulations of Impact.
- Vulnerability Analysis.

However, even though the plan is to ideally make the management of the information easier to the system and have less human input on the outcomes of the decision-making process, the development of such ideas is still in its early stages and the systems still need some human intervention to supervise it and make sure it is working properly. Nevertheless, the fact that this type of technology is not fully explored yet, creates the desire of further investigations by many organizations.

Despite the fact that alarm systems continue to be a huge concern when it comes to the fact that they need to conform according to the current standards and best practices, in order to guarantee that it will ease the operator's during complicated processes and abnormal situations, they continue to show flaws when it comes to continuous improvement causing the constant need for a change and investment by the end users as the alarm systems easily become obsolete.

1.2.1 Objectives

The main goal of *Alarm Management Module for 4.0 Industry* project is to develop an event management platform in industrial context that is flexible and scalable, potentially offering companies a platform that is willing to offer security, consistency, and good decision-making.

There are many specific objectives that should be attained in order to achieve the main goal of the project:

1. The system should be able to handle events from several different sources, such as access control, network intrusion detection, or energy sensors.
2. The platform should allow the receival of events from other existent tools being developed in GECAD.
3. The alarm management should be developed using intelligent technology, based on analysis and event correlation mechanisms, avoiding false positives, and securing a holistic and safe view of every combination possible.
4. The user should be able to interact with the system by a proper user interface, provided by an already established dashboard service or a developed dashboard.

1.2.2 Approach

The project's overall value chain consists in creating an environment where several sensors with different sources are connected directly to a correlation engine, a rule-based system parametrized by the user and willing to adapt to its own domain, with a safe and supervised connection ensuring that the information is reaching its destination properly. The information will then be stored in the database and can be easily accessed by the dashboard. The dashboard will be connected to the database and can store or retrieve data when necessary. The main goal is to allow the user to access all the data from the sensors regarding the alarms in a way that is easy to understand and make use of the receive information. The following figure shows how the whole process and how everything is connected.

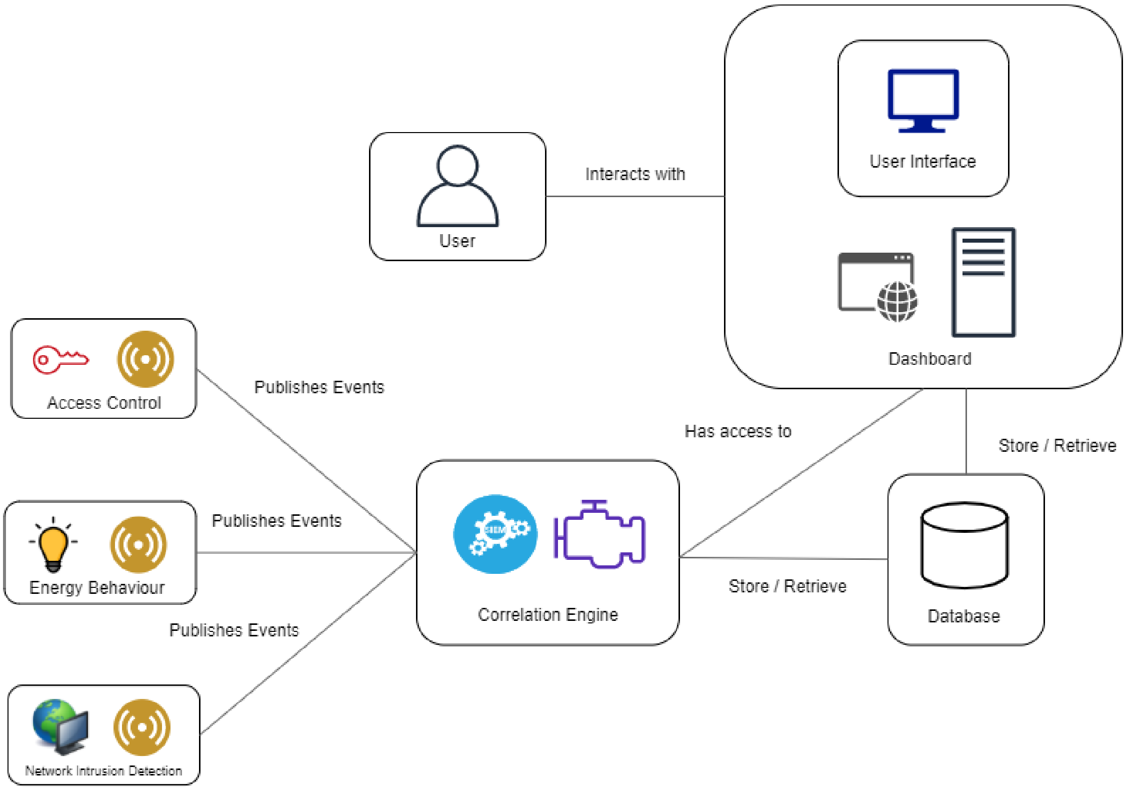


Figure 1 – Alarm Management Module Project Value Chain

1.2.3 Contributions

The main contribution of this project is exploring the various possibilities regarding alarm management systems by making a deep analysis of available technologies and related work and combining those ideas in order to develop a product that is viable for daily usage in companies, with the final result being a practical demonstration of it.

1.3 Report Structure

This report will consist of ten chapters, the introduction, the state of the art, value analysis, analysis, design and experiences and evaluation.

The introduction describes the project, focusing on explaining the problem, the approach, the objectives, and the main contributions.

The state of the art will provide to the reader an insight of the current existing technologies, both commercial and supportive tools, and related work in order to ease the process of understanding the project.

The semantic layer chapter will focus on presenting the proposition of semantic layer and how important it can be to the project.

The Rule-Based systems chapter will present an introduction to rule-based systems and explain how they can be included in the present project.

The value analysis focuses on presenting the value proposition of the project and the business model related to it.

The analysis will focus on presenting the entities of the project and its functional and non-functional requirements.

The design will present two alternative solutions and present both advantages in order to decide which one is being used in the project.

The implementation will focus on explaining the implementation process and justifying the decisions that were made throughout it.

The experiences and evaluation chapter will have as goal present the hypothesis and evaluation methods.

The conclusion will conclude the document making a resume about the project, objectives and future work.

2 Alarm management and existing tools

This section will allow the reader to understand the current state of alarm management, by making a deep analysis on related work, commercial tools and existing technologies, with emphasis on its objectives and impact that they have on the project. The latter will focus on supportive tools, mainly database services and dashboard services.

With the start of this project, some concepts such as Internet of Things (IoT) – defined as the connection between the internet and devices with the goal of automating processes (Oracle, 2021), Smart Alarm Systems, Alarm Management Systems, and Industry 4.0, were immediately associated with it as they seemed to be a key part for the overall understanding of the project. The concept of Industry 4.0, defined as the transformation of regular industries through Intelligent networking and with the help of information and communication technology (Daniel, 2018), is constantly evolving and is the main focus of this project, therefore being the most crucial keyword when trying to understand the project.

2.1 Related Work

This section will provide information regarding related work that can be relevant to the Alarm Management model for 4.0 Industry project, contributing to a better understanding of the concepts surrounding it. For each relevant piece of work related to the subject, there will be a brief presentation and resume of its content and then a explanation on how it is related to the project.

2.1.1 A context-based building security alarm through power and sensors analysis

A context-based building security alarm through power and sensors analysis is a paper focused on building a security alarm that identifies undesired situations when given the proper context (Silva, et al., 2018). With the increase of quantity and availability of information due to technological advance it is mandatory to know how to extract the maximum value from it, and with the development of technologies like IoT it is possible to ensure the information is secure.

The developed system is responsible for monitoring and identifying abnormal behavior in a pre-defined building and it will manage many different variables using sensors such as temperature, CO2, air quality and humidity. The system will confront the real values of those variables with the rules that will indicate the values that can be alarming to that specific situation. Those rules are also dependent on context as those can vary depending on the situation the analysis is being done. In order to detect the context the data is in, the system will separate it and analyze it based on its history. For that propose, clustering and other data mining techniques will be applied.

This paper is related to the present document as it presents a solution that focuses on taking advantage of sensors to increase security, being a crucial aspect to the environment where it is operating. Also, the project incorporates a rule-based system, which is one of the main goals of this project, helping to understand how it should be applied and what role it has on this specific subject.

2.1.2 Evaluation of Smart Alarm Systems for Industry 4.0 Technologies

Evaluation of Smart Alarm Systems for Industry 4.0 Technologies is a paper related to the evaluation of Smart Alarm Systems on the footwear industry and how cost control and labor intensity are managed in these companies by using Industry 4.0's concepts (Chang, 2020). The main goal of this document is to compare the performances of CEGRA (Cause-Effect Grey Relational Analysis) and TOPSIS (Technique for order preference by similarity to the ideal solution) in evaluating large data blockchain technologies and real-time early warning systems for production and raw material supplier management.

The case of study is the Datong Shoes and Materials, company founded in 1992, that with the 2008 global financial crisis decided to invest in new technologies to reduce its operating costs and turn their gross profit margin into positive values again. After the new changes in the way that the company was operating, it culminated in a change of culture that started supporting innovation, believing that the future of the company should be based on automated technologies, therefore reducing the manpower and potential wages. As almost 70% of the company was composed of 50 years or older engineers it increased the necessity of having an artificial intelligence system to preserve the knowledge of the current employees and transfer the essential part of it to the newest generation of workers. This requires a high initial investment to have profit later, but by minimizing the number of employees, the company wanted to maintain the level of substantial grow of 30% that it had before.

The following part of the document consists of comparing the CEGRA and TOPSIS algorithms on other four technology collaborative technology companies similar to the case of study. The CEGRA and TOPSIS procedures are summarized in the following figure.

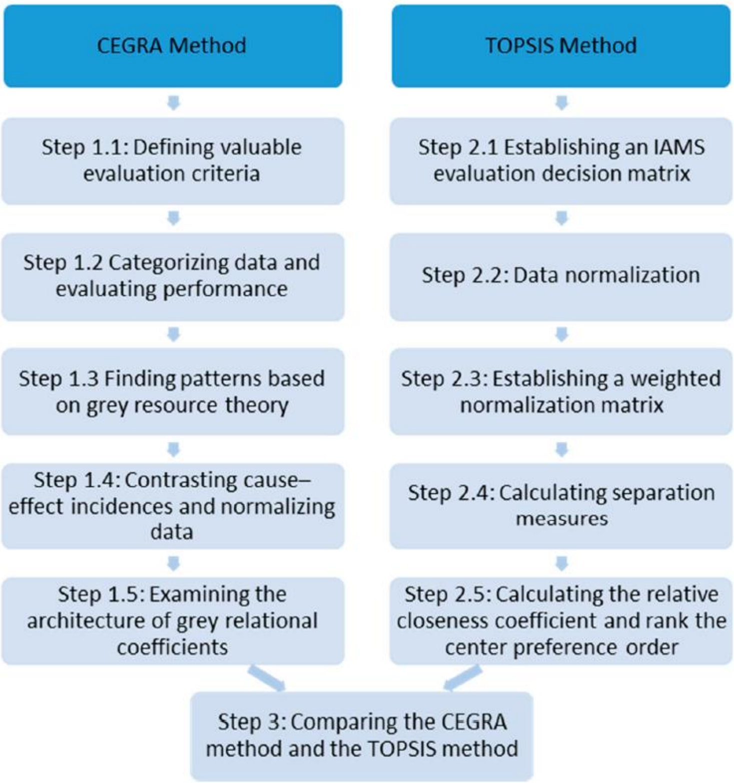

CEGRA Method
TOPSIS Method
Step 1.1: Defining valuable evaluation criteria
Step 2.1 Establishing an IAMS evaluation decision matrix
Step 1.2 Categorizing data and evaluating performance
Step 2.2: Data normalization
Step 1.3 Finding patterns based on grey resource theory
Step 2.3: Establishing a weighted normalization matrix
Step 1.4: Contrasting cause–effect incidences and normalizing data
Step 2.4: Calculating separation measures
Step 1.5: Examining the architecture of grey relational coefficients
Step 2.5: Calculating the relative closeness coefficient and rank the center preference order
Step 3: Comparing the CEGRA method and the TOPSIS method

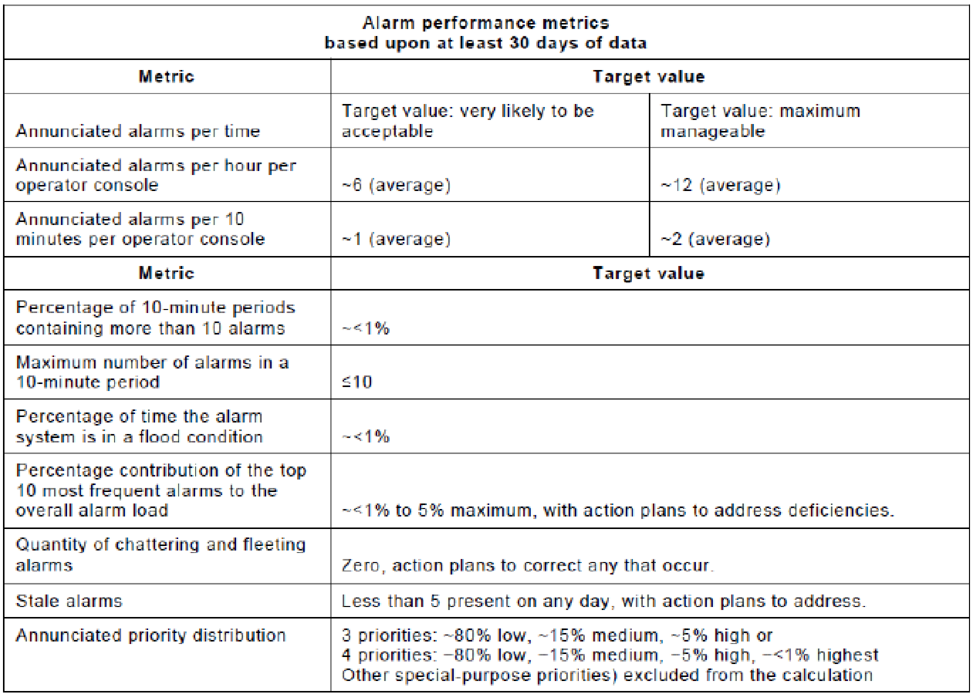
Figure 2 - Comparison between CEGRA and TOPSIS methods (Cheng, 2020)

After applying both methods on the case of study, it was possible to conclude that both techniques reached the same conclusions by ranking the options the same way. However, the CEGRA method proved that it could manage the uncertainty of decision problems, multiple decision evaluation attributes and provides a simpler method for decision making.

With this paper it is possible to understand how 4.0 Industry innovations is affecting companies in terms of structure and how it is replacing humans daily. Based on the impact these technologies have nowadays and after understanding properly what are the pros of having these techniques being applied, it is possible to conclude what are the goals for a project in this subject and what is its role in the final product.

2.1.3 The sense and nonsense of Alarm System performance KPIs

The sense and nonsense of Alarm System performance KPIs is a paper that is focused on addressing the current state of Alarm System performance's analysis. It starts by showing what are the recommended alarm performance metrics and respective target values as it is possible to see in the following figure.



Alarm performance metrics based upon at least 30 days of data		
Metric	**Target value**	
Annunciated alarms per time	Target value: very likely to be acceptable	Target value: maximum manageable
Annunciated alarms per hour per operator console	~6 (average)	~12 (average)
Annunciated alarms per 10 minutes per operator console	~1 (average)	~2 (average)
Metric	**Target value**	
Percentage of 10-minute periods containing more than 10 alarms	~<1%	
Maximum number of alarms in a 10-minute period	≤10	
Percentage of time the alarm system is in a flood condition	~<1%	
Percentage contribution of the top 10 most frequent alarms to the overall alarm load	~<1% to 5% maximum, with action plans to address deficiencies.	
Quantity of chattering and fleeting alarms	Zero, action plans to correct any that occur.	
Stale alarms	Less than 5 present on any day, with action plans to address.	
Annunciated priority distribution	3 priorities: ~80% low, ~15% medium, ~5% high or 4 priorities: ~80% low, ~15% medium, ~5% high, ~<1% highest Other special-purpose priorities) excluded from the calculation	

Figure 3 - Alarm Performance Metric Summary (Dubois, 1957)

After introducing the concept of alarm management metrics and its respective values, the document proceeds to refer the concept of operator whose role usually imposes many limits on alarm handling workload, as it has numerous responsibilities. The alarm rate should never trespass what the operator can handle, with the normal values being close to one alarm per ten minutes. The operator’s workflow is normally defined as in the following image.

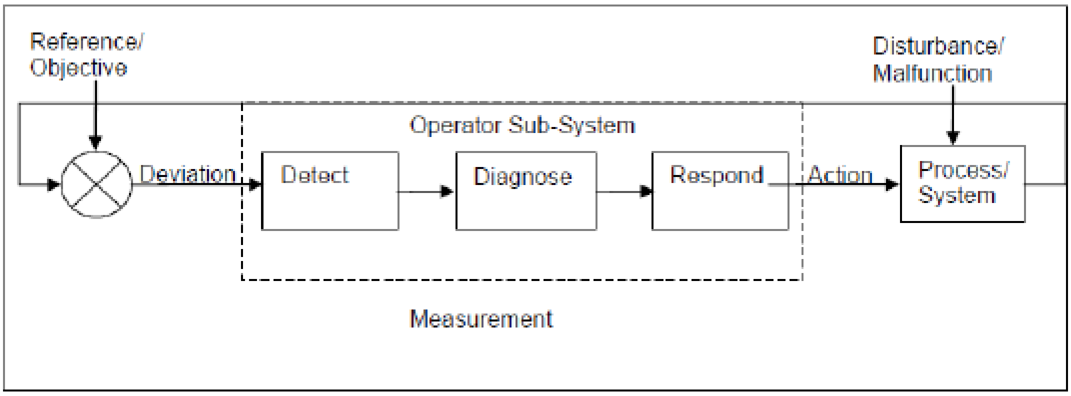


Figure 4 - Feedback Model of Operator Workflow (Dubois, 1957)

The document also provides an insight about the danger of averaging. Essentially it defends that one alarm per minute does not convert proportionally when referring to bigger numbers, as the performance of the operators is affected by the fact that they need to switch attention between innumerous alarms and perform actions accordingly for each one of them. With that being noted, it is important that when companies have interest in developing an alarm management system, they take into account the fact that averaging alarms per minute can be uncertain and harmful to the project.

This paper is important in order to understand how some minimal things like, in this specific case, rounding alarms per minute might have an impact on the overall result of the final product. Overall, it also gives some insight on how to approach this alarm per minute situation and what are the alarm performance metrics to be followed during the development of the project.

2.1.4 Towards intelligent alarm management in the age of IIoT

Towards Intelligent alarm management in the age of IIoT is a paper that addresses the problem of alarm floods in the era of IIoT, aiming to improve the current state of industrial operations by providing an easier way to manage alarms and increasing the reliability of each industry operation (Domova & Dagnino, 2017). The project developed a set of analytic models implemented in Java that were using data mining approaches, allowing to find patterns between the historical dataset and potential variables. The models developed consisted in:

- Hiding Rules Analysis
- Critical Event Analysis
- Sequence Analysis

The first one, Hiding Rules Analysis, is based on the fact that there might be alarms in a management system that may not be needed to be displayed to the operator in a plant as they are likely to be to be directly or indirectly addressed with other alarm, having the overall goal of reducing the number of alarms that can be considered redundant.

On other hand, Critical Event Analysis is one of the most crucial types of analysis in the models implemented as it has the overall goal of identifying the alarms that occur either before or during a critical situation, this way operators have an easier time dealing with potentially dangerous events as well as tracking the whole event from its beginning until its end.

The last one, Sequence analysis, is focused on analyzing the overall logs and historical data of alarms in order to understand what type of alarms usually happen together, and if it is possible to build a logical sequence with that knowledge. Additionally, it there is a situation where a logical sequence is built, it calculates the average time between each alarm on that sequence and the standard deviation for each element.

To sum up, the research developed in the paper provides valuable insight on the current state of intelligent solutions on alarm management, therefore being a great document to analyze before developing a project in this area.

2.2 Commercial Tools

This section will focus on tools that are commercialized and similar to the result that it is supposed to be achieved during this project. It will allow the reader to understand more regarding those similarities and some positive aspects that might be indicated to be implemented in the project.

2.2.1 ProIMS Alarm Management

ProIMS is an information management system based on web used for all kinds of real time processes. It is built on real time databases and is presented as an Industry 4.0 solution for process data integration and calculated data into a single central system (ProAsset, 2021). ProIMS allows archiving process data like process values, events, or messages into its web-based reporting system.

ProIMS Alarm Management is an integrated tool ProIMS used with the goal of managing and viewing alarms in order to assist managers and supervisors in monitoring the behavior of certain alarm setpoints. It provides a set of tools and features that are relevant to improve effectiveness on the operator’s performance, including:

- Long Term Storage
- Web based Alarm historical view
- Web based Alarm Evaluation
    - Most frequent alarms over time
    - Alarm evaluation over time
    - Top alarm identification
    - Operator intervention statistics
    - Alarm sequence analysis
    - Identification of chattering alarms
- Real time alarm view
- Dashboard elements to ensure customized view

The process starts when the system receives alarms and signals using the chosen interface and evaluates the alarms criterion. Then, the evaluation process signals the generation of predictive alarms in order to store it in the designated area so the alarm’s indicators and summaries to use in KPI applications can be calculated. With the first steps being processed, the system will allow real time view of the calculation and display the alarms management information results using the designated dashboard and export the results into a Microsoft Excel file.

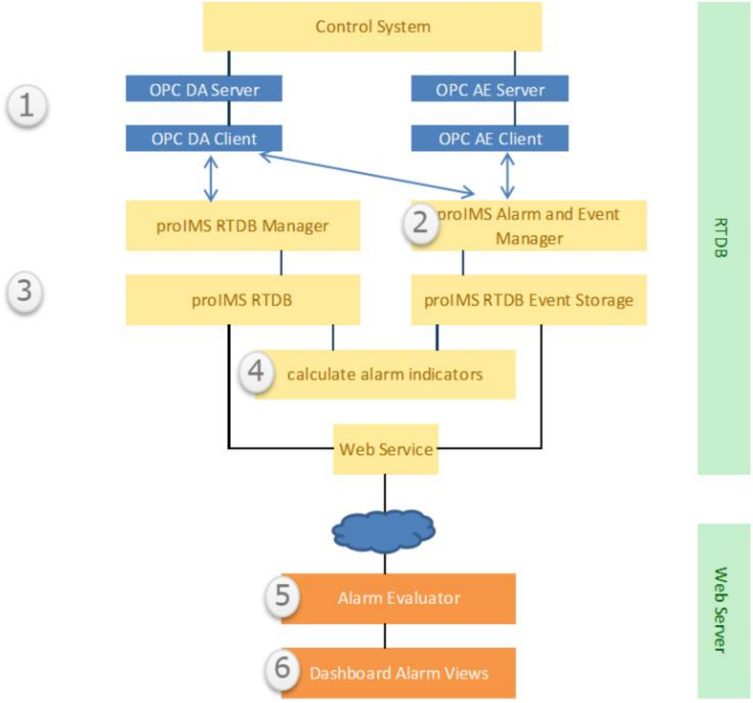


Figure 5 - ProIMS Alarm Management Process (ProAsset, 2021)

The system developed by ProIMS is important to the development of the project as it has many functionalities available that are expected and mandatory in an alarm management system and can be used as example. Even though it seems like a solid choice in terms of how it works overall and the functionalities it has available, the fact that it is only available to certain companies makes it an impossible option to be explored more deeply.

2.2.2 ProcessVue Alarm Management Suite

ProcessVue Alarm Management Suite is an alarm management software that is designed to meet the requirements of most industries, including Gas, Oil, Power Generation, Pharmaceutical and General Manufacturing. It provides key modules that have as main goal managing all the incoming alarms with clear, relevant, and prioritized information. It then plans and complies them based on the global standards of alarm management.

The ProcessVue Alarm Managent Suite is based on three subsidiary modules which compose this tool:

- ProcessVue SOE (Sequence of events)
- ProcessVue Analyzer
- ProcessVue Guardian (Management of Change)

The first one, ProcessVue Sequence of Events, focuses on the collecting and storing on events from different data sources, providing, additionally, the capability to filter and search data. It will allow the user to have access to information like most frequent alarms, its distribution over set time period and an extensive alarm and event routing system.

The second one, ProcessVue Analyzer, provides a KPI presentation using dashboards, Frequency Analysis and Standard Reports. With the inbuilt reporting engine being able to schedule automatically ran reports at a desired frequency. It will help with the identification of chattering alarms and stale alarms, root cause analysis and the unification of all alarms and events in just one tool.

The last one, ProcessVue Guardian, allows the global system to be controlled, is tailored to fit the user's perspective on the requirements of the system and allows the user to respond manually to alarms and react to them.

Overall, ProcessVue Alarm Management Suite seems to be one of the most promising commercial tools and a solid choice to established companies, as it provides the user with a lot of functionalities that are crucial to the good functioning of a proper alarm management system.

2.3 Existing Technologies

This section will provide the reader with information regarding existing technologies that are relevant to understand the project. There are three sub-sections, the first one will present Kylo, the second one will present storage services and the final one dashboard services.

This section and will focus on tools that have supportive role and that can be relevant to the context of this project. It will allow the reader to understand more regarding similarities between the services stated in this section and specific aspects that might be crucial to implement in the project.

2.3.1 Kylo

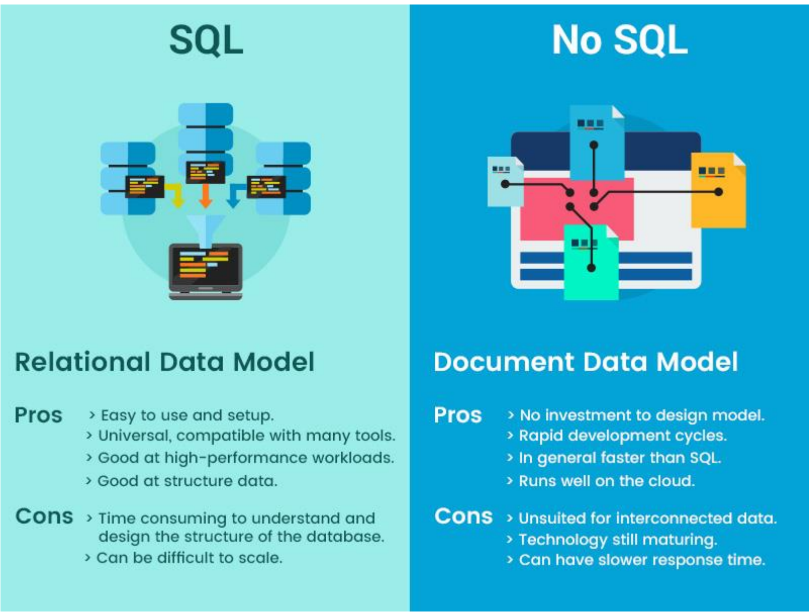
Kylo is an open-source data lake management software platform (Kylo, 2017). It is used by airlines, insurance, banking and retailing companies which proves how reliable and known this platform is. The platform focused on the following features:

- Ingestion – Kylo allows the connection to many sources and infer schema from most data formats, simplifying the ingestion process and allowing a versatility that other services do not have.
- Preparation – Kylo provides the latest capabilities of Apache Spark which can allow the user to have an interactive experience when transforming data.
- Discovery – Kylo includes integrated metadata repositories and capabilities for data exploration.
- Monitoring – Kylo is innovative in terms of monitoring systems as it provides health indicators from a feed-centric perspective, meaning operators can also track service levels associated with data arrival.

2.3.2 Storage Services

Storage Service is one of the most crucial parts of the project as it will include all the data regarding the alarms, so it is necessary that the best service is chosen to take part in the project’s architecture. The first step in the process of deciding a database service is choosing its type, noSQL or SQL. Usually, SQL databases are used in cases where it is known how the data is going to be organized, so there is the need of creating tables with relationships between them. Unfortunately, it will not be always possible to tell immediately what the best-case scenario is since sometimes it can be worth to use a NoSQL database as it has a better performance, even though the tables will have no relation. In the end, it will come down to whether it compensates more to have an organized database and lower performance or the opposite (MongoDB, 2021).

In the case of this specific project, the initial idea is to have a non-relational database as in some other projects from the GECAD research center are already using this type of database and linking it to other tools that have potential interest to the present work. Though, it is important to refer that it is not mandatory to be using a non-relational database, so if after a deeper analysis it is concluded that using a relational database is better for this specific project, then it should be used.



Relational Data Model

Pros
- Easy to use and setup.
- Universal, compatible with many tools.
- Good at high-performance workloads.
- Good at structure data.

Cons
- Time consuming to understand and design the structure of the database.
- Can be difficult to scale.

Document Data Model

Pros
- No investment to design model.
- Rapid development cycles.
- In general faster than SQL.
- Runs well on the cloud.

Cons
- Unsuited for interconnected data.
- Technology still maturing.
- Can have slower response time.

Figure 6 - SQL and NoSQL Comparison (Shah, 2021)

2.3.2.1 MongoDB

MongoDB is a non-relational document-oriented Database Service that allows the storage of JSON formatted documents (MongoDB, 2021). It is one of the most used Database Services as it has a good failure handling, creating huge number of data replicas allowing greater protecting of the data and lesser downtime failures, which can be important in the present subject. Also, the fact that it stores JSON documents is also a very feature of this Database as it is a possibility that all the data received within the scope of this work is in that format.

2.3.2.2 Elastic Search

Elastic Search is defined as a distributed search and analytics engine developed in Java (Search, 2021). It allows real-time storing, searching, and analyzing of huge amounts of data. Just like MongoDB, it uses a document-oriented structure instead of tables so it is possible to see Elastic Search as a server that can process JSON requests and answer it with JSON data. One of other aspects of choosing Elastic Search as the Database Service to the project is that It can easily connect with Kibana, which will allow a better data visualization, by providing histograms, line graphs, piece charts or even maps.

2.3.2.3 Azure Database Services

Azure has many services that are related to databases, and each one of them has its own specification. The Database Services that Azure provides their users with are: Cosmos Database, Azure SQL Database and Azure Database for mySQL.

Cosmos DB is a multi-model database service. It can be defined as schema-agnostic and is classified as a NoSQL Database. Azure SQL Database and Azure Database for MySQL, on other hand, differ from Cosmos DB as they are classified as SQL Databases (TrustRadius, 2021).

Even though the crucial differences are between the NoSQL and SQL Databases, Databases services inside the same type can still be different (Microsoft, 2020):

- Azure Database for MySQL is open source, while Azure SQL Database is not open source and is mainly focusing on being commercial.
- Azure Database for MySQL was programmed in C++, while Azure SQL Server was programmed in C++ and C.
- Azure Database for MySQL was developed by Oracle, while Azure SQL Server, on other hand, was developed by Microsoft.
- Azure Database for MySQL only supports the English language, while Azure SQL Server allows the user to choose between many other available languages.
- Azure Database for MySQL supports many platforms, while Azure SQL Server is restricted to Linux and Windows.
- Azure Database for MySQL has a more complex syntax, while Azure SQL Server has a simpler syntax.

In order to receive all the data from various sources, the Database Services from Azure need to be linked to other service also from Azure, Azure Stream Analytics Job, that is defined as a high-level event processing engine designed to process high volumes of data coming from different sources. Overall, the service consists of defining an input, a query and an output, with the query being based on SQL Language. This service is useful in situations where considerable amounts of data need to be transferred from service to service securely (Microsoft, 2020).

2.3.3 Dashboard Services

A dashboard service is one of the most important parts of this project as it will allow the user to have a more interactive and unique view of how the system is looking which will result in a better experience. This section will focus on presenting Kibana and Angular as Dashboards and what is the positive points about them and how they can help the project.

2.3.3.1 Kibana

Kibana is, simply put, a data visualization dashboard for Elastic Search (Kibana, 2021). It is open-source and is normally used to explore log and time-series analytics, application monitoring and operational intelligence cases. The features that this service offers such as histograms, line charts, pie shorts and heat maps are very simple and user-friendly.

2.3.3.2 Angular

Angular can be defined as a platform and framework to build client applications using HTML and TypeScript (Angular, 2021). In the context of this project Angular will be explored as a Dashboard service and see how it performs compared to Kibana, which seems to be the most reliable option to this project.

3 Semantic Layer

This section will allow the reader to understand the importance of developing a semantic layer in the context of this project, how ontologies are related to it and how efficient they can be to fulfill the requirements of the project.

A semantic layer can be defined as a business representation of data, providing human-readable terms to data sources that in similar situations would be impossible. A semantic layer consists in mapping enterprise data into more familiar concepts with the main goal of achieving a common view of the data, which can be understood and learned easily.

3.1 Ontologies

An ontology is often defined as a formal specification of classes, attributes, and the relationships between them, with the purpose of enabling knowledge sharing and reuse. In the context of computer sciences, they are designed in a way that humans and computers can communicate, share knowledge, and automate the translation of unknown expressions transferring meaning between applications (Grubber, s.d.).

The usage of ontologies to represent knowledge as a set of concepts comes with a lot of advantages, such as increased quality of entity analysis or facilitation of domain knowledge sharing using vocabulary that is common and perceived accordingly across other independent domains (O'Reilly, 2022). Other advantages include enabling automated reasoning about data, provide more coherent and easy navigations as the concepts are related and are common in every domain, they are easy to extend, and they are able to represent any data formats including unstructured, semi-structured or fully structured data (ontotext, 2022).

Building an ontology means developing a knowledge system that includes hidden meanings and all the relationships between related concepts. It is also possible to enhance the original knowledge system by linking similar words’ meaning to new ones. In order to extract maximum potential out of the ontology, it must be:

- Easy to implement and maintain.
- Able to reuse existing code with ease.
- Adaptive to new knowledge.

This chapter will mainly focus on describing ontologies and explaining why they could be useful in the context of this project. It will also describe and compare the available languages, editing software and development methodologies.

3.1.1 Languages

An ontology language is defined as a formal language used to encode an ontology in both syntax and semantics (Ramakrishnan, 2010). The main feature of an ontology language that distinguishes them from regular databases is the capability of making logical deductions and decisions (Anon., 2015), with the allowance of incremental building, sharing and usage of knowledge also being big priorities.

There are numerous ontology languages that differ between themselves based on their properties. Ontology languages can be classified as:

- Logical Languages
    - First Order predicate logic.
    - Rule Based logic.
    - Description logic.
- Frame Based Languages
    - Similar to relational databases.
- Graph Based Languages
    - Semantic Network

Ontology languages are most of the times generalizations of Frame Based languages, usually based on First Order or Description logistics (Vallespir, 2009).

With that being said, in this section of the chapter some different ontology languages will be introduced, with this section being concluded deciding which language is the best moving forward in the project.

3.1.1.1 RDF

Resource Description Framework (RDF) is a standard model for data interchange on the web being first published in 1996 by World Wide Web Consortium (W3C) (Anon., 2004). RDF represents knowledge in a form of semantic graph where the edges represent the link between two main resources, entity and value where the predicate stands for the attribute. The usage of this model, which can also be called Entity Attribute Value (EAV), allows structured and semi-structured data to be mixed, shared and exposed across different applications.

Despite adding providing the basic capabilities needed for the project, its expressivity regarding the subject does not seem sufficient. Therefore, languages that provide additional required features such as cardinality constraints, transitive and inverse properties will have an upper hand when it comes to being chosen to such project (Group, 2014).

3.1.1.2 OIL

Ontology Inference Layer consists of an established layer structure for extension that satisfies all the requirements of semantic web. It is based on a Frame-based system, description logics and web languages. OIL offers hierarchical extensions, inference mechanism based on the description logic and well-defined semantics, but struggles to define the default-value, to provide the meta-class and it is impossible to support the concrete domain (Jong-Soo, 2014).

3.1.1.3 OWL

Web Ontology Language (OWL) can be defined as a semantic web language focused on represent complex knowledge regarding things, groups of things and relations between them (Anon., 2012). It was first introduced in 2004, with the latest version, OWL2, releasing in 2009.

OWL has as main goals:

- Defining classes and their properties within a specific domain.
- Defining relationships between classes.
- Defining relationships between instances.
- Defining properties of properties.
- Defining relationships between properties.

It is also important to refer that the full version of OWL can be considered an extension of RDF, which means consequently that every RDF file is a OWL file by default.

3.1.1.4 Selected Language

The language selected to develop the ontology was OWL. It can easily express complex knowledge and relationships and is very precise when it comes to applying precise constraints on concepts.

3.1.2 Editing Software

It is possible to write ontologies using any text editor available. With that being said, it is recommendable the usage of an editing software to make sure the ontology is functioning properly, because not using it makes it hard to track all the relationships between concepts as the ontology grows bigger just by human interpretation.

This section will introduce Protégé, Fluent and Ontolis as the tools considered for the development of this project, summarizing and establishing a comparison between them.

3.1.2.1 Protégé

Protégé is a free, open source ontology editor and a knowledge management system (Anon., 2016). Its meta-tool version was built in 1987 and has then been published in 2001 by Stanford University. It is the most used ontology development tool using desktop and web applications, having a large community of active users. The down part of this software is that it is a local software, not being able to be used by other users to edit the same ontology. Even though Protégé has some flaws those are minimal, and it is one of the most promising ontology editors on the market. The last updated version of this service was released in 2019.

It allows the user to:

- Import, edit and save existing ontologies.
- Save ontologies in various formats including XML.
- Visualize ontologies in graphical form.
- Populate ontologies with concrete instances of classes
- Execute reasoners that can perform inferences on a ontology.

3.1.2.2 Fluent

Developed by Cognitum Company, Fluent is an ontology editor focused on supporting ontology visualization, reasoning and Semantic Web Rule Debugging. It contains a predictive editor, preventing any incorrect sentences in terms of grammatic and morphological aspects (Anon., 2011).

3.1.2.3 NeOn

NeOn is an open-source multi-platform ontology engineering environment based on the Eclipse platform originally developed as part of the NeOn Project. The last updated version was released by the end of 2011 (Anon., 2014).

3.1.2.4 Selected Editing Software

Based on researching and a comparison done between the available software, it is safe to say that Protégé is the best option available as it is free, popular and covers all the needed features for the development of the project.

3.1.3 Ontologies and Alarm Systems

With the study of available and recent ontologies related to Alarms systems in general it was possible to understand that many of them can be relevant to the progress of this research. In this section those ontologies will be presented together with their characteristics and a brief explanation on how they can be impactful to the definition of this project's ontology.

One of studied papers during the research process (Zamora & Sipele, 2017) combined academic and industrial interests in order to develop an ADAS – Advanced Driver Assistance System – to help the driver in the driving process by providing privileged information regarding the vehicle and the surrounding environment. This type of initiatives, according to this project, are crucial to increase the safety of the drivers as the security on the roads is poor and the increasing number of accidents with deaths is worrisome. The support given to the driver is provided through alarms that can be both visual and sonorous.

The following figure shows the concept of the framework Intelligent Co-Driver that agglomerates the useful concepts for the design and development of the ADAS.

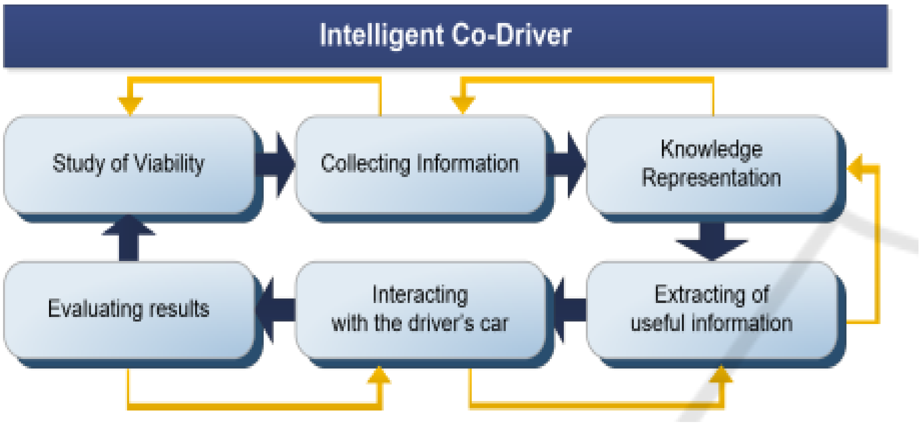


Figure 7 - Intelligent Co-Driver Framework

The Intelligent Co-Driver consists in a never-ending process that combines several research fields including predictive computational models. It starts with the study of respective viabilities and ends with the results evaluation, just to start again in a continuous improvement cycle.

The developed ontology will include concepts such as Car, CarContext, Padestrian, PadestrianCrossing, Actor and Driver and includes an expert system design addressing several risks such as:

- Risk of frontal collision.
- Risk of running over.
- Risk of rear collision.
- Risk of lateral collision.
- Padestrian not visualized.

The final ontology looks as follows:

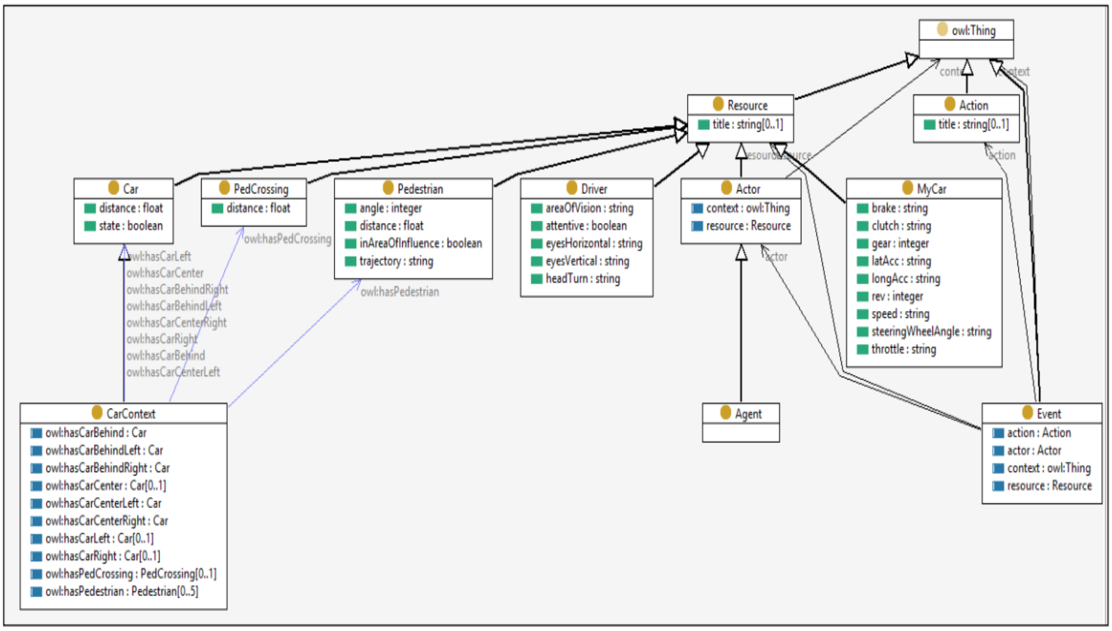


Figure 8 - Intelligent Co-Driver Ontology Diagram

Other paper that was studied during this process of understanding the involvement of ontologies on alarm management systems (García, et al., 2012) consisted of explaining how real-time alarm management systems could be used on guiding platform operators to the essential information regarding potential imminent danger in the petroleum operation domain. In this context, the analysed system is a petroleum process plant, composed of a set of tools that interact with each other and have the sole goal of transforming and conducting safely a fluid. Each tool has its own rules, limits, and safety devices – alarms – which will be represented as a group of agents in constant communication with a set of rules for acting, reacting and interacting with each other to follow in order to guarantee that the process is done safely. In this case, the system is based on an artificial intelligence multi-agent based approach and is bounded to external and non-predictable effects from outside of the system.

The architecture used for creation of the system consists in a process that only ends when the actions taken by the operator do not result in an unexpected alarm or presents a high danger degree that may compromise the proper functioning of the unit.

The process starts with the operator making an action and it is composed by five types of agents that share human-like characteristics including reasoning, proactivity, communication and adaptative behaviour, but have different functionalities, such as:

- Environment Agent – Focused on monitoring information from nature.
- Automation Agent – Focused on automation systems related to events and alarms embedded to the equipment. Also has the responsibility of creating a log history merging the information coming from the sensors and recording it on the Blackboard agent.

- Log Handler Agent – Focused on reading and parsing the log information stored in the Blackboard agent so that it can be analysed.
- Log Analyzer Agent – Focused on analysing the logs and responsible for triggering alarms when it is necessary. It contains stored procedures to make sure unexpected behaviours are under control.
- Blackboard Agent – Focused on handling the information that will be displayed to the operators. It handles information synchronization since many agents are reading and writing into it.

The following image provides a view on how the whole process was built, connected and the relationships between the agents previously described, making it easier to understand the whole multi-agent architecture.

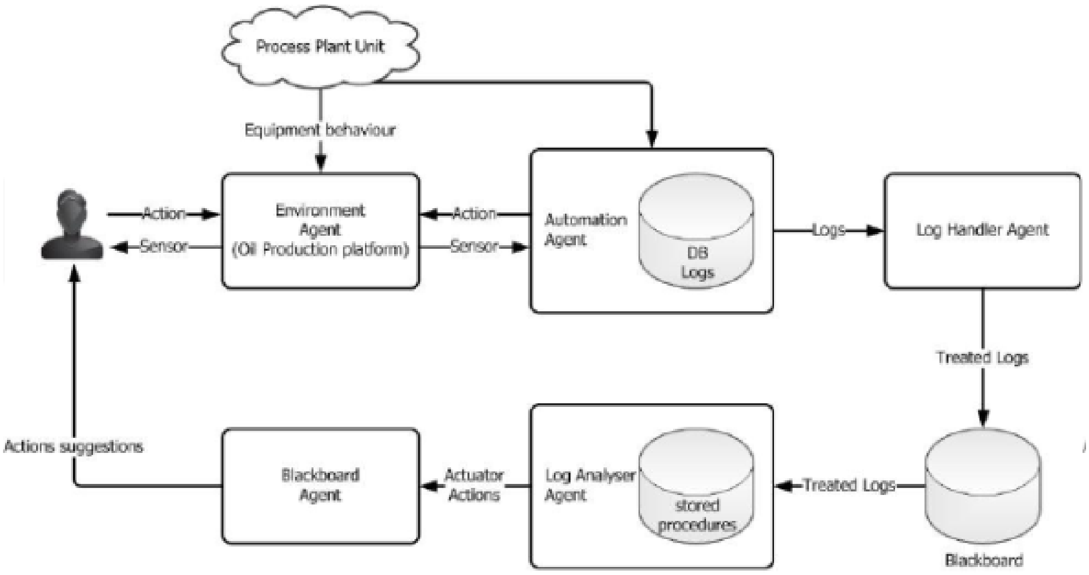


Figure 9 - Multi-Agent Architecture

The ontology was built as a model for the process plant domain. The alarm management system will be focused on certain main entities such as:

- Equipment – Component of a process plant unit.
- Actuator – Devices that control different equipment behaviours.
- Equipment Behaviour – Represents the way the equipment behaves in order to achieve the functionalities assigned to them.
- Event – Action over an actor that might cause a change in the behaviour of the alarm.
- Alarm – It describes the abnormal state of the equipment. They can fall into four categories, such as Very High (HH), High (H), Low (L) and very Low (LL). HH values can lead to an entire shutdown of the system.

- Sensor – Device that measures control variables. Two types of sensors are considered in this ontology – Analog and Alarm sensors.
- Control Variable Status – Indicates the measurement. A control variable status indicates for instance if the temperature is increasing.

The following image provides a view on how the entities are related and an understanding on how an ontology can be implemented on a system similar to this one.

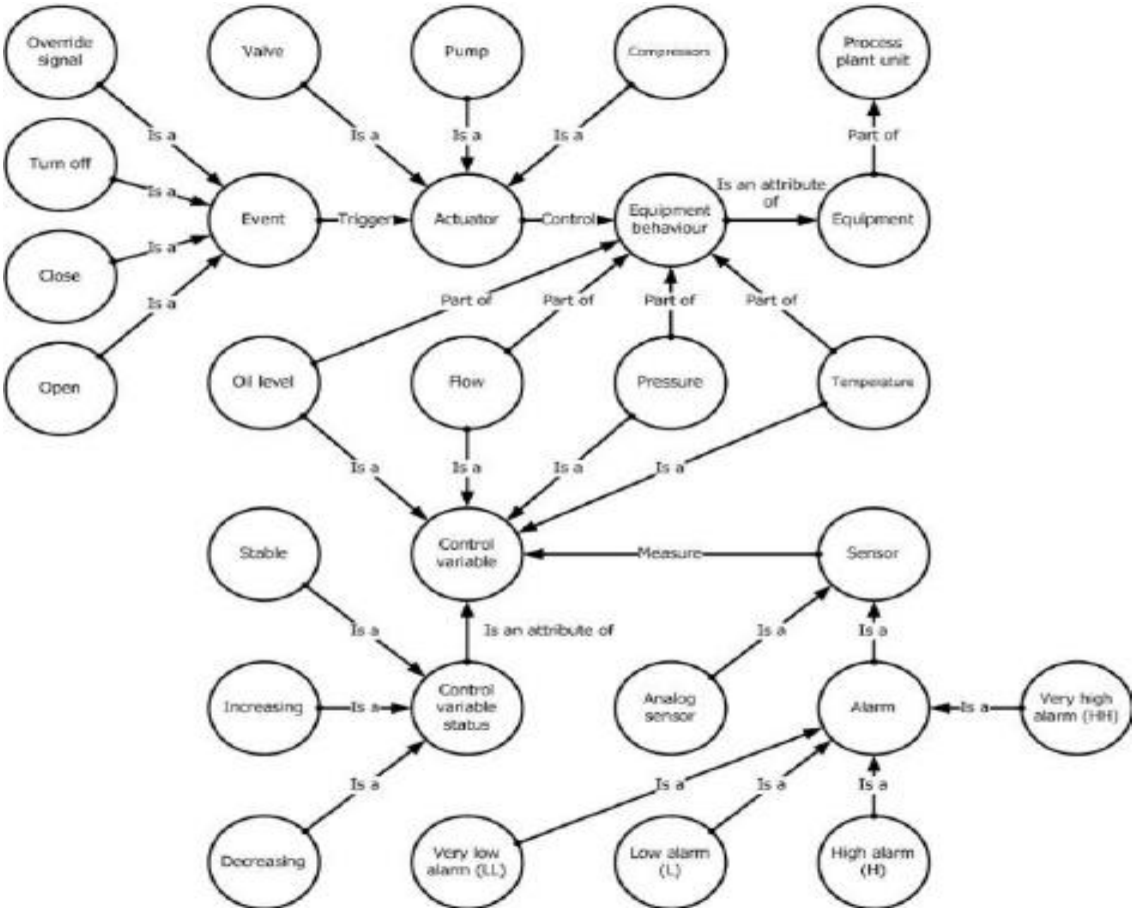


Figure 10 - Ontology for the Alarm Management System

Other approach of ontologies application based on the usage of alarm systems is, for example, in geographical context. The following paper (González & Marichal, 2005) is important to understand how it is possible to integrate already existing ontologies on a newer one by extension – one of the main advantages of using ontologies is its reusability, providing a considerable decrease of time and resources spent in the development of an application from scratch while at the same time providing more robustness and consistency. In this specific application, three dimensions were involved – sensors, geographic elements and alarms systems. The first two were added through the ontology integration functionality with the latter one being implemented entirely.

The chosen ontology to integrate the sensor part of the system was the SSN ontology with the integration of a System Capabilities Model. The module was proposed by W3C under the horizontal segmentation of SSN

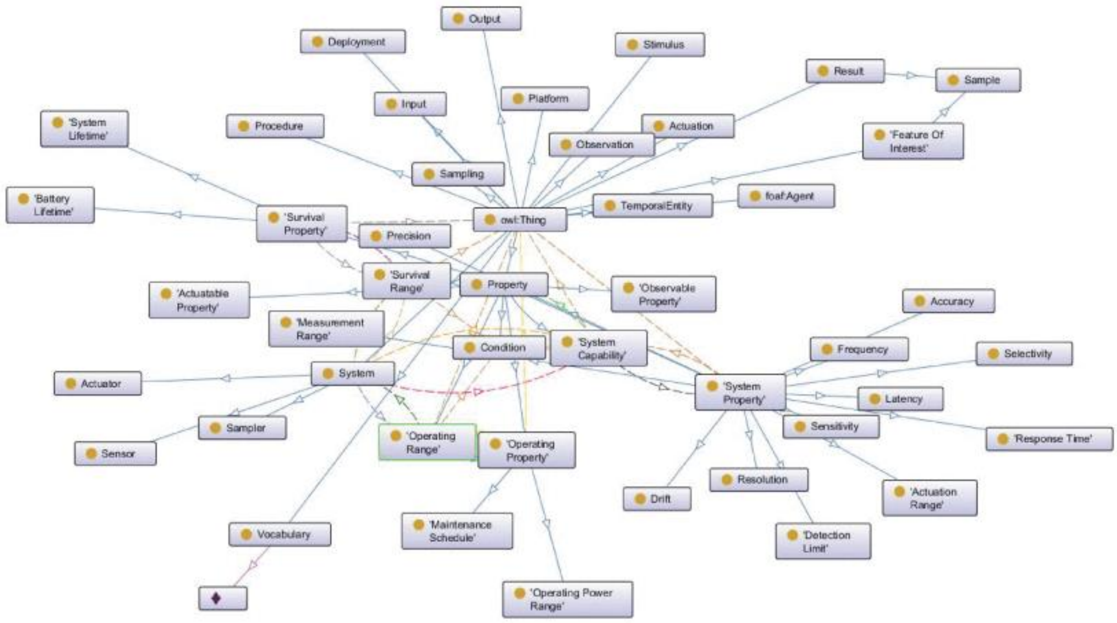


Figure 11 - Overview of the SSN ontology

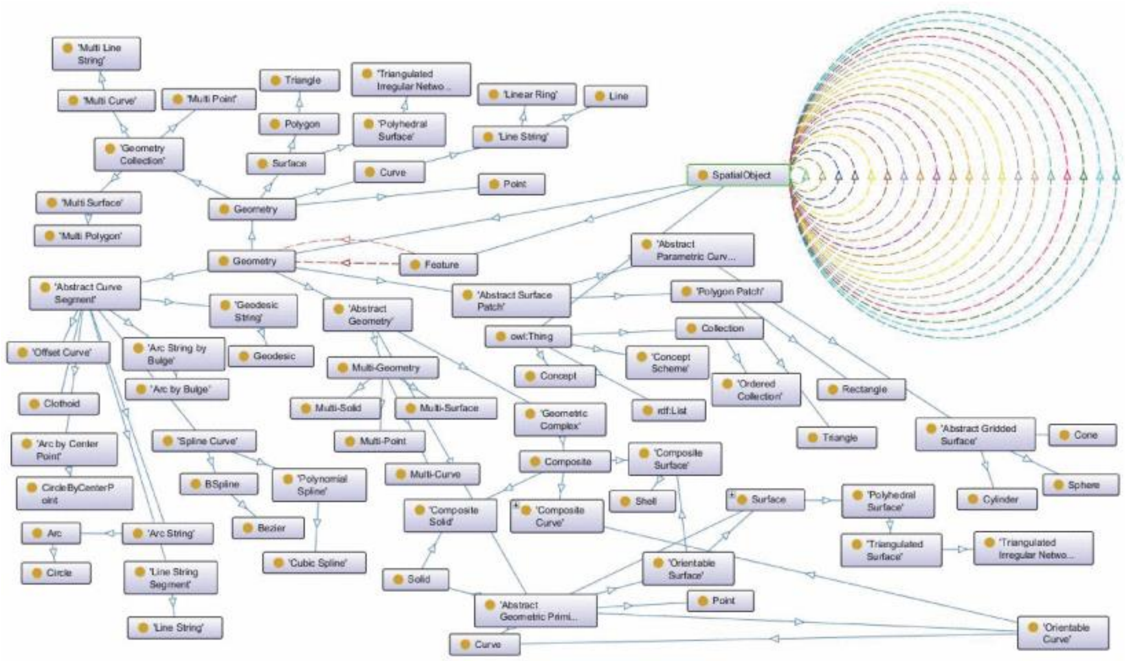


Figure 12 - Overview of the GeoSPARQL ontology

4 Rule-Based Systems

This section will be focused on explaining the importance residing in the development of a rule-based system in the context of alarm management in industrial facilities.

A rule-based system can be defined as a system used to apply human-made rules, mostly developed by experts in the area, in order to store, sort and manipulate data with the goal of replicating human intelligence on the matter (ThinkAutomation, 2017). Similar to other expert systems, rule-based systems have been increasing their popularity commercially, and even though they are defined as a special type of expert system, due to being dependent of a set of *if-then* rules, rule-based systems have a lot of areas where they can be applied on such as healthcare, transportation or security (Liu, et al., 2014).

The way this type of system works is simple. Firstly, it needs the injection of the data or the new business event in order to start the process. When the first step is done, the system will automatically start the analysis part that consists of conditionally processing the injected data against his set of rules. To finish the process the follow-up actions will react automatically based on the results of the second step. It is also important to refer that in order to work properly a rule-based system should include a reliable set of rules – also known as the Knowledge base – and a set of facts. A set of facts can be defined as set of confidence and forceful statements of fact or belief and a set of rules is usually defined as a group of actions that should be taken based on the presented problem, normally acting according to the set of facts provided, and as referred previously should always be represented in the *if-then* form (Engati, 2017).

Just like other expert systems, rule-based expert systems also have their own specific characteristics such as (Engati, 2017):

- Combining knowledge of human beings with expertise on the subject.
- Representing knowledge in a declarative way.
- Supporting the implementation of non-deterministic search and control strategies
- Being robust and having the possibility of operating using uncertain or incomplete knowledge.
- Helpful when it comes to rule-based decision making

4.1 Rule-Based Systems Main Components

Despite being easy to understand how the system is working, it is important to make sure that in the moment of creating a new rule-based expert system that all the components necessary to the proper functioning of the system are properly connected, are reliable and have been revised several times as a potential user will interact with it often. The structure of the newly created system should be, like referred previously, based on Knowledge Base (Rules) and a Working Memory (Facts) that will connect to an Inference Engine. The role of the Interference Engine is to link the rules that are defined on the knowledge base with the facts that are stored on the working memory and then it will decide which rule is satisfied by the supplied facts, prioritizing the rules and executing the highest priority rule. The Interference Engine will also be responsible for the communication with the Explanation Facility to know what the final result was so then it can be presented to the user through the User Interface (Panda, et al., 2012). The interaction between all the components can be observed on Figure 13.

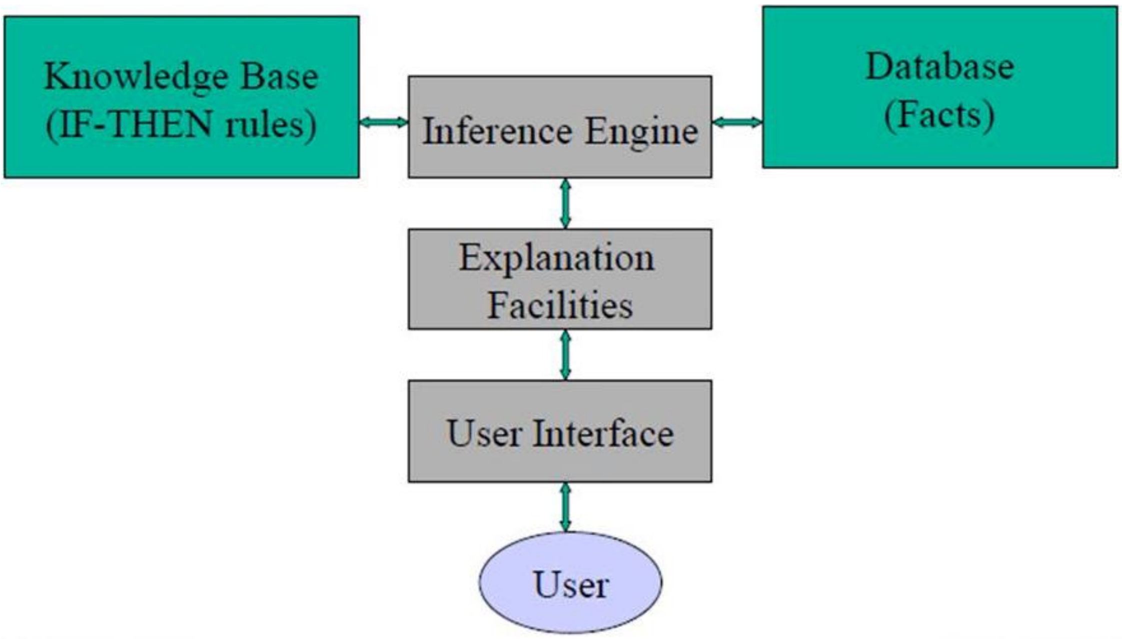


Figure 13 - Structure of a Rule-Based System (Williamson, 2015)

4.2 Rule-Based Systems vs Machine Learning

Machine Learning can be defined as a type of Artificial Intelligence that allows applications to be more productive, able to learn based on patterns and references and more accurate on predicting outcomes without using any direct instructions or being explicitly programmed to do so (Burns, 2018). Artificial Intelligence is normally portrayed as a field of study capable of developing smart machines with the final purpose of performing tasks similarly to humans (Velazquez, 2022). Deep Learning is a subset of machine learning in which neural networks with multiple layers learn from vast amount of data. The way these three concepts are related to each other is possible to view on Figure 14.

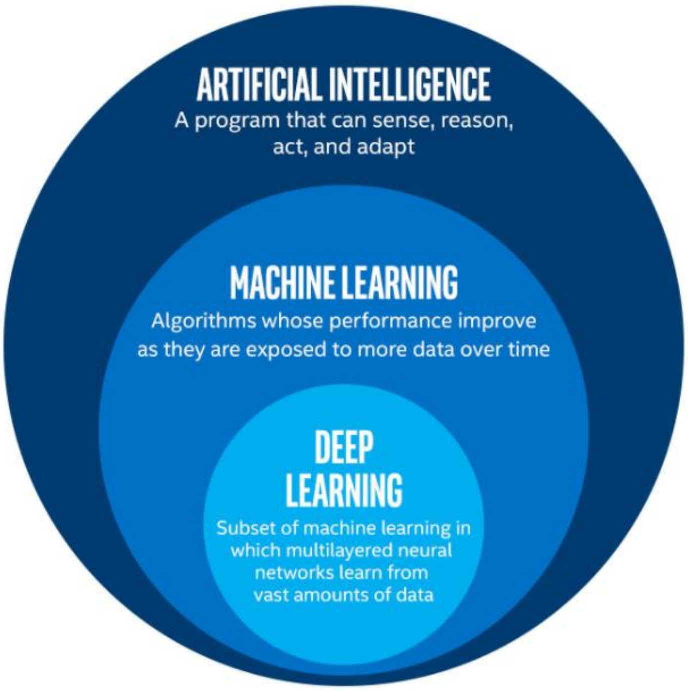

ARTIFICIAL INTELLIGENCE
A program that can sense, reason, act, and adapt
MACHINE LEARNING
Algorithms whose performance improve as they are exposed to more data over time
DEEP LEARNING
Subset of machine learning in which multilayered neural networks learn from vast amounts of data

Figure 14 - Artificial Intelligence, Machine Learning and Deep Learning relationship (tbyyf, 2022)

Machine Learning can be divided into four distinguished types which are Supervised Learning, Unsupervised Learning, Semi-Supervised Learning and Reinforcement Learning.

Supervised Learning focuses on labelling the data, making the algorithms learn to predict the input data based on a correct input previously known. Unsupervised Learning works as the opposite version of the Supervised Learning, working with non-labelled data, and with the respective algorithms learning to inherent structure exclusively from the input data (Brownlee, 2020).

Semi-Supervised Learning is often viewed as a middle point between the Supervised and Unsupervised Learning, using only part of the data labelled but keeping most of it non-labelled (Brownlee, 2020).

Reinforcement Learning is a special type of Machine Learning as it enables an agent to learn exclusively from the interactive environment it is in, by trial and error and using its own feedback from the experiences and actions it has been taking (Bhatt, 2018).

The usage of Machine Learning technologies comes with a lot of advantages as it helps the society building ways of modernizing the world and revolutionizing the areas of technology that influence our daily life, but the disadvantages of its application remind the human being that it is a new technology and that it has its problems and side-effects, being necessary to progress with caution (techvidvan, 2022).

The advantages of using Machine Learning reside on the fact that it does not need human intervention, achieving automation with relative ease at least for most of its process, meaning that it can reach conclusions faster than a human being based on the information that its being given to the algorithm. Secondly, it identifies with ease possible trends or and patterns through the review of large amount of data, that could have different sources, different types or different dimensions, in a way that would not be clear to humans. Also, Machine Learning algorithms are programmed in a way that they gain experience based on the number of iterations that they ran and on the number of data that they went through, keeping astonishing levels of improvement when it comes to accuracy and efficiency, helping them making better decisions overall on the long run. Lastly, the usage of Machine Learning has a wide number of applications and possibilities, being able to create personal experiences with ease and targeting the right path (DataFlair, 2021).

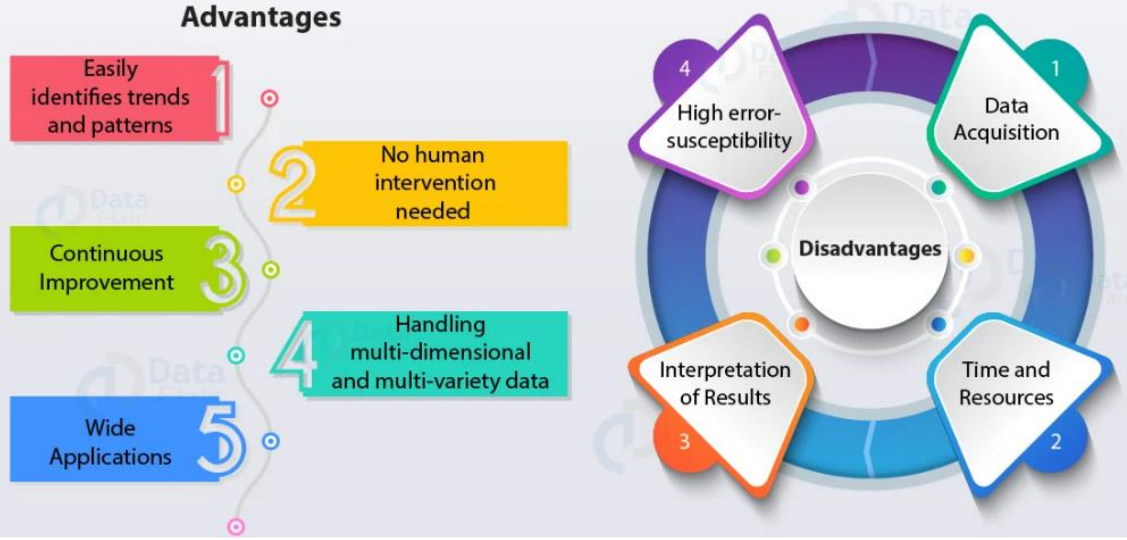

Advantages
Easily identifies trends and patterns
1
No human intervention needed
2
Continuous Improvement
3
Handling multi-dimensional and multi-variety data
4
Wide Applications
5
Disadvantages
1
Data Acquisition
2
Time and Resources
3
Interpretation of Results
4
High error-susceptibility

Figure 15 - Advantages and Disadvantages of using Machine Learning (DataFlair, 2021)

Despite all the advantages that using Machine Learning comes with, it is not perfect, also having some disadvantages that need to be acknowledged before making the decision of moving forward with the usage of this technology.

The disadvantages of using Machine Learning start with the fact that to work to its full potential, this type of technology needs a good pool of data, which means that the data must be of good quality to achieve the pretended results – a process that can take time and will always bring uncertainty. Another disadvantage of using ML is how expensive it is when it comes to time and resources – Machine Learning needs time in order for the algorithms to perform at the level they are expected to, reaching the desired levels of accuracy and relevancy and can be really expensive, monetary-wise, needing a lot of funding in order to develop high-level algorithms. Also, the interpretation of results is another major challenge regarding the usage of Machine Learning, being necessary to understand accurately the results that are being generated by the chosen algorithm in order to understand whether or not it is the best algorithm in the context it is being applied into. Lastly, when using ML there are a lot of high error-susceptibility situations, meaning that inaccuracies are a common occurrence when it comes to the development of algorithms due to the fact that they are developed by human beings, caused by mistakes included in the code that have a outcome on the veracity of the result (Sahu, 2020).

Just like Machine Learning, Rule-Based Expert Systems are included in the Artificial Intelligence category despite not sharing a connection with one another, as shown in Figure 16. Although both technologies are widely used in order to reach conclusions, based on large amounts of data, they differ in some ways that make them reasonably different (Carew, 2020). Machine Learning models are advised for projects or situations where there is a need for pure coding processing, pace of change and where simple guidelines do not apply while Rule-Based Expert Systems are used in situations where there is a need for speedy outputs and danger of error (Smith, 2020). A more detailed comparison between the two can be viewed in Table 1.

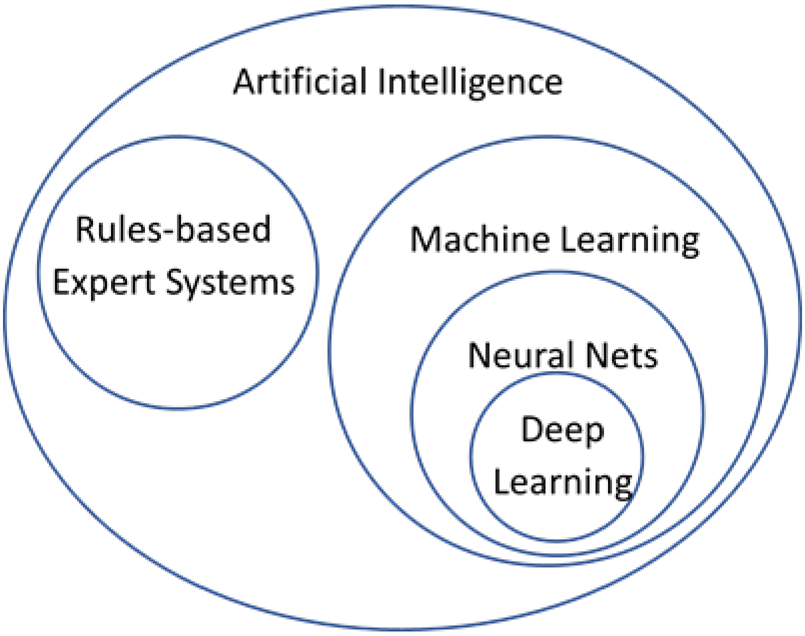


Figure 16 – Relationship between Rule-Based Systems and Machine Learning (Lessware, 2022)

Rule-Based Expert Systems	**Machine Learning**
Rule-Based Expert Systems models are deterministic.	Machine Learning models are probabilistic.
Rule-Based Expert Systems models are not scalable.	Machine Learning models are scalable.
Can work with basic data and information.	Requires more complexity and amount of data.
Rule-Based Expert Systems are immutable objects.	Machine Learning models are mutable objects.

Table 1 - Rule-Based Expert Systems versus Machine Learning (Swaminathan, 2020)

A rule-based system, as stated previously in this chapter, is normally defined as a system used to apply human-made rules developed by experts in order to store, sort and manipulate data with the goal of replicating human intelligence on the matter (ThinkAutomation, 2017). Just like Machine Learning systems, rule-based systems have its own list of advantages and disadvantages that should always be taken into account when developing a project of this magnitude.

Some of the advantages of using a rule-based system are:

- It is generally cost-efficient and accurate when it comes to results.
- The generated outputs are consistent due to being dependent on set rules, meaning the stability will be increased and the randomness levels will be lower.
- A rule-based system will always provide a high accuracy as it is based on pre-defined rules.
- Easily optimizable in terms of speed.

The disadvantages of using a rule-based system are:

- A rule-based system requires a lot of data, manual work and expertise regarding the subject.
- Writing and generating rules for a complex system is challenging and time consuming and most of the times the coverage for different circumstances will not be as developed as an evolving system.
- The self-learning capacity of a rule-based system is close to zero as it is generated purely based on the rules.
- Complex pattern identification is really challenging as it requires a lot of resources.

Based on this chapter, it is possible to understand that both Rule-Based systems and Machine Learning systems have positive and negative sides with sometimes being a good option to run both systems at the same time, especially when transitioning from Rule-Based to Machine Learning. Despite Machine Learning being the most promising when it comes to the long run, due to the fact that it has a higher potential due to its evolution not being directly dependent on the human being, it was decided that, In the context of this project, Rule-Based systems was the technology to be picked because it would make more sense to do so.

4.3 Rule-Based Systems on Industry 4.0

Rule-based systems are being applied more and more often in the context of Industry 4.0 and many different companies share a wide range of projects to help this vision turn into reality.

Bosch is a company from Germany that has a wide collection of products and services including home appliances, car services and industrial solutions. They are related to the development of driving and controlling solutions, energy, large thermal plants solutions and industrial software solutions (Bosch, 2022). The idea of more transparency, traceability and process optimizations to reduce costs to increase efficiency are the expectations for this company and they do believe that rule-based analysis and the processing of production data with the help of a software designed for that effect is the way to go. The developed software started with a series of interviews with experts on the subject such as production planners, setters, team leaders and plant workers and the obtained knowledge from those interactions helped defining the parameters and rules that were going to be decisive in supporting the manufacturing the process of Electrical Discharge Machines, also known as EDM. In daily operation, the software works as follows:

- Data is transferred from the machines and relevant systems to the software
- The rules monitor processes the data
- The software recognizes whether any limit was compromised or not, generating a warning in positive cases.
- Notifications are sent to the respective workers responsible for each line.
- Machines are adjusted based on the notifications provided
- The manufacturing expert team can always work towards improving the set model in case any malfunction happens.

The usage of rules-based systems in this company allows for the workers to detect potential malfunctions at an early stage so a reaction to it can be faster and more prepared in order to prevent production downtime and loss of quality control. Bosch assures that a knowledge system continually improves the quality standards and increases manufacturing transparency, being flexible and always dependent on the experience of the employees (Eisenbart, 2022).

The next example of rule-based systems applied in the context of Industry 4.0 comes from a research at the University of Manchester that managed to develop a framework based on a rule-based system with the purpose of demonstrating how can they influence decision making positively and benefiting the transition to Industry 4.0 (Kourtis, et al., 2020). The developed framework will have as its main objectives:

- Describe the structure of production knowledge, focusing on scheduling and material requirements planning.
- Be able to model the temporal progression of the production.
- Practical and easy to implement.

In the context of this project, experts develop a model of production, together with a knowledge base, planning and scheduling software. The knowledge base will play the main role in the value chain of this project as it will be responsible for arranging the necessary actions based on the set rules defined by the expert team when a decision must be made. The point of contact between the other components of the system and the knowledge base is the controller which will be responsible of updating the knowledge base and taking into consideration the information in it whenever it is necessary. On figure 17 it is possible to understand how the production is managed by the expert team and on figure 18, it is shown how the production is centered around the knowledge base also known as controller.

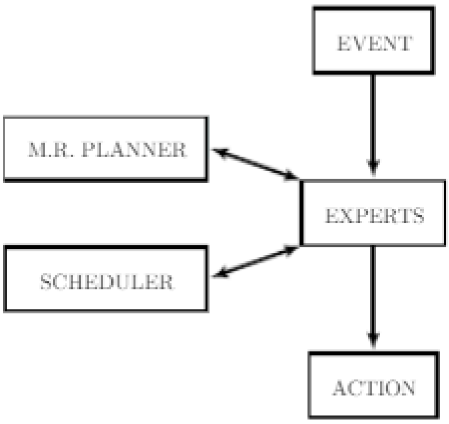


Figure 17 - Production managed by the expert team

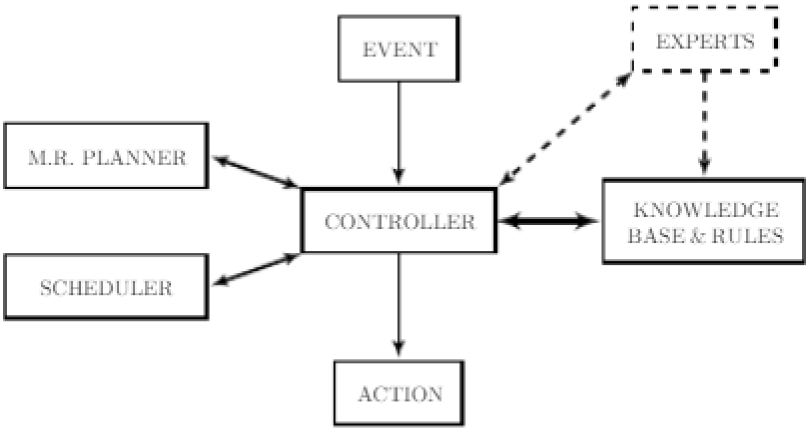


Figure 18 - Production around the Controller

With the end of the project, it was concluded that it is possible that a system, which has as its foundation the domain knowledge of a group of experts and a set of defined rules also known as a rule-based system, with emphasis on production scheduling, material requirements and productivity to be crucial in modern manufacturing enterprises. The biggest hold back when choosing a rule-based system is the fact that, according to the authors, in modern production it is disadvantageous to only be able to provide definitive statements about a different series of events that could happen at any time. A solution to this problem would be to add a machine learning layer to be able to deal with uncertain reasoning and learn with it.

4.4 Rule-Based Systems Software

In order to implement a rule-based system there are a lot of steps that must be done and a lot of variables that need to be taken into consideration such as the veracity of the set of facts, the set of rules and the quality of the database that is going to be used, but, just as important as the expert system, is the software that is going to be used in helping or fully developing the system. There are many different types of software that can be used in the context of this project, each one with different possibilities, depth and values.

4.4.1 DecisionRules

Launched in late 2021, DecisionRules software works based on the premise of speeding up decision making in the context of e-commerce, logistics, health and many others (DecisionRules, 2021).

DecisionRules provides, when it comes to a business perspective, the possibility of sharing and editing with ease the rule flows created by different users with different permissions. It is Excel-compatible, enabling the import and export of excel sheets, and allowing the development of a codeless approach solution (DecisionRules, 2021).

In terms of technology, the DecisionRules software provides a Seamless integration with already existing applications, whether it is a JavaScript, NodeJS, Angular or a Python tool, it is ready to handle both frontend and backend systems integration and provides the possibility of Kafka Connector.

Despite the fact that DecisionRules seems like a great asset to include in the context of this project, it is important to refer that it is not fully Open-Source, making it difficult to use it to its full potential as Excel importation and Exportation and API Management (The ability to manage rules and spaces via an external API) is limited. It is enough, though, to understand and get an idea of how good it can be in the scenario where there are no restrictions regarding software usage.

4.4.2 Drools

Drools is an Open-Source business rule management system (BRMS) with a forward chaining inference-based rules engine that processes facts and produces an output as a result of the rules and facts considered. Drools supports the Java Rule Engine API (JSR94) standard for its business rule engine and enterprise framework for the construction, maintenance, and enforcement of business policies in an organization, application or service (baeldung, 2022).

Drools is a rules engine implementation ready for the Java language but is able to run in both Java and .NET. It is also designed to allow pluggable language implementations with the rules being able to be written in Java, MVEL, Python and Groovy. Drools also provides for declarative programming and is flexible in a way that it matches the semantics of the problem with Domain Specific Languages (DSL) via XML (Drools, 2012).

There are many key advantages of using Drools as your bussiness rule management system, such as (Paragyte Technologies, 2017):

- The fact that rules are defined in a declarative way, which is easier when compared to application code.
- Speed and scalability.
- The business logics and rules are centralized, being easy to maintain, enhance or update the rule set.
- Having many options of tool integration.
- The fact that the rules are very close to natural language.

Obviously, Drools, as any other tool, has its own downsides and in this specific case is the fact that when compared to newer tools it shows a bigger challenge utilizing it when it comes to the ability level required to use it and it is not as intuitive as more recent software.

4.4.3 Red Hat Decision Manager

Red Hat Decision Manager is a platform for developing containerized microservices and applications that automate business decisions. Red Hat Decision Manager provides many possibilities when it comes to modelling business decisions and has support available for Decision Model & Notation (DMN), which is a standard for decision modelling. Decision Manager allows the organizations to incorporate an advanced and developed decision logic and maintain and evolve the established rules as the market shifts towards a different direction. It supports complex event processing, and resources that focus on solving complex scheduling and planning issues through optimization engines (Red Hat Decision Manager, 2022).

4.4.4 Azure Logic Applications and Azure Functions

Microsoft Azure, developed by Microsoft is a group of services that help organizations create, manage and implement applications on a larger scale, providing a solid range of cloud services that include computing, analytics, storage, networking and IoT services. Microsoft Azure is one of the most known public cloud server providers worldwide, helping businesses manage challenges and reach their organizational goals (Bigelow, 2022). It offers software as a Service (SaaS), Infrastructure as a Service (IaaS) and Platform as a service (PaaS), inclusively providing support for many different programming languages (McCoy, 2022).

Despite the fact that Azure does not have a dedicated tool focused exclusively on creating a Rule-Based System, the combination of Azure Logic Applications and Azure Functions can have the same role as one.

Azure Logic Applications is a cloud-based application that can be used to build automated workflows for a business and can be integrated with relative ease with other services that are part of Azure (Microsoft, 2022). Azure Logic Applications allows the creation of workflows that can be built with or without coding, providing an appealing visual to the flow created. It has auto-scale capabilities and in terms of security it shows promise and quality as all the data during a logic run is encrypted under TLS (Transport Layer Security).

Azure Function Applications, just like Azure Logic Applications, allows users to build their workflows but works using event triggers that will eventually initiate the flows created. It supports many programming languages, including C#, Javascript, Python and Typescript and the main focus of the service is to create a working environment where it is possible to do a great workflow without having a high skill level when it comes to coding but keeping a good and informative output (Microsoft, 2022).

Despite the fact that Azure Logic Applications and Azure Function Applications have some similar approaches when it comes to the way they work and what are the goals for the applications, they still have some differences that are worth noting in order to understand why they complete one another and how they can be used in the context of this project. In the following table the main differences between Azure Logic Applications and Azure Function Applications are described (LogicV, 2021).

Azure Function Applications	**Azure Logic Applications**
Requires Coding	Requires Visual Design
Developed and run locally on any self-managed server or in the cloud	Built and run inside Azure
Limited number of “bindings” as it works using triggers, input and output bindings	Can connect to both cloud and on-premises applications using SaaS applications
Each activity is an Azure function requiring code for activity functions	Large collection of already-made actions
In terms of management, it uses Azure Application Insights, an APM service for developers	In terms of management, it uses Azure Portal and Azure Monitor Logs
In terms of scaling, is built based on the consumption model depending on the plan chosen for the business	Has autoscaling capabilities

Table 2 - Azure Function Applications versus Azure Logic Applications (Microsoft, 2022)

5 Value Analysis

The procedure of analyzing value is a mandatory, yet complicated, part of the development of a new product or service as the value is easily interpreted differently by different people. However, in order to identify and avoid unnecessary costs or poor investments, this process becomes a crucial part in any business.

This section will be divided in three parts, with the first one being the New Concept Development analysis, the second one the Value Proposition and the last one being the Business Model Canvas of the present project.

5.1 New Concept Development

Innovation is crucial in the continuous success of any organization, so those entities always seek for new ideas and innovative processes that can enhance their position. With the constant development of technology and its globalization, it makes it mandatory for companies to have a constant understanding of the market they are inserted in and be ready to adapt according to the new innovations. Usually, innovations emerge due to the dissatisfaction of the current state of art and the goal is always to increase the value of the company, even if it means changing the way the organization operates overall.

5.1.1 Opportunity Identification

The introduction of new technologies, together with the constant increase of accessible information, creates numerous vulnerabilities so it makes It necessary that there is a way to ensure the safety of the information. On established companies with numerous responsibilities, it is safe to say that the quantity of alarms is huge with many different varieties among them which can bring dangerous situations and consequently problems to the company. It is safe to say that many companies are trying to enter the digital era and based on this information and on the fact that most companies have trouble dealing the constant daily problems, it is certain that an alarm manager would be necessary on this type of situations.

5.1.2 Opportunity Analysis

Now that the opportunity is identify, is it mandatory to justify and analyze it so it is possible to understand if the implementation process is advantageous or not. So, in the specific case of this project, it is necessary to understand if the companies feel like entering the digital era is positive and if not, what can be the reasons to not be worth it to do it. In a second stance, it is necessary to comprehend how alarm systems affect companies in general and how does an alarm management system differ from other alarms in a company's perspective.

In an article written by PTC, an American software company, it is estimated that forty percent of companies' technology spendings are spent in digital transformations, being essential in the point of view of the companies to create opportunities and to stay relevant (White, 2019). Nowadays in companies' eyes it is mandatory to have digital transformation going on inside the company or at least the idea of doing one to stay competitive.

In terms of using alarm management systems companies feel like it is positive to use them as they provide easy access to data, can store significantly more data, they are easy to maintain and are more cost-effective. So, in a global opinion alarm management can give you an edge when compared to other companies as it will not only upgrade the safety of the information regarding the company as it will also make it more competitive (Black, 2015).

5.1.3 Idea Generation

With the opportunity being analyzed and proved to be worth of investment, it is now up to discussion and idea generation. In order to conclude the best approach to this project that can fulfill the requirements it is up to the student and the project supervisor to debate this subject.

After the discussion of the approach, the idea generation process ended up deciding that the project should:

- Develop a new customized alarm system.
- Use an already existent alarm system.
- Develop a new alarm system using existing technologies.

5.1.4 Idea Selection

Now that the ideas are identified it is time to understand which one should be implemented. As it is a crucial choice to project, the method that was chosen to make the decision was a Multi-criterion decision analysis method. The two most known methods inserted in this category are TOPSIS, Technique for the Order of Prioritization by Similarity to Ideal Solution, and AHP, Analytic hierarchy process. In the case of TOPSIS, it will take the alternatives and the chosen criterion to generate a decision matrix. The biggest advantage between TOPSIS when compared to AHP is that the calculation in the first method takes into consideration all the alternatives are the same time while the second one compares all the alternatives on pairs, meaning that the latter has a more complicated calculation part. However, the fact that AHP considers every possibility makes it more reliable, and in the context of this project where only three ideas were generated justifies the usage of this method (Hussein, 2018).

The AHP method is divided in four different steps. The first one consists of developing a hierarchic decision tree where the initial problem is defined, followed by the criterion and the alternatives to make the final decision. In the Idea generation chapter, the alternatives were defined, and in order to make the final decision the criterions that are going to be followed are:

- Possibility of integrating with sensors from different sources.
- Capable of storing the information in safety.
- Capable of showing the user the alarm results properly in real time.

The first criterion, possibility of integrating with sensors from different sources, is crucial to the project as it can provide a more diverse alarm system ready for any type of situation. The second criterion, capacity of storing the information in safety, is also very important as it is crucial to make sure that the information is safe and that it is not compromised. The last criterion, capacity of showing the user the alarm results properly in real time is probably the most important part as the ability of the user being able to interact with the stored information is basically the core part of this whole project, the fact of it being real time adds even more depth to it as the user can have more time to react to certain problems that the system cannot solve by itself.

In the following figure an image regarding the AHP method hierarchic system is presented:

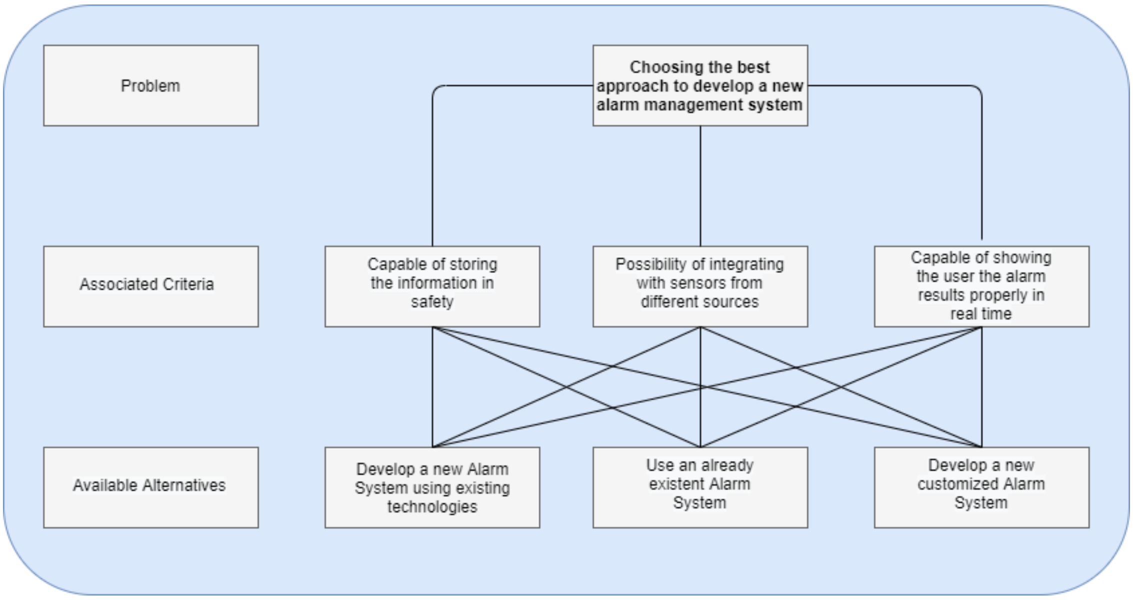

Problem
Choosing the best approach to develop a new alarm management system
Associated Criteria
Capable of storing the information in safety
Possibility of integrating with sensors from different sources
Capable of showing the user the alarm results properly in real time
Available Alternatives
Develop a new Alarm System using existing technologies
Use an already existent Alarm System
Develop a new customized Alarm System

Figure 19 - AHP Hierarchic structure

The second step of AHP is the comparison between alternatives and criterion, therefore a table with a comparison matrix between criterion was developed and to define priorities, Saaty scale was used.

The development of the table started taking into consideration the following:

- A – will be the criterion “Capable of storing the information in safety”.
- B – will be the criterion “Possibility of integrating with sensors from different sources”.
- C – will be the criterion “Capable of showing the user the alarms properly in real time”.

	A	B	C
A	1	0.5	0.25
B	2	1	0.5
C	4	2	1

Table 3 - Comparison matrix between criterions

By the observation of the previous table, it is possible to understand that the criterion C, capability of showing the user the alarms properly in real time is the most important part of the work, as it is the main objective of the project, but it is followed by criterion B, possibility of integrating with sensors from different sources, which is the second most important as the system should work with real data and not only with manipulated examples. The least important, even though is still very important to any project, is the criterion A, which consists of the capability of storing the information in safety. Even though this last one is very important it is only crucial to when the project is going to be commercialized and not on the testing phases.

The third step of AHP is to identify the relative priority of each criterion, by normalizing the comparison matrix. This process results in the following matrix called normalized matrix.

	A	B	C
A	1/7	1/9	1/7
B	2/7	2/9	4/14
C	4/7	2/3	4/7

Table 4 - Normalized matrix

After calculating the normalized matrix, it is necessary to calculate the priority vector by calculating the average between each line of the normalized matrix.

	Relative Priority
A	0.13
B	0.26
C	0.60

Table 5 - Priority Vector

The fourth step consists of evaluating the priority vector, for that it is necessary to calculate the Consistency Ratio.

Consistency Ratio is calculated by dividing the Consistency Index, IC, for the set of judgments by the Index for the corresponding random matrix, IA.

$$RC = \frac{IC}{IA}$$

The following table corresponds to the set of judgements by index corresponding to the random matrix, with the first line being the matrix dimension and the second line being the consistency index.

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
0.00	0.00	0.58	0.90	1.12	1.24	1.32	1.41	1.45	1.49	1.51	1.48	1.56	1.57	1.59

Table 6 - Random Matrix

With the matrix dimension being three, the IA value will be 0.58. In order to calculate IC, it will be necessary to use the following formula.

$$IC = \frac{(\lambda_{max} - n)}{(n-1)}$$

After that, It will be possible to multiply the comparison matrix with the priority vector, getting the value matrix which is in the following image:

A	0.41
B	0.82
C	1.64

Table 7 - Value Matrix

With this being done, it is now possible to calculate the λ_{max} value that can be done by the average of the values obtained in the previous matrix divided by the priority vector previously calculated.

$$\lambda_{max} = \frac{\frac{0.41}{0.13} + \frac{0.82}{0.26} + \frac{1.64}{0.60}}{3} = 3.01$$

With the λ_{max} calculated it is then time to calculate the IC value doing the following calculation:

$$IC = \frac{(\lambda_{max} - n)}{(n-1)} = \frac{(3.01 - 3)}{(3-1)} = 0{,}005$$

To finalize, it is now possible to use the consistency ratio formula:

$$RC = \frac{IC}{IA} = \frac{0.005}{0.58} = 0.0086$$

With 0.0086 < 0.1, it is possible to conclude that the priority vectors are consistent.

Now, it is possible to start doing the fifth step of the AHP method, developing comparison matrix for all the existing alternatives, analyzing each criterion as an individual matter. The considered alternatives are:

- A1 - Develop a new customized alarm system.
- B1 - Use an already existent alarm system.
- C1 - Develop a new alarm system using existing technologies.

	A1	B1	C1
A1	1	5	1.2
B1	0.2	1	0.25
C1	0.8	4	1

Table 8 - Comparison Matrix A Criterion

A1	0.5
B1	0.1
C1	0.4

Table 9 - Priority Vector A Criterion

	A1	B1	C1
A1	1	0.33	0.5
B1	3	1	1.2
C1	2	0.8	1

Table 10 - Comparison Matrix B Criterion

A1	0.17
B1	0.47
C1	0.36

Table 11 - Priority Vector B Criterion

	A1	B1	C1
A1	1	5	3
B1	0.2	1	0.5
C1	0.33	2	1

Table 12 - Comparison Matrix C Criterion

A1	0.67
B1	0.13
C1	0.20

Table 13 - Priority Vector C Criterion

The sixth step will consist of obtaining the composed priority for the alternatives, for that the priority vectors from the previous matrixes can be put together in a single matrix.

	A	B	C
A1	0.5	0.17	0.67
B1	0.1	0.47	0.13
C1	0.4	0.36	0.2

Table 14 - Alternative Relative Priority Matrix

By multiplying the matrix with the relative priority vector, it is possible to obtain the composed priority vector present on the following table:

A1	0.51
B1	0.21
C1	0.27

Table 15 - Composed Priority Vector

The last step is to conclude what is the best option. By observing the previous table is possible to understand that the development of a new alarm management system will be the best option.

5.2 Value proposition

Value proposition has a crucial role on the development and definition of a new project as it presents the business, showing its flaws and positive aspects and what is offered to the client, both the benefits and disadvantages on acquiring it.

In order to properly organize the ideas on process of the creation of value, there is a tool called *Value Proposition Canvas* normally used from small to established companies on its early stages.

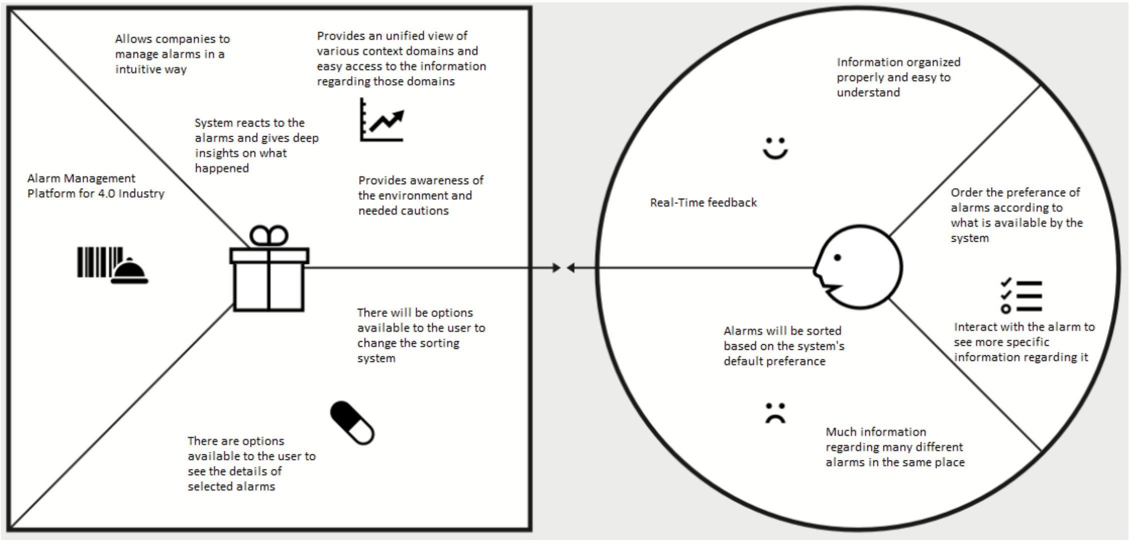

Allows companies to manage alarms in a intuitive way
Provides an unified view of various context domains and easy access to the information regarding those domains
System reacts to the alarms and gives deep insights on what happened
Alarm Management Platform for 4.0 Industry
Provides awareness of the environment and needed cautions
There will be options available to the user to change the sorting system
There are options available to the user to see the details of selected alarms
Information organized properly and easy to understand
Real-Time feedback
Order the preferance of alarms according to what is available by the system
Alarms will be sorted based on the system's default preferance
Interact with the alarm to see more specific information regarding it
Much information regarding many different alarms in the same place

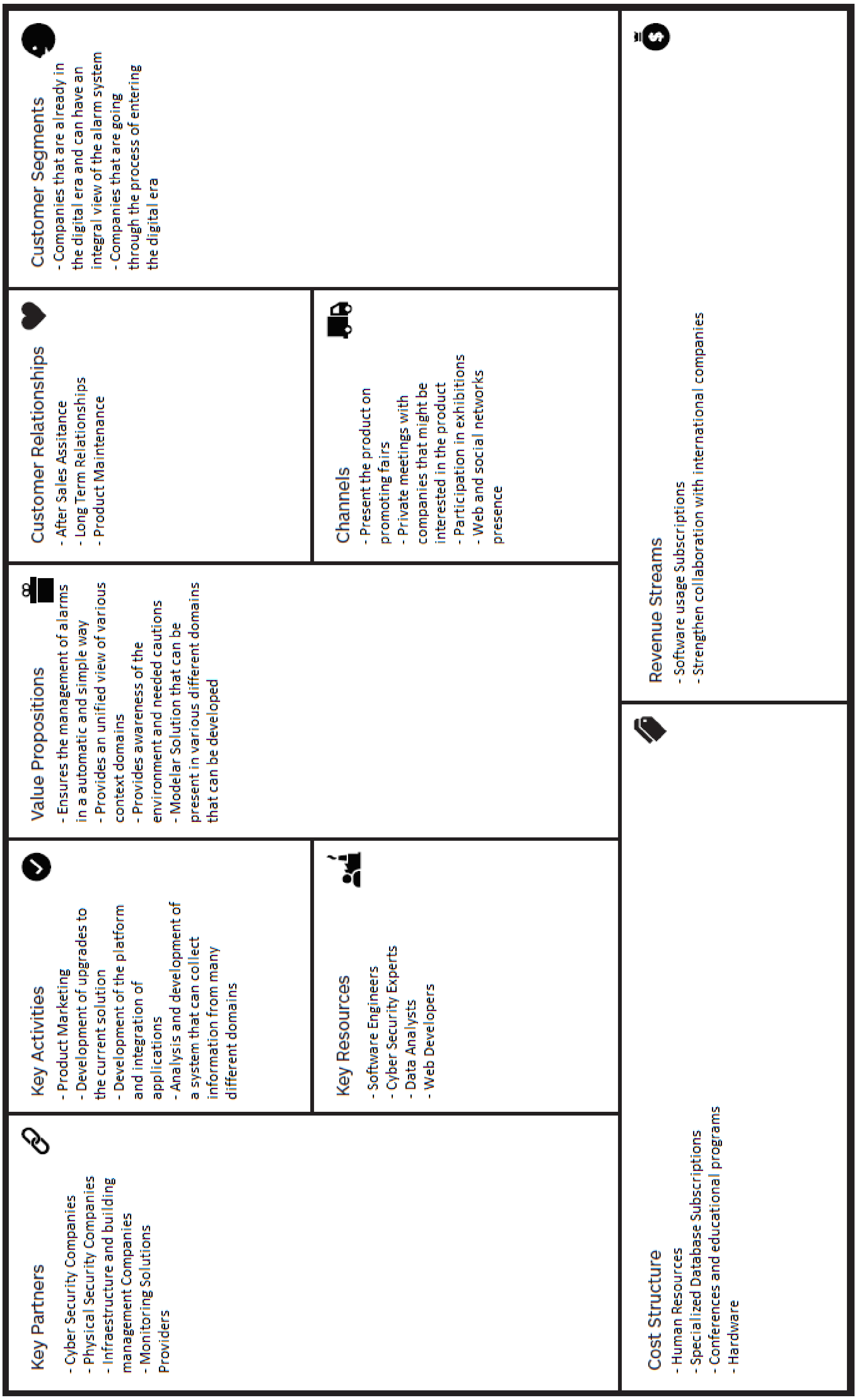
Figure 20 - Value Proposition Canvas

As it is possible to see, the image shows a clear connection between the product and the client, showing how it facilitates some processes in the client’s point of view.

The value proposition of this project consists of developing a product where the client can manage alarms with safety, simplicity and autonomy, allowing an easy understanding of how the product is working and a much faster decision-making.

5.3 Business Model Canvas

Business Model Canvas is defined as strategic management tool to define and communicate a new or an existing business model. The model should be simple in order to be easy to understand and should work through the fundamental elements of the business or product, structuring the idea in a coherent way.



Key Partners
- Cyber Security Companies
- Physical Security Companies
- Infraestructure and building management Companies
- Monitoring Solutions Providers

Key Activities
- Product Marketing
- Development of upgrades to the current solution
- Development of the platform and integration of applications
- Analysis and development of a system that can collect information from many different domains

Key Resources
- Software Engineers
- Cyber Security Experts
- Data Analysts
- Web Developers

Value Propositions
- Ensures the management of alarms in a automatic and simple way
- Provides an unified view of various context domains
- Provides awareness of the environment and needed cautions
- Modelar Solution that can be present in various different domains that can be developed

Customer Relationships
- After Sales Assitance
- Long Term Relationships
- Product Maintenance

Channels
- Present the product on promoting fairs
- Private meetings with companies that might be interested in the product
- Participation in exhibitions
- Web and social networks presence

Customer Segments
- Companies that are already in the digital era and can have an integral view of the alarm system
- Companies that are going through the process of entering the digital era

Cost Structure
- Human Resources
- Specialized Database Subscriptions
- Conferences and educational programs
- Hardware

Revenue Streams
- Software usage Subscriptions
- Strengthen collaboration with international companies

Figure 21 - Business Model Canvas

The Business Model Canvas provides a view with nine building blocks, each of them consisting in relevant information to the understanding of how the business is working. The nine blocks consist of:

- **Key Partners –** This block includes the key partners that will be crucial in accomplishing the product objectives. In this case, cyber security companies, physical security companies and infrastructure and building management companies are the key partners to be considered during the development of this project.
- **Key Activities –** This block includes the key activities that will take an important role on accomplishing the product roles. Product marketing, upgrade to the current solution and the development of the platform and integrations of related applications are the key activities to be considered.
- **Key Resources –** This block includes the key resources to achieve the established goals. Software engineers, cyber security experts, data analytics and web developers should work together in the context of this project.
- **Value Proposition –** This block describes the products or services that ensure value to the product. Guarantee of simplicity and automation on the management of alarms, awareness of the environment and needed cautions and modular solutions that can be present in various domains are examples of valuable services provided by the product.
- **Costumer Relationships –** This block includes the customer relationships that are relevant to the product in order to make it consistent and keep its clients. Therefore, after sales assistance, the development of long-term relationships and product maintenance are just examples of what is needed to attract the costumers to buy the product.
- **Channels –** This block describes the Channels to make the product a successful one. Promotion in fairs and exhibitions, web and social media presence and private meetings with companies that might be interested in the product are the best options in the context of this project.
- **Costumer Segments –** This block describes the customer segments of the project. Companies that are already in the digital era or companies that are through the process of entering it are the best options to fill this block.
- **Cost Structure –** This block represents the cost structure of the development of the project. Human resources, specialized database subscriptions, conferences and educational programs and the hardware are the costliest parts in the context of the project to be developed.
- **Revenue Streams –** This block represents the income of the project, with software usage subscriptions and strengthen collaboration with international companies being the best source of income.

6 Analysis

This section will focus on presenting the developed ontology and giving the reader an insight on the functional requirements using a use case diagram and non-functional requirements by presenting them using the FURPS+ model.

6.1 Ontology

As referred previously an ontology is defined as a representation of classes, designed in a way that humans and computers can communicate and share knowledge. In the context of this project, the development of an ontology is mandatory, and it was created in order to simplify the understanding of the project and the connections between concepts. The goal was to promote a unique way of understanding the proposed problem by all the parties involved. It is possible to see the developed ontology on the following figure.

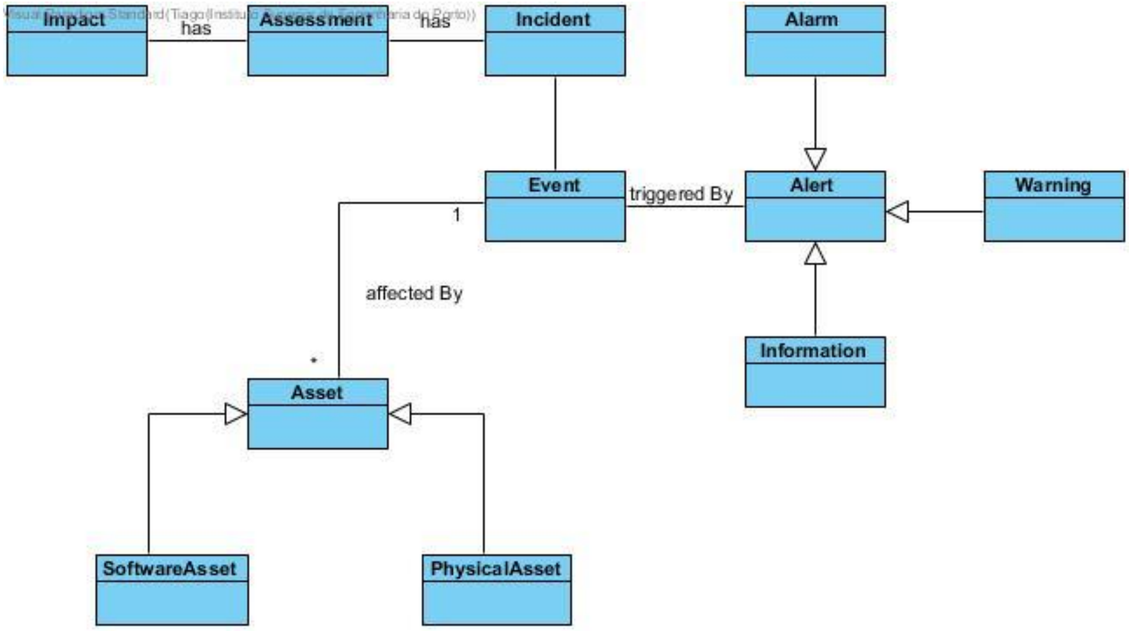


Figure 22 - Ontology

The concept classes represented in the previous figure have the following roles:

- **Event –** Can be defined as a change of status of an Asset or group of Assets, containing the information about a specific event. The event concept class affects the Asset class and triggered by the Alert class. It is related to an Incident.

- **Asset –** Resource that brings value to the organization involved. It is affected by the Event class and is divided into two sub classes: Software Asset and Physical Asset.
- **Alert –** The concept of alert is defined as a notification that one of components is not working properly. Its sub-classes are Alarm, Warning, and Information.
- **Incident –** Event that compromises an Asset. It has an impact level, which is described by its severity and completion level.
- **Assessment –** Contains specific information regarding the incident.
- **Impact –** Contains specific information regarding the impact level of the incident.

6.2 Functional Requirements

In order to present the functional requirements of the system, a use case diagram was developed. In the following figure it is possible to observe the use case diagram:

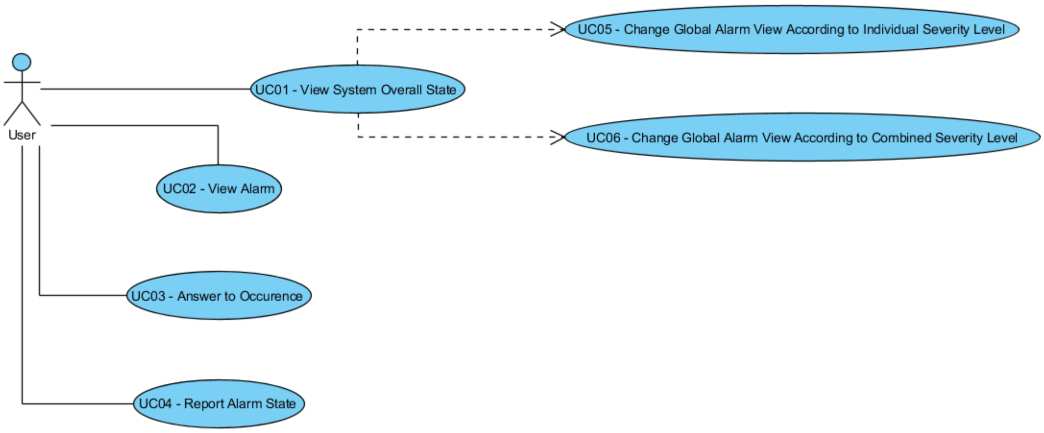

User
UC01 - View System Overall State
UC05 - Change Global Alarm View According to Individual Severity Level
UC06 - Change Global Alarm View According to Combined Severity Level
UC02 - View Alarm
UC03 - Answer to Occurence
UC04 - Report Alarm State

Figure 23 - Use Case Diagram

In the system there will only be one actor called User and will have the role of starting each use case. The use cases presented in the image above are described as follows:

- **UC01 – View System Overall State:** The user is presented with an overall view of the system, showcasing all the alarms based on default preference and summarized information regarding them, such as severity level or date of occurrence.
- **UC02 – View alarm:** The user is presented with a view of a specific alarm it will showcase all the alarm's specifications and give detailed information regarding it.
- **UC03 – Answer to Occurrence:** This use case will allow the user to react to an alarm according to various options based on the nature of the alarm by changing its state.

- **UC04 – Report Alarm State:** This use case will allow the user to report the state of a specific alarm in situations where the system is not behaving how it is supposed to, enabling a possibility of reacting to it.
- **UC05 – Change Global Alarm View According to Individual Severity Level:** This use case will change the view of the Overall State of the System by modifying the order of how the alarms are being presented individually. It will give preference to alarms on a higher or lower severity level, based on the option that the user selects.
- **UC06 – Change Global Alarm View According to Combined Severity Level:** This use case will change the view of the Overall State of the System by modifying the order of how the alarms are being presented based on their combined severity level. It will give preference to combinations of alarms on a higher or lower severity level, based on the option that the user selects.

Regarding UC03 and UC04, the first one will be focused on answering to occurrences based on the available scripted options given by the system according to the nature of the alarm, the latter will have as main focus reporting the alarm state in cases where the system is not behaving how it is supposed to and the User has to act and make sure that every other User has access to the information regarding that specific alarm.

The biggest differences between UC05 and UC06 is that the first one will view alarms and judge dangerous occurrences individually so each alarm will be evaluated individually based on its danger or time to react, while the latter will be focused on combined danger. For instance, there might be situations where an alarm can be dangerous if evaluated individually but can be less dangerous than two alarms that individually might not be as dangerous as it but combined show higher levels of danger.

Now that all the use cases are defined it is important to present how the use cases of higher importance are going to be structured. For that, it was decided to use *System Sequence Diagrams (SSD)* in order to present these specific use cases.

The following use case corresponds to UC03 – Answer to Occurrence. Basically, the user will have to join the system, which will automatically show the list of alarms available to him. Then, he will have to choose, from the list of alarms, the alarm which he wants to answer. It is important to notice that not every alarm will be able to be answered if the user does not have the power to respond to that specific kind of alarms. When the alarms specifications are entered, he will be presented with a lot of scripted reactions that the system can have to deal with that specific situation. A scripted reaction should be chosen with the knowledge available. The system will then inform the user that the operation was a success, and that the situation is solved.

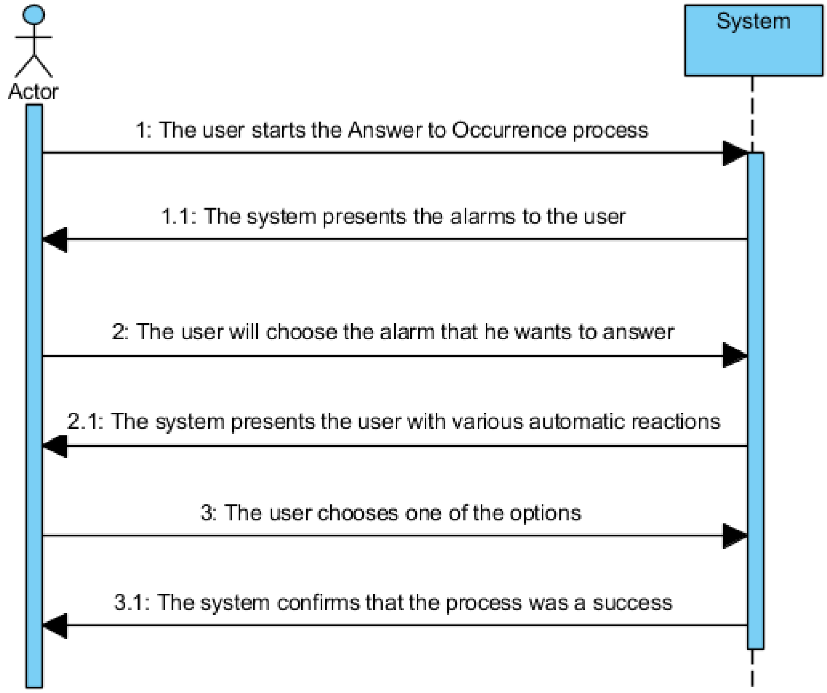

Actor
System
1: The user starts the Answer to Occurrence process
1.1: The system presents the alarms to the user
2: The user will choose the alarm that he wants to answer
2.1: The system presents the user with various automatic reactions
3: The user chooses one of the options
3.1: The system confirms that the process was a success

Figure 24 - UC03

The following use case corresponds to UC05 – Change Global Alarm View According to Combined Severity Level. It starts after the user joins the system and it offers the default view of the alarms. The user will then have the option to change how the alarms are shown by selecting the change view option. The system will then show the user all the options regarding the changing of view and the user will then be able to select the option that he wants. In this specific case the user will select the view according to combined severity level which will result on a change on how the alarms are viewed, ranked the combined severity level from higher to lower.

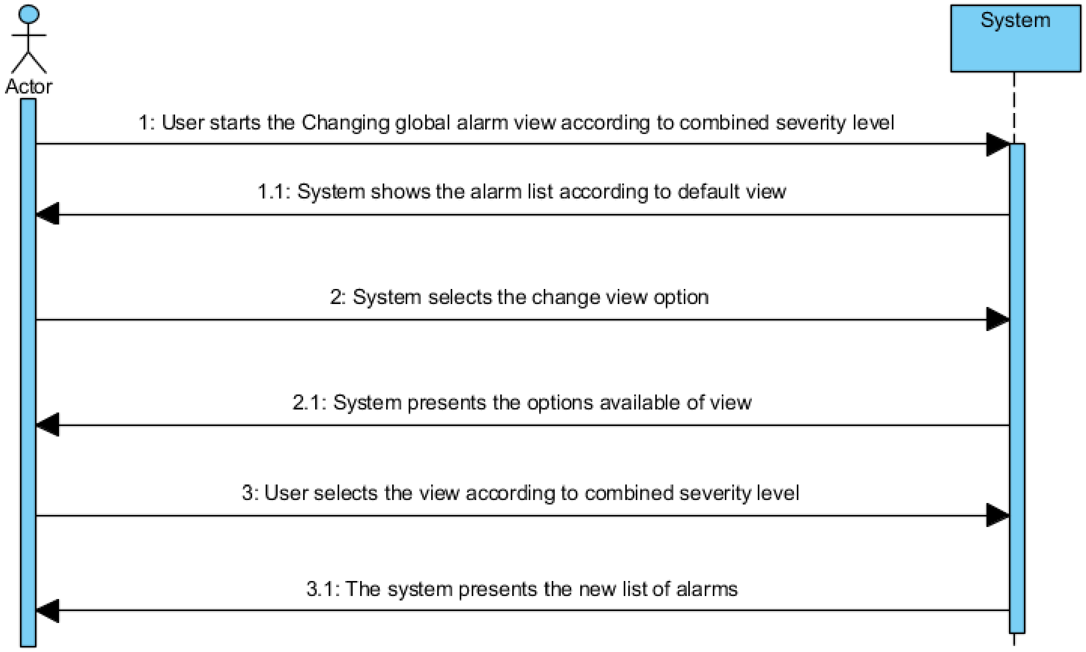

Actor
System
1: User starts the Changing global alarm view according to combined severity level
1.1: System shows the alarm list according to default view
2: System selects the change view option
2.1: System presents the options available of view
3: User selects the view according to combined severity level
3.1: The system presents the new list of alarms

Figure 25 - UC05

6.3 Nonfunctional Requirements

Nonfunctional requirements are crucial to a project as they define the way the system will execute certain functionalities even though they are not related to them. In order to present the nonfunctional requirements, it was decided to use the FURPS+ model, as it seemed to be the most viable option to showcase them.

FURPS+ can be defined as a technique to validate the prioritized aspects of the client's needs and necessities. It is an acronym for Functionality, Usability, Reliability, Performance and Supportability, and the + has the role of identifying other requirements, such as Design Restrictions, Implementation Restrictions, Interface Restrictions or even Physical Restrictions (Hyderabad, 2014).

This section is divided in other sub-sections that will include all the identified nonfunctional requirements.

6.3.1 Usability

The system should be easy to understand, and very intuitive so the user can easily react to what is happening and always understand what is going on the system to give a fast a better answer. To help that, the information regarding the alarms should be clear in order to ease the understanding and processing the information. The user interface that is going to be used in the context of this project should be easy to understand and use.

6.3.2 Reliability

The system should be reliable, its failure frequency should be very low, and the predictability of those failures should be close to perfection. In case of failure, its time length should be close to none in other to guarantee that the system is fully operational most of the time.

6.3.3 Supportability

The system should be able to run in any device.

6.3.4 Performance

Due to the nature of the project, it is mandatory that the response time is very low as it can be the difference between reacting to a chain of alarms properly or not. The user experience should not be compromised by server latency or any other causes.

6.3.5 Implementation Restrictions

The system should adopt good implementation methods so in potential cases of maintenance it eases the process of adapting the system to fulfill new requirements.

7 Design

Software Design focuses on creating a software artefact that has goals to fulfill and some rules and restrictions to follow.

This section will be divided in three subsections, presenting two different Design approaches of the problem, using once again UML notation. The third section will be focused on comparing both approaches and deciding which one should be chosen and used regarding advantages and disadvantages. The technologies being used in the subsections are described earlier in the document in the state of art section.

7.1 System overview using open source technologies

The following image presents the system overview from the point of view of using open-source technologies only.

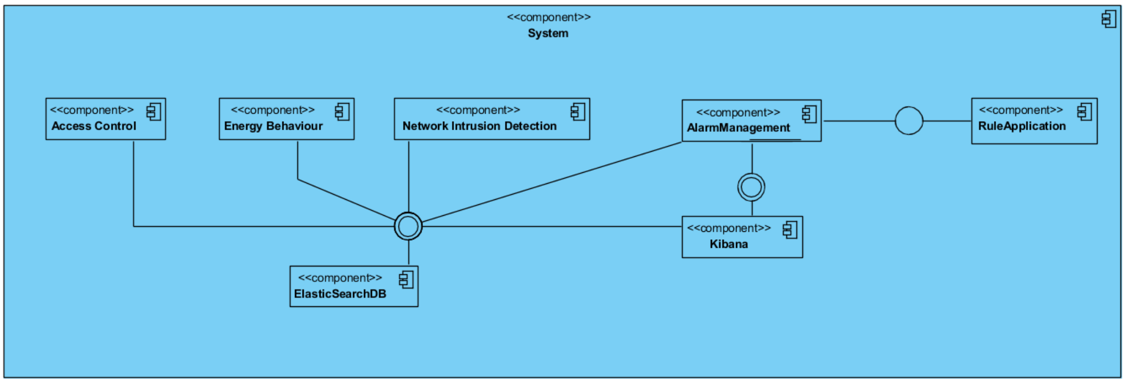


Figure 26 - Component Diagram of the system using open source technologies

The components described in the image have the following roles:

- **Access Control –** Will manage the accesses in a building to certain restricted areas and will send information in json format to database regarding the situation from time to time.
- **Energy Behavior –** Will manage the energy behavior of some buildings and structures and will send information in json format to database regarding the situation from time to time.
- **Network Intrusion Detection –** Will control potential network intrusions or attacks and will send information in json format to database regarding the situation from time to time.
- **Elastic Search Database –** Will store all the information regarding all the types of json formatted data so it can be analyzed later.

- **Kibana –** Kibana will serve as a support tool to the Alarm Management application. It will
- **Alarm Management –** Will have as main focus acting as the first contact between the user and the application.
- **Rule Application –** Will have as role the execution of the rule's module.

7.2 System overview using Azure technologies

The following image presents the system overview from the point of view of using azure technologies.

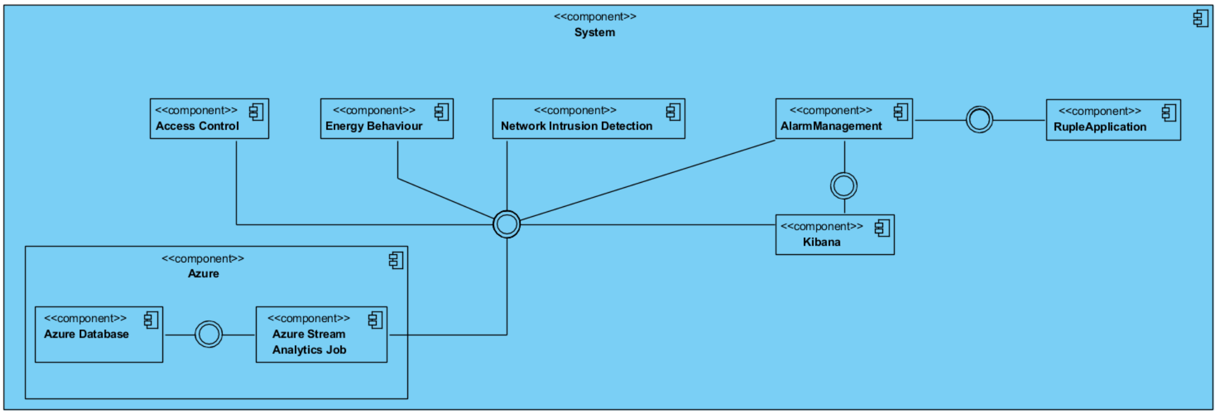


Figure 27 - Component Diagram of the system using Azure technologies

- **Access Control –** Will manage the accesses in a building to certain restricted areas and will send information in json format to database regarding the situation from time to time.
- **Energy Behavior –** Will manage the energy behavior of some buildings and structures and will send information in json format to database regarding the situation from time to time.
- **Network Intrusion Detection –** Will control potential network intrusions or attacks and will send information in json format to database regarding the situation from time to time.
- **Azure Database –** Will store all the information regarding all the types of json formatted data so it can be analyzed later.
- **Azure Stream Analytics job –** Will act as a linking system between Kibana and the Database.
- **Kibana –** Kibana will serve as a support tool to the Alarm Management application. It will

- **Alarm Management –** Will have as main focus acting as the first contact between the user and the application.
- **Rule Application –** Will have as role the execution of the rule’s module.

7.3 Comparison and Choice of Design

The azure technologies provide a security service provided by a third-party company, Microsoft, which will allow the system to have one less responsibility and focus the attention on other parts of itself, although it is important to say that using azure technologies can have some limitations unless the unlimited service is acquired which can be very expensive to some companies. It is safe to say that azure is the greater option here as it is a paid service with a lot of evidence of success, but the open-source option is free so it tends to be more appealing on the long run as it is also more flexible.

8 Implementation

This section will focus on describing the process of implementation that is divided in two sub-chapters:

- Ontology Implementation
- Expert System Implementation

As the name indicates, the Ontology Implementation part will focus exclusively on the ontology implementation and how it affects the overall system and the Expert System part will be focused on the expert system implementation.

8.1 Ontology Implementation

As previously described in the document, an ontology is often defined as a formal specification of classes, attributes, and the relationships between them, having the sole purpose of enabling knowledge sharing and reuse, therefore showing a great potential when it comes to the development of this project. In order to implement the idealized ontology, it is mandatory to use a selected software, which was Protégé, as described on the Ontologies chapter. The implemented ontology was named *Alarm Management Module Ontology* (AMMO). This part of the document will detail the implementation of the ontology, including the classes representing concepts, object properties and their relationships and the concept's attributes.

8.1.1 Classes

This section will focus on presenting the classes after implementing them on the Protégé software and will be displayed together with their specifications in order to make it easy to understand the context provided. The implemented ontology has the following classes:

- *AMMOThing* – The classes' implementation process starts with the definition of a local class that will be a direct subclass of the default *OWL:Thing* class that is automatically generated. The *AMMOThing* class will serve as a parent to all of the developed classes and will have as its main role establishing the relationships between classes and how they are grouped and distributed. The definition of the *AMMOThing* class and the introduction to its direct subclasses are provided in Figure 22.

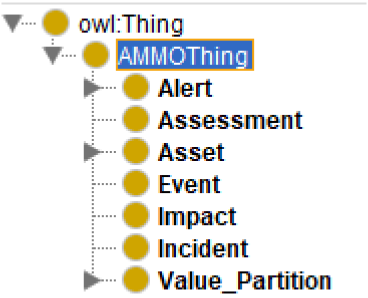


Figure 28 - Ontology overview

- *Alert* – The concept of *Alert* is defined as a notification that implies that certain protocols have been broken and therefore the system, or part of it, is not behaving as it is supposed to. It is divided into three subclasses:
    - ○ *Alarm – Defined as a specific type of Alert, usually referred to Alerts with a high level of severity.*
    - ○ *Information – Contains synthetized information regarding the Alert in order to make it easily understandable.*
    - ○ *Warning – Contains information regarding the procedures that should be taken in case of a specific alert, if it ever happened before, in order to ensure that the decision-making process when it comes to reacting to the alert is precise.*

  Information regarding the *Alert* class is provided visually in Figure 23.

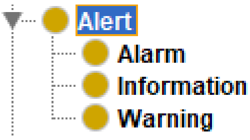


Figure 29 - Alert class overview

- *Assessment* – The *Assessment* class has as main role providing, in detail, everything that happened after an incident occurred, especially when it comes to the procedure route so basically if it was followed or if there was improvisation in the between. This class is important as it will allow for future incidents that are eventually similar to one that happened before being analyzed beforehand.
- *Asset* – The concept of *Asset* was introduced on the ontology with the purpose of representing resources that have an important role in the proper functioning of the general industry system that it is inserted into. It is divided into two subclasses – PhysicalAsset and SoftwareAsset – which, as the names indicate, have the purpose of distinguishing physical assets from software assets. The information regarding the *Asset* class is in Figure 24.

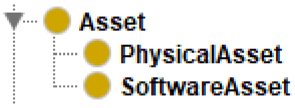


Figure 30 - Asset class overview

- *Event* – The *Event* class represents random events directly related to a change of criticality, severity or status of an *Alert* or *Alarm* or the change of status of at least an *Asset*. The events are generated by an occurring *Incident*.
- *Impact* – Defines the severity level of the impact that an occurring *Incident* had based on the *Assessment* that is related to it.
- *Incident* – The *Incident* class is introduced as the responsible for generating events from the *Event* class. It can generate more than one event and has a specific start time and end time. Every *Incident* is related to an *Assessment* related to it, making an evaluation and report of what happened in the specific *Incident*.
- *Value_Partition* – The *Value_Partition* class has the goal of maintaining as subclasses all the variables with set pre-defined values. They have the goal of restricting the number of pre-defined values and providing clear and objective values, so it is easier to understand. The subclasses related to *Value_Partition* are:
    - *Severity* – Provides the severity level and is divided into five categories: Very_Low, Low, Medium, High, Very_High.
    - *Criticality* – Provides the criticality level and is divided into five categories: No_Effect, Minor, Major, Emergency, Catastrophic.
    - *Status* – Provides the status and is divided into two categories: Enabled, Disabled.

  The information regarding the *Value_Partition* class is in Figure 25.

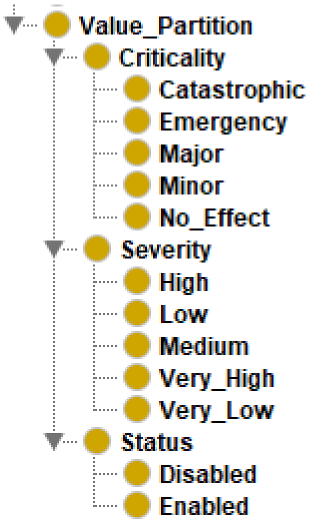


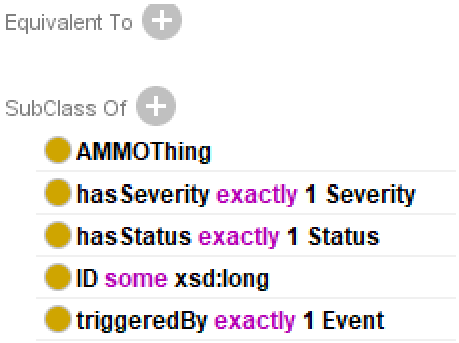
Figure 31 - Value_Partition class overview

All the classes referred in this section were added manually to the ontology and a synthesised explanation about them can be found on **Appendix A** in this document.

8.1.2 Properties

In OWL, properties are divided into two different types: *Object Properties* and *Data Properties*. The role of *Object Properties* is to set or add some restrictions regarding specific classes or varied relationships between those classes. OWL provides a default top-level property called *topObjectProperty* which can contain different groups of *Object Properties*, created by the user, and acts as a parent to all of these groups. In this project we did not create any group and defined the *Object Properties* of interest immediately as sub-properties of the *topObjectProperty*. *Data Properties* differ from the *Object Properties* as they are responsible of defining simple restrictions such as numbers or characters and are not object related. Just like *Object Properties*, *Data Properties* has a default top-level property called *topDataProperty* that allows the creation of different groups but once again it will not be necessary as all the *Data Properties* will be created as sub-properties of the default one.

- Alert – Is a subclass of AMMOThing and has as Object Properties a hasSeverity and hasStatus relationship with singular instances of Severity and Status respectively of the Value_Partition's class. It also has a triggeredBy relationship with one instance of the Event class. As Data Properties, the Alert class has ID. The properties of the Alert class can be viewed in Figure 26. The Alert class also includes the Alarm, Warning and Information classes but they all share the same properties of their parent class. Their relationship can be viewed in Figure 27.



Equivalent To

SubClass Of

- AMMOThing
- hasSeverity exactly 1 Severity
- hasStatus exactly 1 Status
- ID some xsd:long
- triggeredBy exactly 1 Event

Figure 32 - Alert Class Object and Data properties

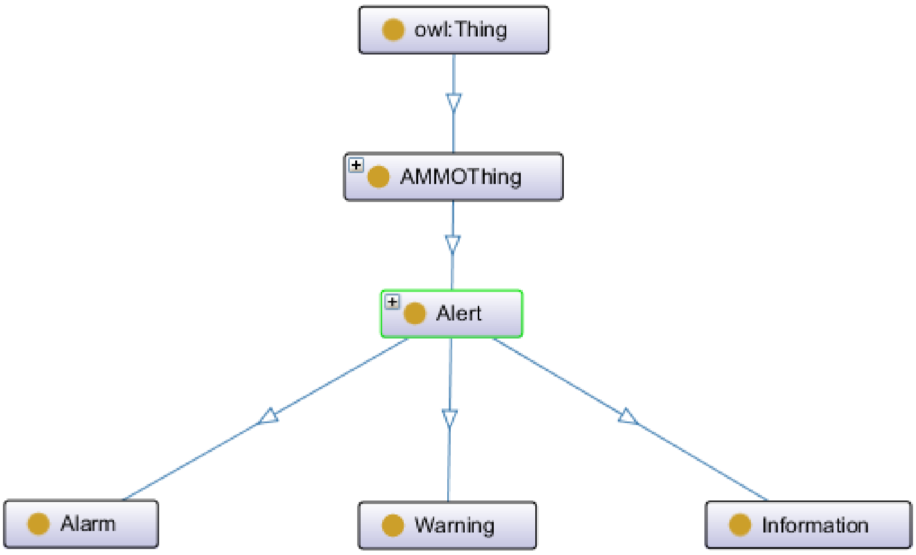
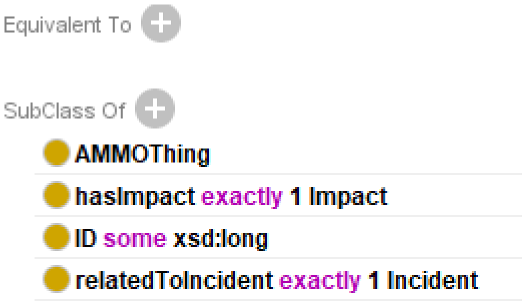


Figure 33 - OntoGraf of the Alert class

- Assessment – Is a subclass of AMMOThing and has as Object Properties a hasImpact relationship with exactly one instance of the Impact class and a relatedToIncident relationship with a singular instance of the Incident class. The Assessment class also has a Data Property, an ID, and it can be observed in Figure 28.



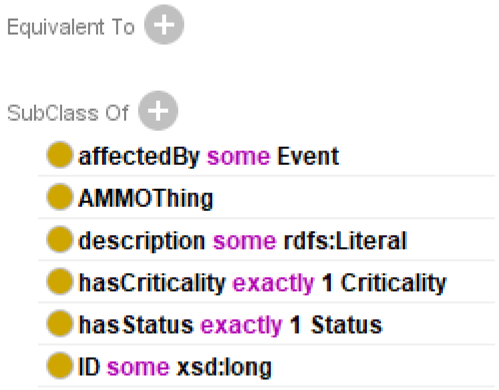
Equivalent To

SubClass Of

AMMOThing
hasImpact exactly 1 Impact
ID some xsd:long
relatedToIncident exactly 1 Incident

Figure 34 - Assessment Class Object and Data properties

- Asset – Is a subclass of AMMOThing and has as Object Properties hasCriticality and hasStatus relationship with singular instances of Criticality and Status respectively of the Value_Partition's class. In addition, it has an affectBy relationship with the Event class. As Data Properties, the Asset Class has an ID and a description. The properties of the Asset class can be viewed in Figure 29. The Asset class also includes the PhysicalAsset and the SoftwareAsset classes, but they all share the same properties of their parent class. Their relationship can be viewed in Figure 30.


Equivalent To

SubClass Of

affectedBy some Event

AMMOThing

description some rdfs:Literal

hasCriticality exactly 1 Criticality

hasStatus exactly 1 Status

ID some xsd:long

Figure 35 - Asset Class Object and Data properties

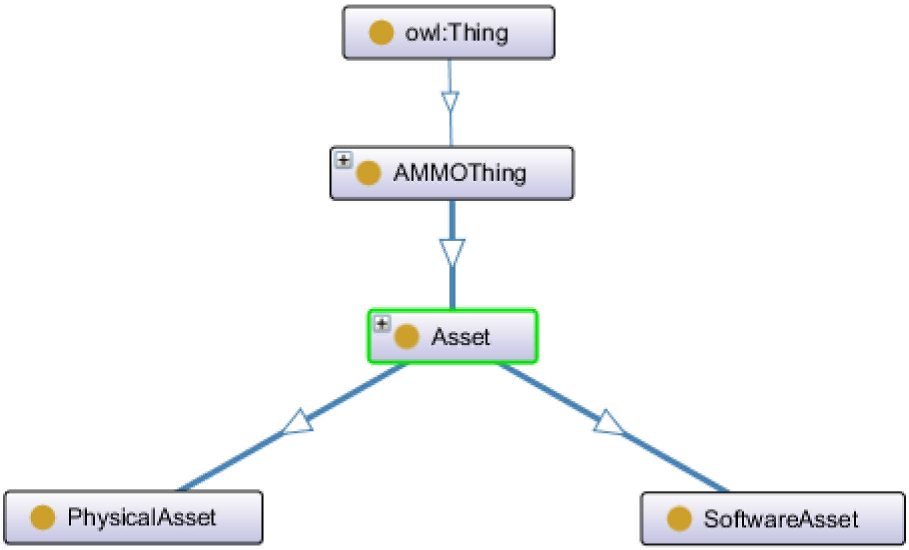
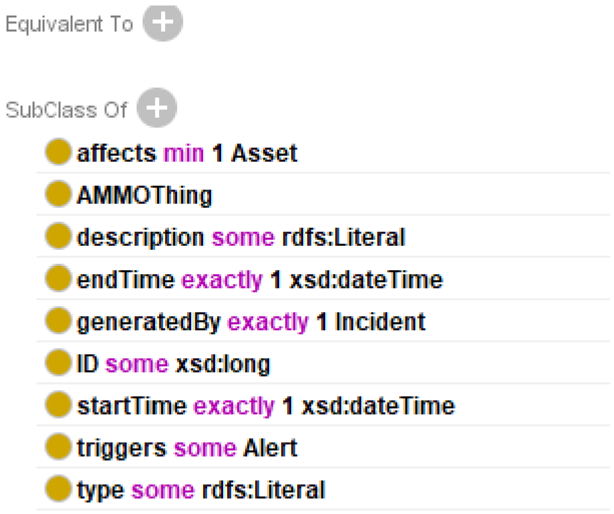


Figure 36 - OntoGraf of the Asset class

- Event – Is a subclass of AMMOThing and has as Object Properties a affects relationship with the Asset class, a triggers relationship with the Alert class and a generatedBy relationship with the Incident class. The class event also has some Data Properties, including an ID, a description, a type and a start and end time. The properties of the Event class can be seen in Figure 31.



affects min 1 Asset

AMMOThing

description some rdfs:Literal

endTime exactly 1 xsd:dateTime

generatedBy exactly 1 Incident

ID some xsd:long

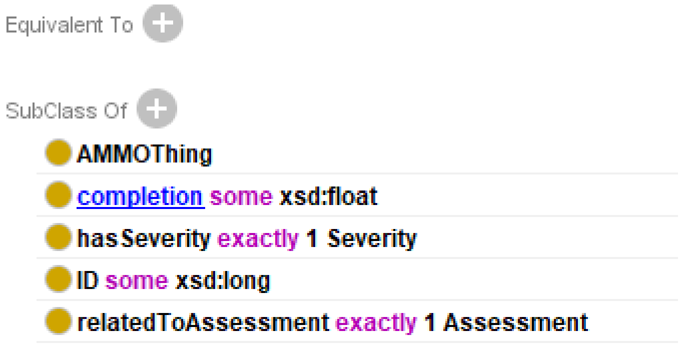
startTime exactly 1 xsd:dateTime

triggers some Alert

type some rdfs:Literal

Figure 37 - Event Class Object and Data properties

- Impact – Is a subclass of AMMOThing and has as Object Properties a hasSeverity relationship with a single instance of the Severity class included in the Value_Partition class. Also, it has a relatedToAssessment relationship with exactly one Assessment instance. When it comes to Data Properties, the Impact class has an ID and competition property. The properties of the Impact class can be seen in Figure 32.



AMMOThing

completion some xsd:float

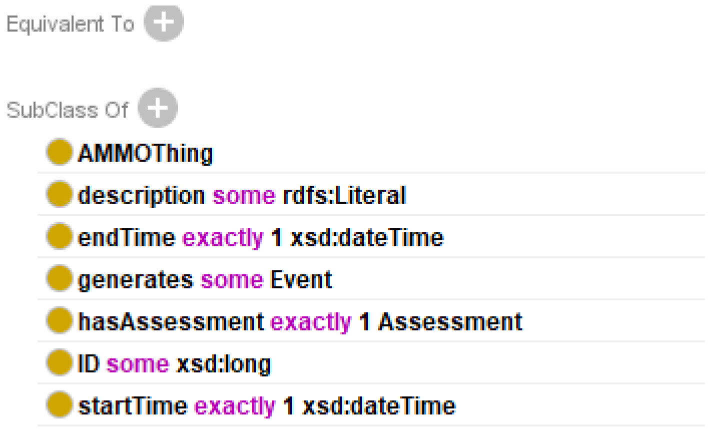
hasSeverity exactly 1 Severity

ID some xsd:long

relatedToAssessment exactly 1 Assessment

Figure 38 - Impact Class Object and Data properties

- Incident – Is a subclass of AMMOThing and has as Object Properties a hasAssessment relationship with a single instance of the Assessment class. The Incident class also has a generated relationship with the Event class. When it comes to Data properties, the Incident class has an ID, a description and a start and end time. The properties of the Incident class can be seen in Figure 33.



Equivalent To

SubClass Of

● AMMOThing
● description some rdfs:Literal
● endTime exactly 1 xsd:dateTime
● generates some Event
● hasAssessment exactly 1 Assessment
● ID some xsd:long
● startTime exactly 1 xsd:dateTime

Figure 39 - Incident Class Object and Data properties

- Value_Partition - The Value_Partition class is a subclass of AMMOThing and does not have any Object Properties or Data Properties. Its sole purpose is to serve as a parent to other classes including the Severity, Criticality and Status classes but they all share the same properties of their parent class. Their relationship can be viewed in Figure 34.

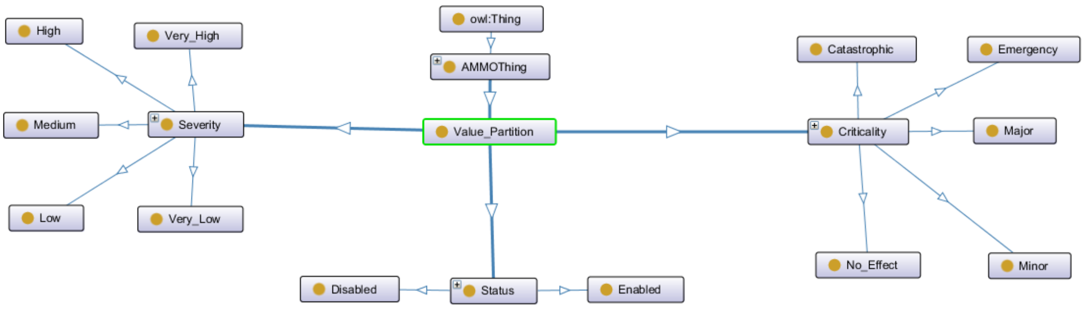
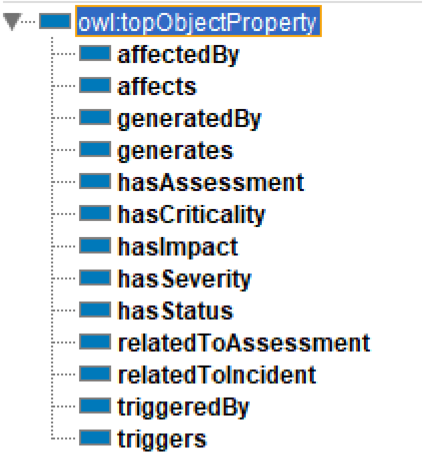


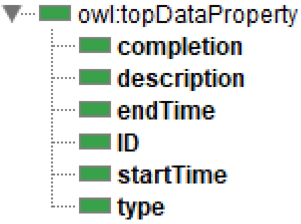
Figure 40 - OntoGraf of the Value_Partition class

As referred previously, during the implementation of the ontology many Object Properties and Data Properties were used totalizing 13 Object Properties and 6 Data Properties. On Figure 35 and Figure 36 it is possible to observe the full list of used Object Properties and Data Properties respectively. Table 14 represents the data type of the Data Properties.



- owl:topObjectProperty
  - affectedBy
  - affects
  - generatedBy
  - generates
  - hasAssessment
  - hasCriticality
  - hasImpact
  - hasSeverity
  - hasStatus
  - relatedToAssessment
  - relatedToIncident
  - triggeredBy
  - triggers

Figure 41 - List of used Object Properties



- owl:topDataProperty
  - completion
  - description
  - endTime
  - ID
  - startTime
  - type

Figure 42 - List of used Data Properties

Data	Data Type
Completion	Float
Description	Literal
EndTime	DateTime
ID	Long
StartTime	DateTime
Type	Literal

Table 16 - Data type of Data Properties

The following figure provides to the reader a detailed information about the relationships between the classes that were referred throughout this section.

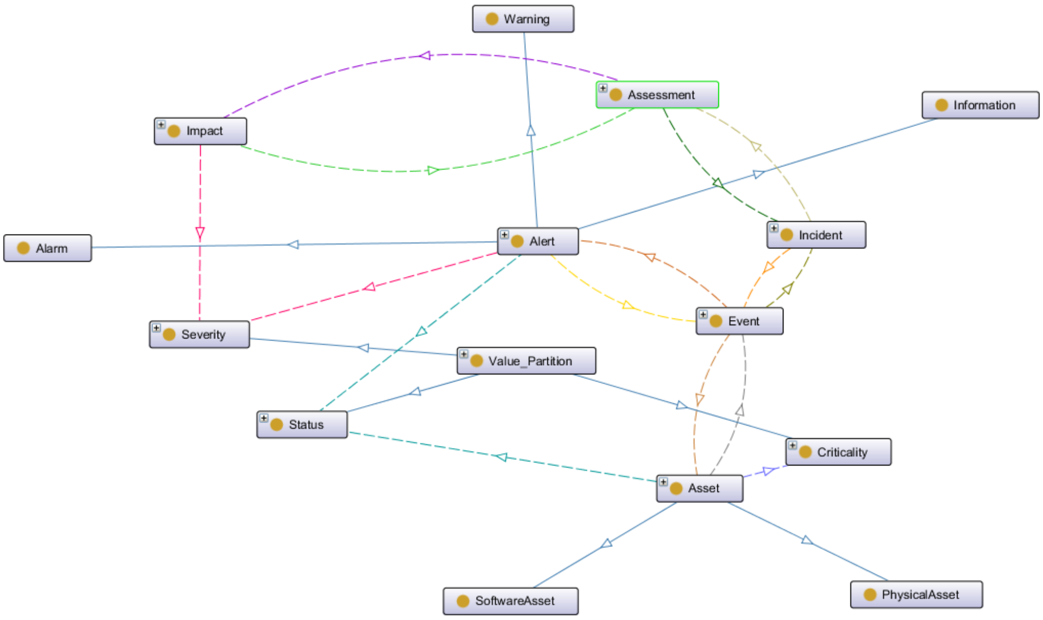


Figure 43 - OntoGraf of the full ontology

8.2 Expert System Implementation

As referred in previous chapters, a rule-based system can be defined as a system used to apply human-made rules, mostly developed by experts in the area, in order to store, sort and manipulate data with the goal of replicating human intelligence on the matter. In the context of the expert system development, as stated previously, it was decided that Drools was going to be the software used.

The developed expert system, in order to fulfill the requirements established at the beginning of this document, has to be prepared to control energy values, access and potential network intrusion. The system is going to have as base the JSON files obtained from sensors that generate alarms, so then it is able to classify whether or not the alarm is justified in that situation and what is it related to.

When it comes to the energy levels, the data collected will include:

- Start time

- End time
- Actual Energy Value
- Predicted Energy Value
- Difference between Energy Values
- Energy Values Threshold

The variables start time and end time, as the name indicates have the start and end time of the period of time where the energy values were out of ordinary based on the configurations that were done and on the expected values. These variables have the day, hour, minute and second.

The variable Actual Energy Value and Predicted Energy Value consist of the value that was collected in a given moment by the sensor and the expected value of energy that a certain asset is supposed to run on. In the context of this project, it is advised that the Actual Energy Value and Predicted Energy Value do not differ to much from each other. The variable Difference between Energy Values serves this purpose, understanding what the difference in terms of energy levels between the two variables is.

The last variable that will be considered is Energy Values Threshold. It is basically defined as the limit or maximum value that the difference between energy values can have. Different assets can have different energy values threshold, and in some cases, it can be only in superior values or in inferior values. For instance, there can be cases where the energy threshold is a certain value, but only applied when:

$$Actual\ Energy\ Value > \ Predicted\ Energy\ Value + Energy\ Values\ Threshold$$

OR

$$Actual\ Energy\ Value < \ Predicted\ Energy\ Value + Energy\ Values\ Threshold$$

OR

$$Actual\ Energy\ Value > \ Predicted\ Energy\ Value + Energy\ Values\ Threshold$$

AND

$$Actual\ Energy\ Value < \ Predicted\ Energy\ Value + Energy\ Values\ Threshold$$

When it comes to network intrusion, in order to understand the origin of the alarm, how potentially dangerous it could be and if it is possible or not to react to it, it is advisable to classify them (Atymtayeva & Kozhakhmet, 2014). They are commonly divided into three categories:

- Occasional vulnerabilities.

- Objective vulnerabilities.
- Subjective vulnerabilities.

Occasional vulnerabilities are usually related with the characteristics of the object environment or unforeseen circumstances. These factors can be predictable to some degree, but their removal is only possible during the engineering activities for decreasing or mitigating threats to network security. Some examples of occasional vulnerabilities would be faults and failures, such as malfunctions of technical equipment, software failures, power supply failure and equipment aging or damage to lifeline enclosing structures. It is hard to avoid occasional vulnerabilities, but it is more than necessary to prepare for when they are bound to happen.

The next type of vulnerabilities are Objective vulnerabilities. Objective vulnerabilities depend on features usually related with construction and technical parts of the equipment used. The removal of these type of vulnerabilities is not possible but it is possible to weaken them having the preparation to counter these types of threats. Examples of objective vulnerabilities include technical means of radiation such as electromagnetic, electrical, or even sound, clickable vulnerabilities that are included in the hardware and software tab, vulnerabilities determined by the characteristics of the projected objects, such as location object or organization communication channels.

The last type of vulnerabilities, Subjective vulnerabilities, and potentially the most important ones to the context of this project, are divided in two big groups: External vulnerabilities and Internal vulnerabilities. The first one depends mostly on attacks with external sources with malicious intents. Those external attacks can include database invasions, SSH or VNC scans or DDOS attacks for example and to avoid this type of attacks it is necessary to have a system ready for it and most of the related possibilities and in case there are some attacks that cannot be avoided at all or could not be avoided in a specific situation, it is necessary to detect them as soon as possible – an alarm notification should take place – and have a document with the protocols of reaction that the employees should have in specific situations. The second type of Subjective vulnerabilities depends mostly on the actions of the employees. Usually, the main ways employees can harm the company is through individual mistakes that can and most likely win happen regardless of the training that they have. The other way that employees can be harmful to the company is through leaks, acting as an insider for hackers or potential direct competition, both resulting in unauthorized access to private information. These types of vulnerabilities can be almost fully eliminated in most of the cases with an enhanced trainee programs and proper organizational software and hardware methods.

In this context, the data that will be collected and presented in the JSON file include:

- Origin
- Timestamp
- ID
- Event Type
- Source IP

- Source Port
- Destination IP
- Destination Port
- Protocol
- Signature

The variable Origin is related to the network from which the alarm comes from, making it possible to exclude every other network that could be considered.

The variable Timestamp, just like in the data collected to energy, will include the timestamp where the alarm event was generated. This variable includes the day, hour, minute and second.

The ID and Event Type variables will serve as a way to distinct alarms and its type respectively. In the context of this project, the only Event Type that will be considered in JSON files is Alerts. The decision to keep this variable comes with the possibility of adding other Event Types.

The next variables, Source IP and Source Port, Destination IP and Destination Port will serve to understand from what IP and port the alarms are being generated and what is the destination of this information.

The Protocol variable will indicate what protocol was used in the transmission of a specific alarm. In the context of this project, UDP and TCP protocols will be considered.

The last variable is Signature and basically it will consist on explaining why the alarm was triggered, whether it is suspicious behavior or with certainty.

The Expert system will be divided in three parts, each one considering a relevant part for understanding why a certain alarm was triggered and to be able to classify them.

The first part of the expert system will be focused on energy-related questions in order to understand if it is an energy problem overall. The first rule is based on understanding if there is any alarm that is active, so, in this case the system will be just making sure that it is being run for a specific reason. In case there is not any active alarm the process will end sending an *"OK"* message to the operator in duty. If the system locates there is an active alarm, then the next part of the process is to check the threshold energy values that are being applied on that kind of alarms and understand if the set value respects the value that it is supposed to have. If it is not respecting the value that it is bound to have, then the system will advise a re-configuration of the Threshold energy values, assuming it is not properly configured sending a message *"Threshold energy values are not properly configured"*. The next step of the system's process is to check the energy values levels according to the source of the alarm and compare it to the expected value. The system will then check the difference between the actual level of energy and the recommended level of energy and in case the difference is lower than the threshold it will assume everything energy-related is working as it should and start the Network Analysis, which is the second part of the expert system. In case the difference is higher than the threshold, the conclusive message can be “Caution – Energy value too high” if the actual energy value is higher than the expected or recommended value or “Caution – Energy value too low” if the actual value is lower than the expected value. Usually, a higher energy value means more danger. A detailed diagram with this first part of the expert system can be observed in Figure 44.

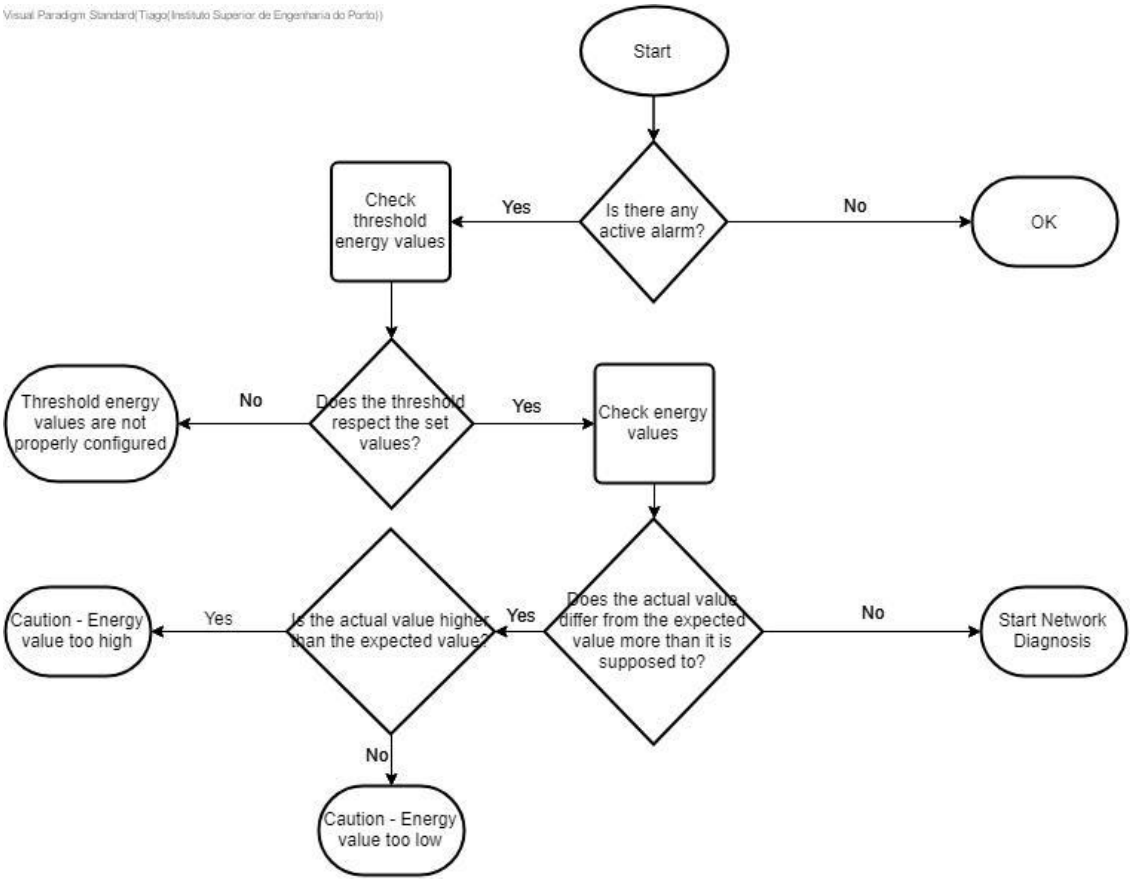

Visual Paradigm Standard(Tiago(Instituto Superior de Engenharia do Porto))
Start
Is there any active alarm?
Yes
No
OK
Check threshold energy values
Does the threshold respect the set values?
No
Threshold energy values are not properly configured
Yes
Check energy values
Does the actual value differ from the expected value more than it is supposed to?
No
Start Network Diagnosis
Yes
Is the actual value higher than the expected value?
Yes
Caution - Energy value too high
No
Caution - Energy value too low

Figure 44 - Expert System - Energy

The second part will be focused on the network analysis and will try to understand what type of vulnerability is being faced. The first thing that the system will do when starting this process is to run tests for faults, failures and potential infrastructure damage in order to understand if there is something wrong in that aspect. If it is the case, the system will classify the alarm as an occasional vulnerability and will advise to contact the technical support showing the message *“It is an occasional vulnerability – Call the technical support”.* In case there is not any physical damage, the system will then focus on checking for technical radiance levels and clickable vulnerabilities and in case there is evidence of vulnerabilities it will classify the alarm as an objective vulnerability and then, just like before, advise to contact the technical support. The shown message in this case will be *“It is an objective vulnerability – Call the technical* *support".* If everything looks controlled in those aspects, the system will assume the vulnerability is of subjective type and will start the third and last part of the process that consists of subjective vulnerabilities analysis. A detailed diagram with this first part of the expert system can be observed in Figure 45.

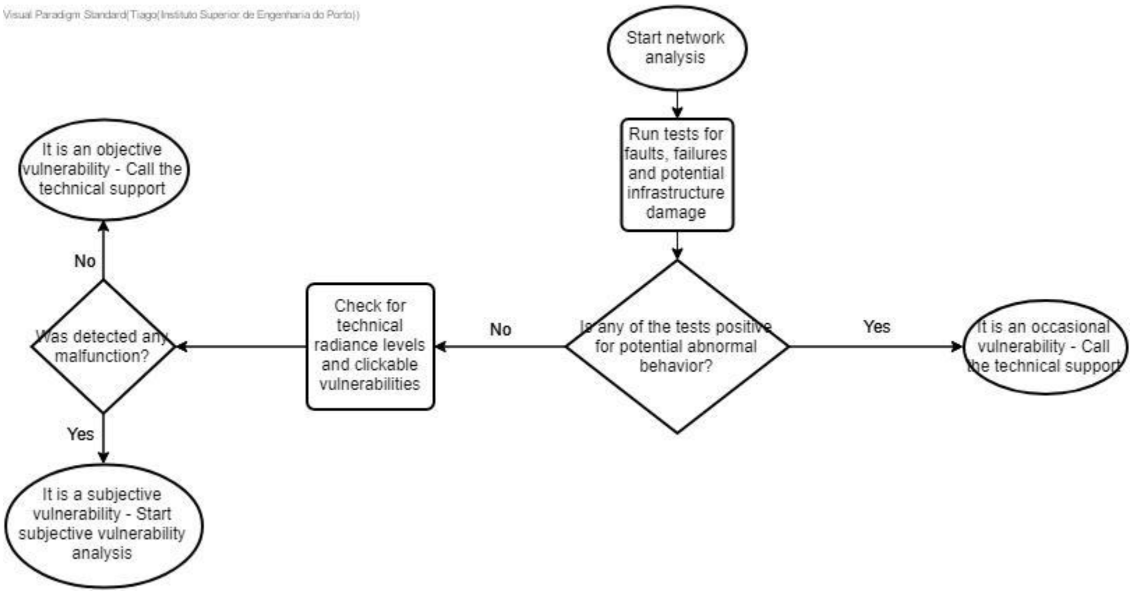

Visual Paradigm Standard(Tiago(Instituto Superior de Engenharia do Porto))
Start network analysis
Run tests for faults, failures and potential infrastructure damage
Is any of the tests positive for potential abnormal behavior?
Yes
It is an occasional vulnerability - Call the technical support
No
Check for technical radiance levels and clickable vulnerabilities
Was detected any malfunction?
No
It is an objective vulnerability - Call the technical support
Yes
It is a subjective vulnerability - Start subjective vulnerability analysis

Figure 45 - Expert System - Network Analysis

The third and last part of the expert system is subjective analysis, and it starts by checking for abnormal behaviour or potential breaches of information. In case there is abnormal behaviour but there was no breach of information then there is a high chance of staff mistake, with a slight possibility of a false positive. In this specific scenario the system will send the message *"High chance of staff mistake with a possible chance of a false positive".* Then, in case there was breach of information but with little to no abnormal behaviour, the system will send the message *"High chance of an insider's leak of information"*. To finish, and in case there was no leaks of information or blatant abnormal behaviour, the system will apply the message *"High chance of false positive or new actions to the system happened"* and in case both apply, the system will check for operative system inbound. After checking it, in case it seems positive, the system will end with a *"High chance of database being compromised"* and in case it is negative the system will end with a "*Relative chance of SSH or VNC scan*". A detailed diagram with this first part of the expert system can be observed in Figure 46.

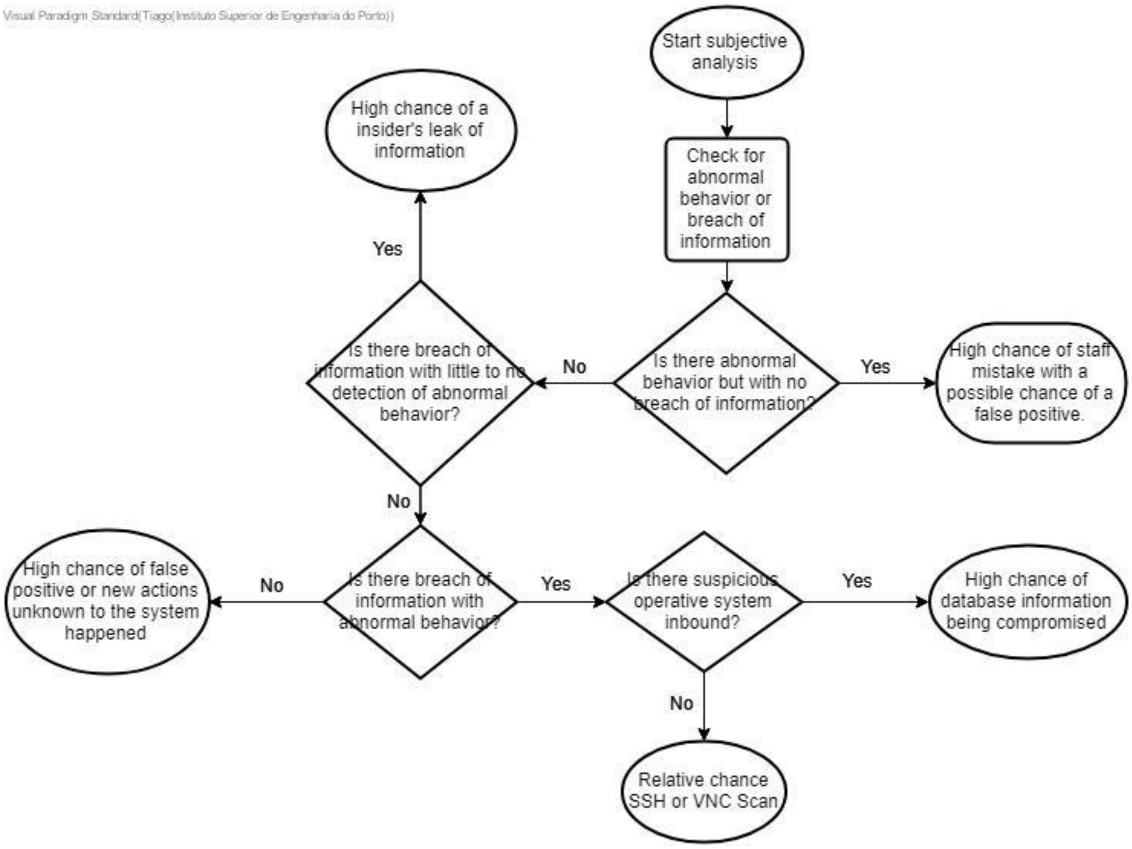

Visual Paradigm Standard(Tiago(Instituto Superior de Engenharia do Porto))
Start subjective analysis
Check for abnormal behavior or breach of information
Is there abnormal behavior but with no breach of information?
Yes
High chance of staff mistake with a possible chance of a false positive.
No
Is there breach of information with little to no detection of abnormal behavior?
Yes
High chance of a insider's leak of information
No
Is there breach of information with abnormal behavior?
No
High chance of false positive or new actions unknown to the system happened
Yes
Is there suspicious operative system inbound?
Yes
High chance of database information being compromised
No
Relative chance SSH or VNC Scan

Figure 46 - Expert System - Subjective Analysis

The rules that are included in the expert system will have a question number associated, so it can be easier to understand, program and communicate. Every single of these questions will then have rules related to them, creating a hypothesis and potentially a solution.

- Question 1 – Is there any active alarm?
- Question 2 – Is the threshold value properly configured?
- Question 3 – Does the energy value differ from the expected value more than it is supposed to?
- Question 4 – Is the actual value higher than the expected value?
- Question 5 – Is any of the tests positive for potential abnormal behaviour?
- Question 6 – Was detected any malfunction?
- Question 7 – Is there abnormal behaviour but with no breach of information?
- Question 8 – Is there breach of information but with little to no detection of abnormal behaviour?
- Question 9 – Is there breach of information with a lot of abnormal behaviour?
- Question 10 – Is there suspicious operative system inbound?

Rules answer	Conclusion
Q1 – “no”	“OK”
Q1 – “yes, Q2 – “no”	“Threshold values are not properly configured”
Q1 – “yes”, Q2 – “yes”, Q3 – “yes”, Q4 – “no”	“Caution – Energy value too low”
Q1 – “yes”, Q2 – “yes”, Q3 – “yes”, Q4 – “yes”	“Caution – Energy value too high”
Q1 – “yes”, Q2 – “yes”, Q3 – “no”, Q5 – “yes”	“It is an occasional vulnerability – Call the technical support”
Q1 – “yes”, Q2 – “yes”, Q3 – “no”, Q5 – “no”, Q6 – “no”	“It is an objective vulnerability – Call the technical support”
Q1 – “yes”, Q2 – “yes”, Q3 – “no”, Q5 – “no”, Q6 – “no”, Q7 – “yes”	“High chance of staff mistakes with a possible chance of false positive”
Q1 – “yes”, Q2 – “yes”, Q3 – “no”, Q5 – “no”, Q6 – “no”, Q7 – “no”, Q8 – “yes”	“High chance of an insider’s leak of information”
Q1 – “yes”, Q2 – “yes”, Q3 – “no”, Q5 – “no”, Q6 – “no”, Q7 – “no”, Q8 – “no”, Q9 – “no”	“High chance of false positive or new actions unknown to the system happened”
Q1 – “yes”, Q2 – “yes”, Q3 – “no”, Q5 – “no”, Q6 – “no”, Q7 – “no”, Q8 – “no”, Q9 – “yes”, Q10 – “no”	“Relative chance of SSH or VNC Scan”
Q1 – “yes”, Q2 – “yes”, Q3 – “no”, Q5 – “no”, Q6 – “no”, Q7 – “no”, Q8 – “no”, Q9 – “yes”, Q10 – “yes”	“High chance of database information being compromised”

Table 17 - Different outcomes of the conclusions in the Expert System

When it comes to the application, it will be divided in three different packages:

- Main
- Model
- View

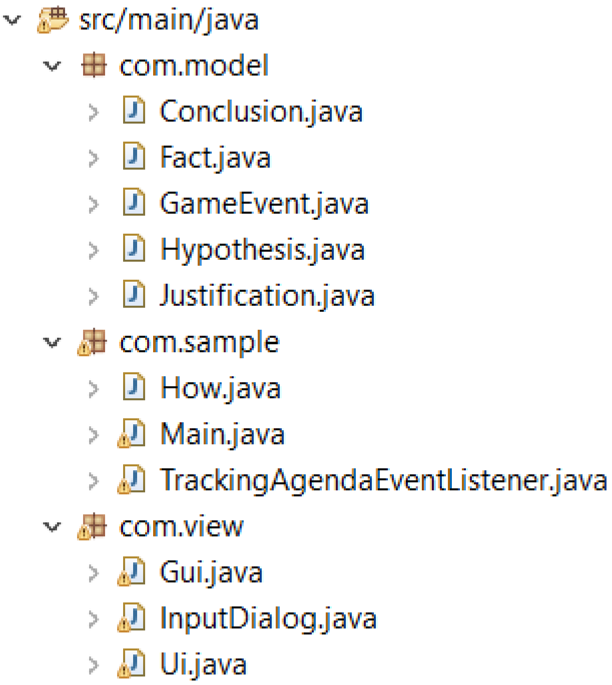
The View package will include everything related to front-end classes, such as GUI and UI.

The Model package includes the class Fact that serves as extension to every other class in this package. Those classes are:

- GameEvent – includes every event that could potentially happen, being related to the questions that are asked to reach conclusions later on.
- Conclusion – As the name indicates the class conclusions has the information regarding all the possible outcomes in the context of this project.
- Hypothesis – The class hypothesis contains information regarding possible hypothesis that are created when a conclusion is not given yet.

- Justification – Has the information regarding about a single part of process that culminated in a specific conclusion.

The Main package will include the main class, responsible for running the application. It will also include the How class that will include the combination of Justification class instances that culminated in a certain conclusion.



```
src/main/java
    com.model
        Conclusion.java
        Fact.java
        GameEvent.java
        Hypothesis.java
        Justification.java
    com.sample
        How.java
        Main.java
        TrackingAgendaEventListener.java
    com.view
        Gui.java
        InputDialog.java
        Ui.java
```

Figure 47 - Application Packages and respective classes

There will also be a drools file called GameRules that will be including all the rules that are going to be taken into account.

After we run it as a Java application, the user that is operating will have to answer a set of questions related to the alarm. In case he hits cancel, the whole process will be cancelled. The presented UI when the user launches the application is on Figure 48.

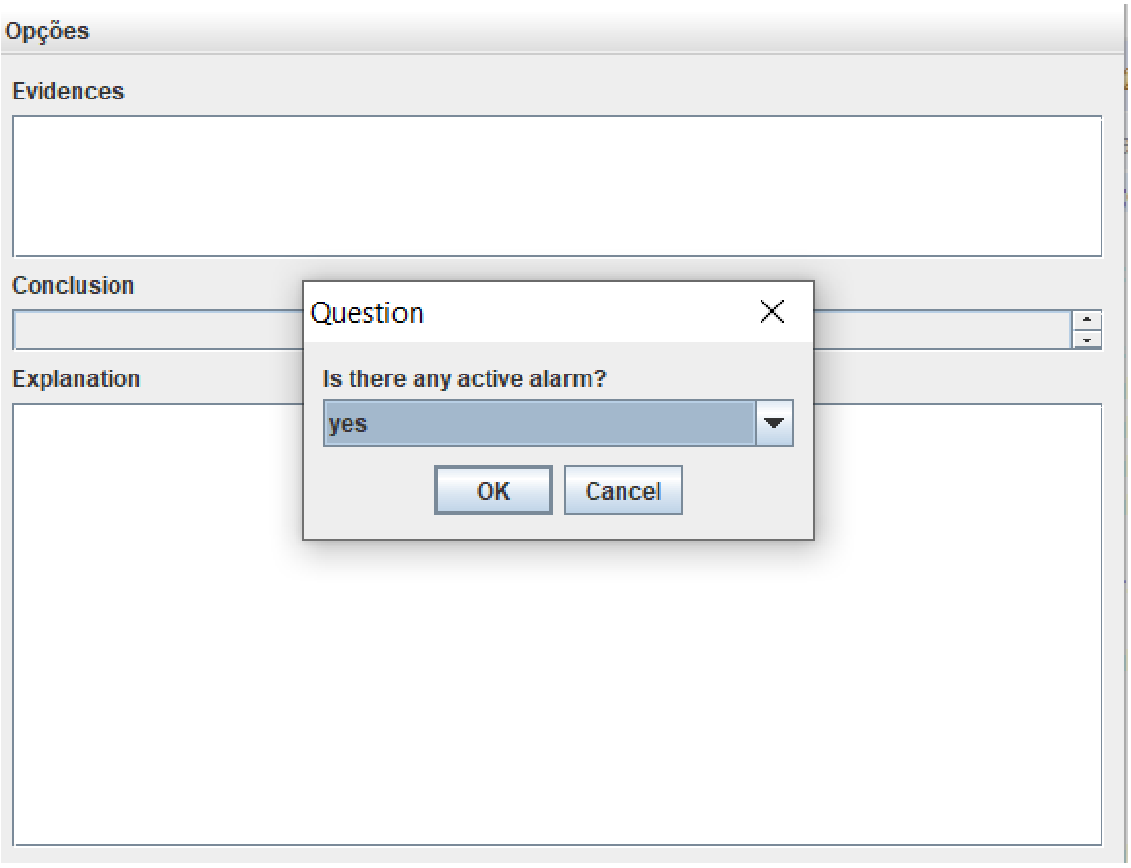
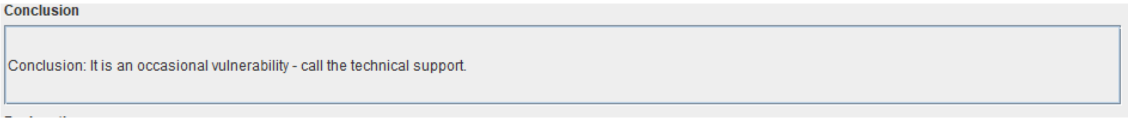


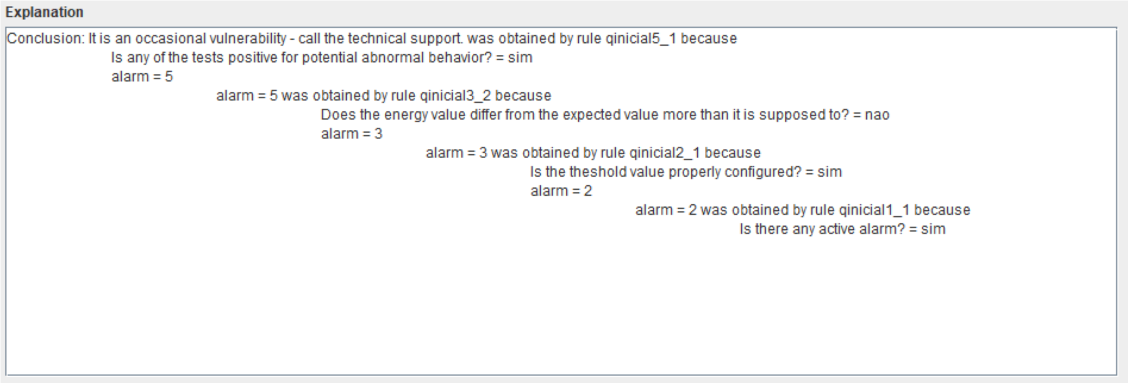
Figure 48 - Rule Based System UI

After answering all the questions that are presented, the user will then be allowed to see the conclusion, the explanation, the evidence and the tracking that was done to reach the final answer as it is possible to see on Figure 49, Figure 50 and Figure 51 respectively.



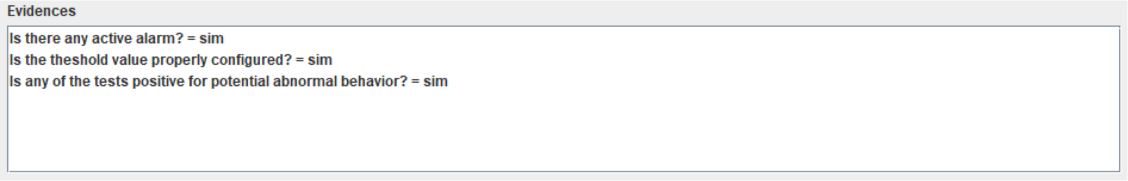
Conclusion: It is an occasional vulnerability - call the technical support.

Figure 49 - Conclusion



```
Conclusion: It is an occasional vulnerability - call the technical support. was obtained by rule qinicial5_1 because
          Is any of the tests positive for potential abnormal behavior? = sim
          alarm = 5
                    alarm = 5 was obtained by rule qinicial3_2 because
                              Does the energy value differ from the expected value more than it is supposed to? = nao
                              alarm = 3
                                        alarm = 3 was obtained by rule qinicial2_1 because
                                                  Is the theshold value properly configured? = sim
                                                  alarm = 2
                                                            alarm = 2 was obtained by rule qinicial1_1 because
                                                                      Is there any active alarm? = sim
```

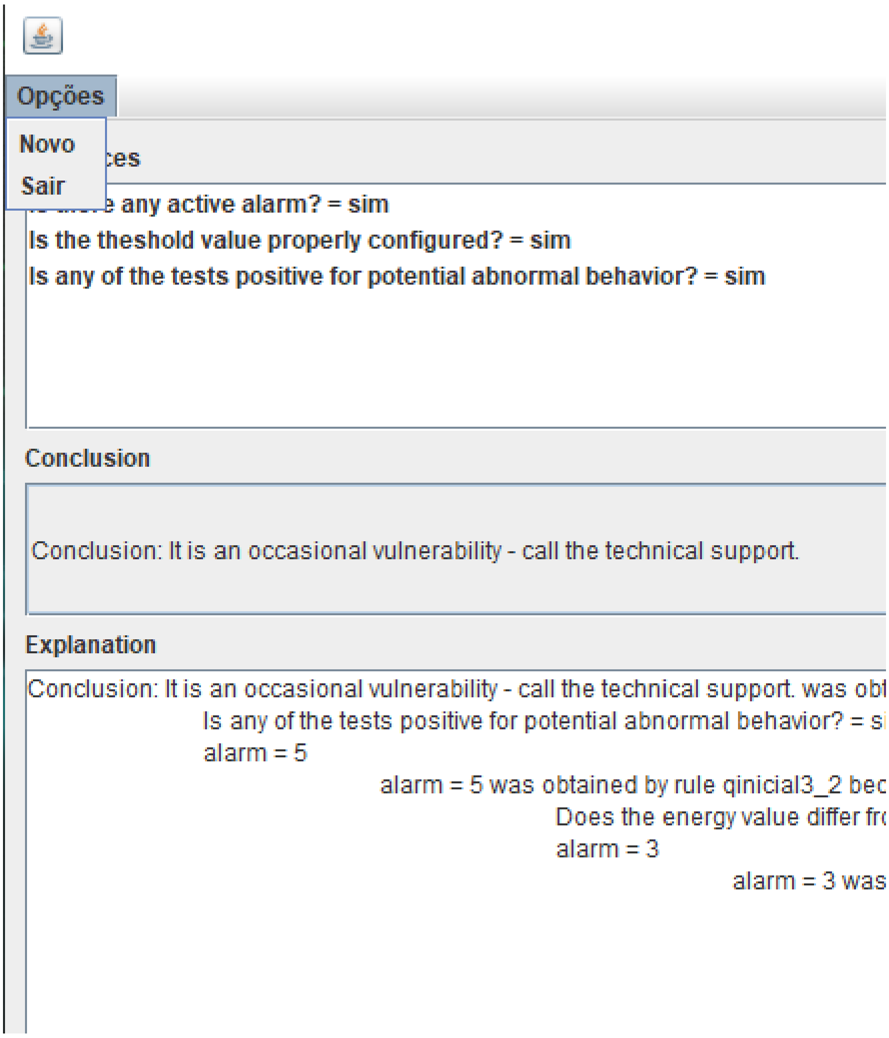
Figure 50 - Explanation


Evidences

Is there any active alarm? = sim
Is the theshold value properly configured? = sim
Is any of the tests positive for potential abnormal behavior? = sim

Figure 51 – Evidence

After finishing this whole process, the user will then be able to click options and start a new process in case he wants or simply hit the leave button to shut down the application.



ces

e any active alarm? = sim
Is the theshold value properly configured? = sim
Is any of the tests positive for potential abnormal behavior? = sim

Conclusion

Conclusion: It is an occasional vulnerability - call the technical support.

Explanation

```
Conclusion: It is an occasional vulnerability - call the technical support. was obt
                    Is any of the tests positive for potential abnormal behavior? = si
                    alarm = 5
                                        alarm = 5 was obtained by rule qinicial3_2 bec
                                                            Does the energy value differ fr
                                                            alarm = 3
                                                                                alarm = 3 was
```

Figure 52 - Options

9 Experiences and Evaluation

In this section, the main focus will be defining the hypothesis in order to validate the developed software, the indicators and the information sources and describe the evaluation methods to analyze the hypothesis referred previously. This chapter includes manual validation and accesses feedback related to the system towards an anonymous questionnaire.

9.1 Hypothesis definition

Null Hypothesis (H0): The system cannot fulfill the companies' requirements and is not helpful or needed to manage alarms in a proper way. The hypothesis can be divided according to the following criterion:

- The system does not have enough capabilities to be helpful or required to manage alarms.
- The system usage is not considered intuitive.

Alternative Hypothesis (H1): The system can fulfill the companies' requirements and is helpful or needed to manage alarms in a proper way. The hypothesis can be divided according to the following criterion:

- The system has enough capabilities to be helpful or required to manage alarms.
- The system usage is considered intuitive.

9.2 Indicators and Information Sources

This section will have as main focus presenting the indicators that will be needed to evaluate the hypothesis exposed previously.

9.2.1 Indicators

In order to take conclusions according to the hypothesis previously exposed, it is necessary to define which are going to be the indicators that will evaluate and conclude if the hypothesis is accepted or not. The chosen indicators to take part in this project are:

- **Follow the Functional and Non-functional requirements –** Fulfilling both requirements will mean that the project can succeed according to the objectives described in this document.
- **Usability –** Having a good usability and the companies feeling that the system is helpful is one of the most crucial indicators to be pointed.

9.2.2 Information Sources

The information sources include:

- **User Rating Inquiry –** Inquiry that will be able to evaluate the user opinion based on his satisfaction and usability.
- **Manual Validation –** Validating the work manually using guidelines provided.

9.3 Evaluation Methods

This section focuses mainly on understanding if the hypothesis referred earlier can be verified or not in the context of this project.

9.3.1 User Rating Inquiry

In order to analyze the user's opinion a user rating inquiry will be developed and will be focused on specific questions with restricted answering. The answers will consist of 5 options, following the *likert* scale:

- Completely Agree
- Agree
- Indifferent
- Disagree
- Completely Disagree

The target users are the companies that are going to use the software most specifically the employees that will have direct contact with it and the mainly focus of the inquiry will be the difficulty to interact with the system and how easy it is to understand the information being displayed. The questionnaire that was delivered is on **Appendix B.**

9.3.2 Manual Validation

The project will be exposed to a lot of actions that are going to be taken to ensure that the implementation is proper. These include ontology OWL validation and analysis of sample data, for example.

9.4 Conducted Evaluation

This section has as main goal ensuring the validity of the project that was developed. OWL validations, analysis of sample data from relevant systems will be done.

Initially, the classes in AMMOThing were checked to ensure that the hierarchy was properly configured and made sense, so basically the 'is-a' relationships were analysis, and it was made sure there was no cycles within the classes. It was also checked the possibility of duplicate classes – with same function but different name – but no redundant class was found.

When it comes to naming conventions, the idea was to use camelCase for the name of the classes, but as this version of Protégé, the ontology editor, was not allowing it forced the project to use capital letter on the first character. After a check of all the classes and siblings, it was confirmed that every one of them was following the set rule.

The third part was to check for instances that belong to two disjoin classes. Missing the disjoint is a common mistake in ontology modelling, so this step was manually checked to make sure no inconsistencies were created.

The last part of the manual consisted of checking the siblings in the class hierarchy in order to make sure they have siblings but not too many. It was not found any class that had no siblings or a class that had too many siblings, resulting in a test pass.

10 Conclusion

Technology is in constant evolution and Industry 4.0 takes part in this process as it is starting to have an increasing impact in human life. Despite its development in recent years, Industry 4.0 as a whole is a very recent concept and is yet to reach its full potential. When it comes to alarm management, it is safe to say that some work has been done towards the development of the topic, but most part of it, is still in embryonic state and great part of the software developed is most likely not documented publicly, therefore not being easy to access reliable information in most cases regarding this subject.

Alarm management is the keyword of this project and in order to understand fully about this concept and technologies and domains referred in this document, a state of the art was conducted to enhance the knowledge related to the subjects involved and to have a better perspective of what the system would need moving forward. After reading innumerous publications it was time to make decisions, being decided that ontologies and expert systems were advisable in the context of this project. The relevant related publications were then cautiously filtered, studied, summarized and presented in this document in a simple and engaging approach in order to make it a better and easier experience for the reader.

The usage of ontologies in the context of this project serves to achieve an easier integration of all the systems involved by using a common vocabulary, which is the reason why this work focused with some detail the methodologies behind this subject matter, helping decide in the end what methodology would be used and the tool that would help create the final ontology. It proved to be a great decision as it creates an environment where every part of the project and people involved can communicate easily regarding each different concept and it allows the possibility of an expansion in the future as ontologies are maintain, add new knowledge and reuse other ontologies.

As for the usage of the expert system in the context of this project, it focuses on the main problem of the document, despite being just as important as an ontology, which is dealing with the alarms that are being published and helping a company's operator understand what the reasoning behind the alarm publishment is and how he should deal with it. There were other options such as using the Machine Learning technology in this project but given the presented problem and the context where it is being developed on, it proved to be a great decision not to follow the ML path.

The developed ontology was implemented in OWL ontology language and used Protégé as its editor. The main classes were described in this document and a full list of class definition was provided. The object properties were illustrated using OntoGrafs. As for the expert system, it was developed using drools and every rule and potential outcome was also described in the present document.

When it comes to the objectives that were defined at the beginning of this document, they were achieved overall, except for some use cases. When the project started, it was not supposed to have an ontology being developed, which compromised some of the development that could have been directed towards the expert system when the attention shifted towards having a solid ontology that could be used in the future. Overall, an ontology was developed, and an expert system was created. It is safe to say that there is obviously still room for improvement and expansion when it comes to the development of the project in these two senses, but it shows potential.

To understand how this project can be useful in a competitive company environment, it is necessary to test it and monitor its performance throughout different scenarios to be able to receive feedback and ultimately solve the issues presented and upgrade the project into a better version.

To add more depth to potential improvement on the service, a questionnaire was created to understand what the overall opinion regarding the project is. As it is expected to have feedback from an outside perspective it was supposed to have the questionnaire delivered to potential partners or companies interested in the project development. As the deadline to deliver the project will be before having any contact with any group of people that could have a positive impact on the project development moving forward, this document will not include the overall opinion but will include the questions that are going to be asked.

Future prospects for this project include listening to feedback in order to start expanding the ontology according to the necessity related to it and continue developing the expert system when it comes to the knowledge based involved. Other possibility for the development and healthy growth of this project would be developing a machine learning system to replace the expert system or potentially have the project running both to have a more accurate answer or solution.

References

Angular, 2021. *Angular.* [Online]
Available at: https://angular.io/guide/architecture
[Acedido em 05 03 2021].

Anon., 2004. *Resource Description Framework (RDF): Concepts and Abstract Syntax.* [Online]
Available at: https://www.w3.org/TR/rdf-concepts/
[Acedido em 20 02 2022].

Anon., 2011. *Cognitum Software House.* [Online]
Available at: https://cognitum.eu/semantics/fluenteditor/
[Acedido em 20 02 2022].

Anon., 2012. *OWL.* [Online]
Available at: https://www.w3.org/OWL/
[Acedido em 20 02 2022].

Anon., 2014. *NeOn Toolkit.* [Online]
Available at: http://neon-toolkit.org/wiki/Main_Page.html
[Acedido em 20 02 2022].

Anon., 2015. *Ontology Language.* [Online]
Available at: https://www.sciencedirect.com/topics/computer-science/ontology-language
[Acedido em 20 02 2022].

Anon., 2016. *Protégé.* [Online]
Available at: https://protege.stanford.edu/
[Acedido em 20 02 2022].

Atymtayeva, L. & Kozhakhmet, K., 2014. *Building a Knowledge Base for Expert System in Information Security.* [Online]
Available at:
https://www.researchgate.net/publication/278027130_Building_a_Knowledge_Base_for_Expert_System_in_Information_Security
[Acedido em 11 10 2022].

B.R. Mehta, Y. R., 2015. *Industrial Process Automation Systems.* [Online]
Available at: https://www.sciencedirect.com/book/9780128009390/industrial-process-automation-systems
[Acedido em 14 04 2022].

baeldung, 2022. *Introduction to Drools.* [Online]
Available at: https://www.baeldung.com/drools
[Acedido em 27 08 2022].

Bhatt, S., 2018. *Reinforcement Learning 101.* [Online]
Available at: https://towardsdatascience.com/reinforcement-learning-101-e24b50e1d292
[Acedido em 19 08 2022].

Bigelow, S., 2022. *Microsoft Azure.* [Online]
Available at: https://www.techtarget.com/searchcloudcomputing/definition/Windows-Azure
[Acedido em 29 08 2022].

Black, P., 2015. *Alarm Management Systems Third-Party Software Pros and Cons.* [Online]
Available at: https://www.processindustryinformer.com/alarm-management-systems-third-party-software-pros-cons
[Acedido em 05 03 2021].

Bosch, 2022. *Bosch - Products and Services - Industry and Trades.* [Online]
Available at: https://www.bosch.com/products-and-services/industry-and-trades/
[Acedido em 09 09 2022].

Brownlee, J., 2020. *A tour of Machine Learning Algorithms.* [Online]
Available at: https://machinelearningmastery.com/a-tour-of-machine-learning-algorithms/
[Acedido em 19 08 2022].

Burns, E., 2018. *Machine Learning.* [Online]
Available at: https://www.techtarget.com/searchenterpriseai/definition/machine-learning-ML
[Acedido em 18 08 2022].

Carew, J., 2020. *How to choose between a rules-based vs Machine Learning System.* [Online]
Available at: https://www.techtarget.com/searchenterpriseai/feature/How-to-choose-between-a-rules-based-vs-machine-learning-system
[Acedido em 21 08 2022].

Chang, C.-W., 2020. *Evaluation of Smart Alarm Systems for Industry 4.0 Technologies.* [Online]
Available at:
https://www.researchgate.net/publication/339995010_Evaluation_of_Smart_Alarm_Systems_for_Industry_40_Technologies
[Acedido em 05 03 2021].

Cheng, C.-W., 2020. *Comparing the CEGRA method and the TOPSIS method.* [Online]
Available at: https://www.researchgate.net/figure/Comparing-the-CEGRA-method-and-the-TOPSIS-method_fig1_339995010
[Acedido em 05 03 2021].

Daniel, D., 2018. *Industry 4.0.* [Online]
Available at: https://searcherp.techtarget.com/definition/Industry-40
[Acedido em 05 03 2021].

DataFlair, 2021. *Advantages and Disadvantages of Machine Learning.* [Online]
Available at: https://data-flair.training/blogs/advantages-and-disadvantages-of-machine-

learning/
[Acedido em 19 08 2022].

DecisionRules, 2021. *DecisionRules - Business rules engine that streamlines your business workflow.* [Online]
Available at: https://www.decisionrules.io/
[Acedido em 27 08 2022].

DecisionRules, 2021. *DecisionRules - Powerful rule engine features for smooth decision making.* [Online]
Available at: https://www.decisionrules.io/features
[Acedido em 27 08 2022].

Domova, V. & Dagnino, A., 2017. *Towards intelligent alarm management in the Age of IIoT.* [Online]
Available at: https://ieeexplore.ieee.org/document/8016234
[Acedido em 05 03 2021].

Drools, 2012. *Drools.* [Online]
Available at:
https://www.immagic.com/eLibrary/ARCHIVES/GENERAL/WIKIPEDI/W120817D.pdf
[Acedido em 27 08 2022].

Dubois, L., 1957. *The sense and nonsense of Alarm System performance KPIs.* [Online]
Available at: https://www.processvue.com/en/white-p-down?download=24:alarm-system-performance-kpis
[Acedido em 03 05 2021].

Eisenbart, M., 2022. *The next step in Industry 4.0: rule-based analysis of production data.* [Online]
Available at: https://blog.bosch-si.com/industry40/next-step-industry-4-0-rule-based-analysis-production-data/
[Acedido em 18 09 2022].

Engati, 2017. *What is a rule-based System?.* [Online]
Available at: https://www.engati.com/glossary/rule-based-system#toc-what-are-the-characteristics-of-rule-based-systems-
[Acedido em 16 08 2022].

García, A., Portes, L. & Pinto, F., 2012. *Real-time Alarm Management System for Emergency Situations.* [Online]
Available at: https://www.researchgate.net/publication/237021890_Real-Time_Alarm_Management_System_for_Emergency_Situations
[Acedido em 10 05 2022].

GECAD, 2013. *GECAD - Brief Presentation.* [Online]
Available at: http://www.gecad.isep.ipp.pt/GECAD/Pages/Presentation/Home.aspx
[Acedido em 05 03 2021].

González, E. & Marichal, R., 2005. *Software Experience for an Ontology-Based Approach for the Definition of Alarms in Geographical Sensor Systems.* [Online]
Available at: https://ieeexplore.ieee.org/document/8478129
[Acedido em 01 06 2022].

Group, R. W., 2014. *Resource Description Framework (RDF).* [Online]
Available at: https://www.w3.org/RDF/
[Acedido em 20 02 2022].

Grubber, T., s.d. *What is an ontology.* [Online]
Available at: (http://www-ksl.stanford.edu/kst/what-is-an-ontology.html)
[Acedido em 20 02 2022].

Hussein, S., 2018. *Technical Difference between AHP and TOPSIS.* [Online]
Available at:
https://www.researchgate.net/post/What_are_the_basic_and_technical_difference_between_AHP_and_TOPSIS
[Acedido em 05 03 2021].

Hyderabad, B. A. T., 2014. *What is FURPS+?.* [Online]
Available at: https://businessanalysttraininghyderabad.wordpress.com/2014/08/05/what-is-furps/
[Acedido em 05 03 2021].

Jong-Soo, S., 2014. *Ontology Languages.* [Online]
Available at: https://www.slideserve.com/dick/ontology-languages
[Acedido em 20 02 2022].

Kibana, 2021. *Kibana.* [Online]
Available at: https://www.elastic.co/pt/kibana
[Acedido em 05 03 2021].

Kourtis, G., Kavakli, E. & Sakellariou, R., 2020. *A Rule-Based Approach Founded on Description Logics for Industry 4.0 Smart Factories.* [Online]
Available at:
https://www.academia.edu/63472403/A_Rule_Based_Approach_Founded_on_Description_Logics_for_Industry_4_0_Smart_Factories
[Acedido em 02 10 2022].

Kylo, 2017. *Kylo, an open-source data lake.* [Online]
Available at: https://kylo.io/
[Acedido em 05 03 2021].

Lessware, S., 2022. *Why a Rules Based plus a Machine Learning approach.* [Online]
Available at: https://1spatial.com/news-events/2021/why-a-rules-based-plus-a-machine-learning-hybrid-approach/
[Acedido em 21 08 2022].

Liu, H., Gegov, A. & Stahl, F., 2014. *Categorization and Construction of Rule-Based Systems.* [Online]
Available at: https://link.springer.com/chapter/10.1007/978-3-319-11071-4_18
[Acedido em 16 08 2022].

LogicV, 2021. *Logic Apps vs Azure Functions: Which one should you choose?.* [Online]
Available at: https://logicv.com/blog/logic-apps-vs-azure-functions/
[Acedido em 08 09 2022].

Lutkevich, B., 2020. *Association rules.* [Online]
Available at: https://searchbusinessanalytics.techtarget.com/definition/association-rules-in-data-mining
[Acedido em 20 02 2022].

McCoy, L., 2022. *What is Microsoft Azure and why it matters.* [Online]
Available at: https://ccbtechnology.com/what-microsoft-azure-is-and-why-it-matters/
[Acedido em 29 08 2022].

Microsoft, 2020. *SQL Databases.* [Online]
Available at: https://docs.microsoft.com/pt-pt/azure/azure-sql/database/sql-database-paas-overview
[Acedido em 05 03 2021].

Microsoft, 2020. *Stream Analytics Introduction.* [Online]
Available at: https://docs.microsoft.com/pt-pt/azure/stream-analytics/stream-analytics-introduction
[Acedido em 05 03 2021].

Microsoft, 2022. *Azure Function Applications.* [Online]
Available at: https://azure.microsoft.com/pt-pt/services/functions/
[Acedido em 03 09 2022].

Microsoft, 2022. *Azure Logic Apps.* [Online]
Available at: https://azure.microsoft.com/pt-pt/services/logic-apps/#resources
[Acedido em 03 09 2022].

Microsoft, 2022. *Choose the right integration and automation services in Azure.* [Online]
Available at: https://docs.microsoft.com/en-us/azure/azure-functions/functions-compare-logic-apps-ms-flow-webjobs
[Acedido em 08 09 2022].

Microsoft, 2022. *Choose the right integration and automation services in Azure.* [Online]
Available at: https://docs.microsoft.com/en-us/azure/azure-functions/functions-compare-logic-apps-ms-flow-webjobs
[Acedido em 08 09 2022].

MongoDB, 2021. *Relational vs non-relational databases.* [Online]
Available at: https://www.mongodb.com/scale/relational-vs-non-relational-database
[Acedido em 05 03 2021].

MongoDB, 2021. *What is MongoDB.* [Online]
Available at: https://www.mongodb.com/what-is-mongodb
[Acedido em 05 03 021].

ontotext, 2022. *What are ontologies?.* [Online]
Available at: https://www.ontotext.com/knowledgehub/fundamentals/what-are-ontologies/
[Acedido em 02 10 2022].

Oracle, 2021. *Oracle.* [Online]
Available at: https://www.oracle.com/internet-of-things/what-is-iot/
[Acedido em 05 03 2021].

O'Reilly, 2022. *Advantages of Ontologies.* [Online]
Available at: https://www.oreilly.com/library/view/artificial-intelligence-for/9781788472173/1b0925f7-9346-4896-a821-c1742b383335.xhtml
[Acedido em 02 10 2022].

Panda, B., Misra, A. & Naik, A., 2012. *Development of a rule-based expert system with fuzzy sets and fuzzy logic.* [Online]
Available at:
https://www.researchgate.net/publication/232707515_DEVELOPMENT_OF_A_RULE-BASED_EXPERT_SYSTEM_WITH_FUZZY_SETS_AND_FUZZY_LOGIC
[Acedido em 16 08 2022].

Paragyte Technologies, 2017. *Key Advantages of Drools Rule Engine.* [Online]
Available at: https://medium.com/@paragyte2/key-advantages-of-drools-rule-engine-96e3353129b0
[Acedido em 27 08 2022].

ProAsset, 2021. *ProIMS Alarm Management.* [Online]
Available at: https://www.proasset.net/proims-alarm-management.html
[Acedido em 05 03 2021].

Ramakrishnan, S., 2010. *Ontology Languages - A review.* [Online]
Available at: https://www.researchgate.net/publication/269801838_Ontology_Languages_-_A_Review
[Acedido em 20 02 2022].

Red Hat Decision Manager, 2022. *Red Hat Decision Manager.* [Online]
Available at: https://www.redhat.com/en/technologies/jboss-middleware/decision-manager
[Acedido em 28 08 2022].

Sahu, A., 2020. *Advantages and Disadvantages of Machine Learning.* [Online]
Available at: https://www.westagilelabs.com/blog/pros-and-cons-of-implementing-machine-learning-in-your-projects/
[Acedido em 21 08 2022].

Search, E., 2021. *What is elastic search?.* [Online]
Available at: https://www.elastic.co/pt/what-is/elasticsearch
[Acedido em 05 03 2021].

Shah, M., 2021. *Harvest NoSQL Speed With The Combination Of PHP.* [Online]
Available at: https://www.clariontech.com/blog/harvest-nosql-speed-with-the-combination-of-php
[Acedido em 05 03 2021].

Silva, F., Santos, G., Praça, I. & Vale, Z., 2018. *A context-based building security alarm through power and sensors analysis.* [Online]
Available at: https://energyinformatics.springeropen.com/articles/10.1186/s42162-018-0045-z
[Acedido em 05 03 2021].

Smith, R., 2020. *The Key Differences Between Rule-Based AI and Machine Learning.* [Online]
Available at: https://becominghuman.ai/the-key-differences-between-rule-based-ai-and-machine-learning-8792e545e6
[Acedido em 21 08 2022].

Swaminathan, V., 2020. *The Conundrum of using Rule-Based vs. Machine Learning Systems.* [Online]
Available at: https://www.zucisystems.com/blog/the-conundrum-of-using-rule-based-vs-machine-learning-systems/
[Acedido em 21 08 2022].

tbyyf, 2022. *Artificial Intelligence and Machine Learning and Deep Learning.* [Online]
Available at: https://miro.medium.com/max/631/0*Q3PICBlib-932hhH.png
[Acedido em 18 08 2022].

techvidvan, 2022. *Exploring the Advantages and Disadvantages of Machine Learning.* [Online]
Available at: https://techvidvan.com/tutorials/advantages-and-disadvantages-of-machine-learning/
[Acedido em 19 08 2022].

ThinkAutomation, 2017. *What is a rule based system and what is it not.* [Online]
Available at: https://www.thinkautomation.com/eli5/what-is-a-rule-based-system-what-is-it-

not/
[Acedido em 16 08 2022].

TrustRadius, 2021. *CosmosDB vs SQL Azure.* [Online]
Available at: https://www.trustradius.com/compare-products/azure-cosmos-db-vs-sql-azure
[Acedido em 05 03 2021].

Vallespir, B., 2009. *Ontology Languages Classification.* [Online]
Available at: https://www.researchgate.net/figure/Ontology-languages-classification_tbl1_224595199
[Acedido em 20 02 2022].

Velazquez, R., 2022. *Introduction to AI.* [Online]
Available at: https://builtin.com/artificial-intelligence
[Acedido em 18 08 2022].

White, N., 2019. *Digital transformation statistics.* [Online]
Available at: https://www.ptc.com/en/blogs/corporate/digital-transformation-statistics
[Acedido em 05 03 2021].

Williamson, K., 2015. *Rule-Based Expert Systems.* [Online]
Available at: https://slideplayer.com/slide/6848089/
[Acedido em 16 08 2022].

Zamora, V. & Sipele, O., 2017. *Intelligent Agents for Supporting Driving Tasks: An Ontology-based.* [Online]
Available at: https://www.scitepress.org/papers/2017/62476/62476.pdf
[Acedido em 07 05 2022].

Zhang, E., 2018. *What is Event correlation?.* [Online]
Available at: https://digitalguardian.com/blog/what-event-correlation-examples-benefits-and-more
[Acedido em 20 02 2022].

Appendices

Appendix A

Alarm - Defined as a specific type of Alert, usually referred to Alerts with a high level of severity

Alert - The concept of *Alert* is defined as a notification that implies that certain protocols have been broken and therefore the system, or part of it, is not behaving as it is supposed to

AMMOThing - The classes' implementation process starts with the definition of a local class that will be a direct subclass of the default OWL:Thing class that is automatically generated. The AMMOThing class will serve as a parent to all of the developed classes and will have as its main role establishing the relationships between classes and how they are grouped and distributed

Assessment - The Assessment class has as main role providing, in detail, everything that happened after an incident occurred, especially when it comes to the procedure route so basically if it was followed or if there was improvisation in the between. This class is important as it will allow for future incidents that are eventually similar to one that happened before being analyzed beforehand

Asset - The concept of Asset was introduced on the ontology with the purpose of representing resources that have an important role in the proper functioning of the general industry system that it is inserted into

Catastrophic – Catastrophic levels of criticality

Completion – Float data property related to completion

Criticality - Provides the criticality level and is divided into five categories: No_Effect, Minor, Major, Emergency, Catastrophic

Description – Literal data property related to the description

Disabled – Disabled status

Emergency – Emergency levels of criticality

Enabled – Enabled status

EndTime – DateTime data property related to the EndTime

Event - The Event class represents random events directly related to a change of criticality, severity or status of an Alert or Alarm or the change of status of at least an Asset. The events are generated by an occurring Incident

High – High level of severity

ID – Float data property related to the ID

Impact - Defines the severity level of the impact that an occurring Incident had based on the Assessment that is related to it.

Incident - The Incident class is introduced as the responsible for generating events from the Event class. It can generate more than one event and has a specific start time and end time. Every Incident is related to an Assessment related to it, making an evaluation and report of what happened in the specific Incident

Information - Contains synthetized information regarding the Alert in order to make it easily understandable.

Low – Low level of severity

Major – Major levels of criticality

Medium – Medium level of severity

Minor – Minor levels of criticality

NoEffect – Levels of criticality with little to no effect

PhysicalAsset – An asset that is physical

Severity - Provides the severity level and is divided into five categories: Very_Low, Low, Medium, High, Very_High

SoftwareAsset – An asset that is software

StartTime – DateTime data property related to the StartTime

Status - Provides the status and is divided into two categories: Enabled, Disabled

Type – Literal data property related to the type

ValuePartition - The Value_Partition class has the goal of maintaining as subclasses all the variables with set pre-defined values. They have the goal of restricting the number of pre-defined values and providing clear and objective values, so it is easier to understand

VeryHigh – Very High level of severity

VeryLow – Very low level of severity

Warning - Contains information regarding the procedures that should be taken in case of a specific alert, if it ever happened before, in order to ensure that the decision-making process when it comes to reacting to the alert is precise

Appendix B

Questions:

1. Industry 4.0 is a very recent concept, and in this project, Alarm Management Module for Industry 4.0 we apply expert systems and ontologies on it. Do you have any previous experience in the context of expert systems or ontologies?

   Type: choice with single answer

   Choices: - Both, Only ontologies, Only expert systems, No.

2. Have you ever used ontologies or expert systems in your business?

   Type: choice with single answer

   Choices: - Both, Only ontologies, Only expert systems, No.

3. In case you used it before, in what application domain did you do it?

   Type: choices with multiple answers

   Choices: - Cyber Security, IoT, Medicine, Physical Security, Other

4. What would you add to the ontology or to the expert system that you think is crucial and is missing?

   Type: option text input

5. Do you think the expert system is deep enough to cover most part of a company related to Industry 4.0 problems?

   Type: rating

6. In the context of this project, do you feel like ontologies hold great value when thinking about the long run?

   Type: choices with single answer

   Choices – Yes, No

7. How much better do you think ontologies made the communication and overall understanding of the project?

Type: rating

8. How would you rate Alarm Management Module for Industry 4.0 in terms of potential?

   Type: rating

9. Would you use the expert alarm management system or ontology in any of your upcoming projects?

   Type: choice with single answer.

   Choices: - Yes, No.

10. Is there any feedback or suggestions you would like to give regarding the developed project?

    Type: option text input.