CRANFIELD UNIVERSITY

Santiago Quintana-Amate

A New Knowledge Sourcing Framework to Support Knowledge-Based Engineering Development

SCHOOL OF AEROSPACE, TRANSPORT AND MANUFACTURING

PhD Thesis

CRANFIELD UNIVERSITY

SCHOOL OF AEROSPACE, TRANSPORT AND MANUFACTURING

PhD THESIS

Academic Year 2014 – 2015

SANTIAGO QUINTANA-AMATE

A NEW KNOWLEDGE SOURCING FRAMEWORK TO SUPPORT KNOWLEDGE-BASED ENGINEERING DEVELOPMENT

Academic supervisor: Professor Tiwari Ashutosh

Industrial supervisor: Dr. Pablo Bermell-Garcia

August 2015

This thesis is submitted in partial fulfilment of the requirements for the degree of Doctor of Philosophy

ABSTRACT

New trends in Knowledge-Based Engineering (KBE) highlight the need for decoupling the automation aspect from the knowledge management side of KBE. In this direction, some authors argue that KBE is capable of effectively capturing, retaining and reusing engineering knowledge. However, there are some limitations associated with some aspects of KBE that present a barrier to deliver the knowledge sourcing process requested by the industry. To overcome some of these limitations this research proposes a new methodology for efficient knowledge capture and effective management of the complete knowledge life cycle.

Current knowledge capture procedures represent one of the main constraints limiting the wide use of KBE in the industry. This is due to the extraction of knowledge from experts in high cost knowledge capture sessions. To reduce the amount of time required from experts to extract relevant knowledge, this research uses Artificial Intelligence (AI) techniques capable of generating new knowledge from company assets. Moreover the research reported here proposes the integration of AI methods and experts increasing as a result the accuracy of the predictions and the reliability of using advanced reasoning tools. The proposed knowledge sourcing framework integrates two features: (i) use of advanced data mining tools and expert knowledge to create new knowledge from raw data, (ii) adoption of a well-established and reliable methodology to systematically capture, transfer and reuse engineering knowledge.

The methodology proposed in this research is validated through the development and implementation of two case studies aiming at the optimisation of wing design concepts. The results obtained in both use cases proved the extended KBE capability for fast and effective knowledge sourcing. This evidence was provided by the experts working in the development of each of the case studies through the implementation of structured quantitative and qualitative analyses.

ACKNOWLEDGEMENTS

Realising this PhD thesis was a long and laborious journey. Looking back, now I can tell that it would have not been possible to carry out this research without the help and support from those which have been by my side throughout the course of this research.

First of all, I would like to thank my industrial supervisor, boss and friend, Dr. Pablo Bermell-Garcia for his support and making the journey much easier to complete. Pablo, I am really pleased to work in your team and I really want to thank you for your advice and faith in me.

I would also thank my academic supervisor, Professor Tiwari Ashutosh, for his guidance. Ash, thanks a lot for your constructive feedback in the past three years and being always so friendly.

Thanks to my friends in Bristol and Spain for their unconditional support. James and Luis I will always remember your support which really helped me to complete this research. Rafa and Rubens, thank guys for supporting me defying the distance.

Finally, I just want to dedicate this work to my family, parents and sisters, for their unconditional support and especially my wife. Leidy Jazmin, you are the one, the happiness, my life partner. I have no words to thank you for all what you every day does to make me feel happy no matter how good or bad the day was.

A little more persistence, a little more effort, and what seemed hopeless failure may turn to glorious success.

Elbert Hubbard

LIST OF PUBLICATIONS

S. Quintana-Amate, P. Bermell-Garcia, A. Tiwari, Transforming expertise into KBE tools: A survey of knowledge sourcing in the context of aerospace industry, Accepted manuscript (unedited version) available online: 4-APR-2015 DOI information: 10.1016/j.knosys.2015.04.002, “Knowledge-Based Systems” journal, 2015.

S. Quintana-Amate, P. Bermell-Garcia, A. Tiwari, L. Balcazar, A new Knowledge Sourcing Framework to Support KBE Development, In: Proceedings of the 20th International Conference on Engineering Design (ICED15), Milan, Italy, 27.-30.07.2015.

J. Heilala, R. Ruusu, J. Montonen, S. Vatanen, P. Bermell-Garcia, S. Quintana Amate, C. Iwhiwhu, C. Kavka, F. Asnicar, A. Armijo, S. Scholze, M. Insunza, Collaborative product, process and service development with eco process engineering system. The sixth Swedish Production Symposium, SPS14, 16 - 18 September 2014, Göteborg, Sweden, Proceedings of The 6th International Swedish Production Symposium 2014.

J. Heilala, M. Paju, J. Montonen, R. Ruusu, M. Sorli, A. Armijo,P. Bermell-Garcia, S. Quintana Amate, Discrete Part Manufacturing Energy Efficiency Improvements with Modelling and Simulation, Conference: International Conference on Advances in Production Management Systems (APMS). "Competitive Manufacturing for Innovative Products and Services", At Rhodes, Greece, Volume: C. Emmanouilidis, M. Taisch, D. Kiritsis (Eds.): APMS 2012, Part I, IFIP AICT 397, pp. 142–150, 2013.

J. Heilala, M. Paju, J. Kiirikki, R. Ruusu, J. Montonen, P. Bermell-Garcia, S. Astwood, K. Krishnamurthy, S. Quintana Amate, Simulation with sustainability aspects in the manufacturing system concept phase, Winter Simulation Conference, 2012.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

FREQUENTLY USED ABBREVIATIONS

AI	Artificial Intelligence
AFP	Advanced Fibre Placement
AGI	Airbus Group Innovations
ATL	Automated Tape Laying
CAD	Computer-Aided Design
CAE	Computer-Aided Engineering
CBR	Case-Based Reasoning
CFD	Computer Fluid Dynamics
CFRP	Carbon Fibre Reinforced Plastic
CMS	Content Management System
CV	Cross-Validation
CommonKADS	Common Knowledge Acquisition and Documentation Structuring
DFM	Design For Manufacturing
DoE	Design pf Experiments
EC	Expert Collaboration
EI	Expert Involvement
EIN	Expert Interaction
ED	Engineering Design

EKR's	Enterprise Knowledge Resources
eLBD	Engineering Learning By Doing
EKM	Engineering Knowledge Management
KBE	Knowledge-Based Engineering
KC	Knowledge Capture
KD	Knowledge Discovery
KE	Knowledge Elicitation
KLC	Knowledge Life Cycle
KM	Knowledge Management
KMO	Knowledge Modelling
KNOMAD	Knowledge Nurture for Optimal Multidisciplinary Analysis and Design
KSF	Knowledge Sourcing Framework
LSPC	Lightning Strike Protection of Composites
MAE	Mean Absolute Error
MFG	Manufacturing
ML	Machine Learning
MOKA	Methodology and Software tools Oriented to Knowledge-based engineering Applications
REX	Return of Experience

RMA	Rules Management Application
RMSE	Root Mean Squared Error
RSM	Response Surface Method
UML	Unified Modelling Language

1. Introduction

1.1. Research context

The increasing competitiveness in the aerospace industry is forcing organisations to seek their profits beyond manufacturing (MFG). As a consequence, aerospace companies are changing from being *providers of products to providers of Product Service Systems (PSS)* [1]. This new business model involves shifting efforts from manufacturing activities to those related to service systems such as Maintenance, Repair and Overhaul (MRO) [1],[2].

In this direction, aerospace organisations are increasingly assigning MFG tasks to their suppliers, e.g. “Power8” established by Airbus in 2007 was fostering the use of suppliers to realise certain manufacturing processes. By adopting this strategy, Airbus shared the risks while raising its profits derived from those MFG processes [3]. In parallel, aerospace organisations are focussing their efforts on PSS [4] –e.g. *Power-by-the-Hour* approach proposed by Rolls-Royce, aiming to increase the number of MRO tasks which have been traditionally carried out by airline companies (e.g. Lufthansa Technik, Monarch). A clear example of this business trend is the acquisition of aviation maintenance companies by aerospace companies – e.g. Vector Aerospace bought by Airbus in 2011.

This new trend adopted by aerospace organisations together with the knowledge gained by the author as a result of performing the PhD while working in an aerospace organisation led to the identification of two major risks for future MFG. The first risk is associated to the reduced availability of MFG experts in the market. In many cases this implies the loss of relevant knowledge for the traditional aerospace leaders, most of this knowledge is now in hands of the few suppliers responsible for certain MFG activities. The second risk is the inefficient use of the knowledge for the development of future products and improvements, this may take form of: (i) decrease in the existent engineering models (defining structure, behaviour of a system); (ii) increase of raw data from different manufacturers which require the use of advanced data mining tools in order to capture the required knowledge. In order to minimise the impact of these risks, the aerospace industry needs to source expert engineering knowledge in areas ranging from manufacturing to maintenance.

Knowledge-Based Engineering (KBE) has been traditionally used to source engineering knowledge by integrating software and expertise, thus automating repetitive tasks and speeding up the engineering design process. The notion of “knowledge sourcing” here refers to the capacity to capture, retain and reuse engineering knowledge in KBE applications (see APPENDIX A for terminology).

To adequately perform the knowledge sourcing process it is a must to carry out efficient capture, manage and reuse of engineering knowledge [5]. In this regard, many of the existent KBE applications are not fully prepared for the industrial needs in a context where elicitation and management of both explicit and implicit knowledge have become a key feature to carry out fast and robust engineering [6]. In fact, it has been reported that designers spend around 30% of their time looking for information that is already available [7]. Thus, KBE tools will be effectively useful for the industry when designers can focus only on those tasks that add value to the innovation, while the knowledge is accessible and captured through the tools. In doing so, the time required to obtain an optimal design could be significantly reduced, KBE would then be supporting designers to quickly solve knowledge intense problems.

In this context, beyond the mere automation benefits of KBE, some authors have argued the value of this technology as a flexible and adaptable method to capture, manage and reuse engineering knowledge [5] –current KBE tools encode engineering tasks through IF…THEN rules. As a result, research work has attempted to decouple the automation aspect from the knowledge management side of KBE by adopting a strategy where the KBE software application works as an inference engine over a set of engineering rules managed in an external database [8]. Extended research on this concept considers KBE applications as self-contained resources, thus integrating the access to the knowledge and the automation tools [8],[9]. Despite the progress made in the last decades to transform KBE to deliver automated knowledge-based capabilities, a set of challenges remain to be addressed by the research community in order to achieve an extended KBE development process as required by the industry [9]:

- **Robust and reliable methodology**: KBE tools for industrial applications require methodological support to enable the capture, retaining and reuse of engineering knowledge in a systematic way. A successful methodological implementation could be applied beyond the mere automation of repetitive tasks, aiming at acquiring a sustainable stream of knowledge and to adapt to the needs and changes of the industry. This would also open new avenues in helping KBE tools to move from case-based solutions to methodology-guided projects. Currently, when developing KBE systems it is not frequent the use of already existing methodologies, this being one of the commonly existing challenges.
- **Efficient knowledge sourcing**: Another major bottleneck for the existence of fully automated KBE solutions is the way the knowledge is being captured –task encompassed in the knowledge sourcing process as described in APPENDIX A. Even though some effort has been dedicated into standardising the techniques to correctly extract knowledge from the experts, this process can be both time consuming and exposed to subjectivity, which in the long run may produce biased KBE tools. Therefore the effective sourcing of data, information and knowledge has important implications with respect to the quality of the resulting applications. In this context, the appropriate use of Artificial Intelligence (AI) could potentially deliver a more effective knowledge sourcing process by acquiring expert knowledge more efficiently. By doing so, the time required to extract expert knowledge would be reduced, hampering the use of KBE in the industry.

The research interest in this PhD stands at expanding the applicability of KBE technologies beyond the codification of rules with the mind set into their exploitation within industrial automated applications through an extended KBE development process. More precisely, the effort on this thesis will gravitate along:

- The sourcing of engineering knowledge through the use of automated reasoning tools and methods emerging from the Artificial Intelligence (AI) research field [10] [8].
- The methodological support to systematically manage engineering knowledge delivering a more reliable and robust KBE framework [8],[9],[11],[12].

1.2. Industrial motivation

Within the context of the aerospace industry there are industrial challenges associated with the sourcing of knowledge to support decision making in design which became apparent to aerospace engineers, but need to be substantiated and extracted for future developments. Companies such as Airbus have implemented a different research programme to address this problem focused on the design, test and validation of future methods and tools to provide more efficient engineering knowledge. This thesis was particularly motivated by one of this programme that was carried out within the Airbus Group Innovations named Return of Experience (REX).

Within the REX program, the focus was: (i) on the capture of knowledge from downstream engineering expertise (i.e. manufacturing, maintenance); and (ii) on the delivery of this knowledge to upstream phases of the Airbus products and services (i.e. design offices).

An example of the KBE methods within REX was the transfer of manufacturing lessons learned from the A350XWB aircraft program into the new design tools for future composite aircraft design; these tools were later used for the A30X aircraft concept [9]. Furthermore, in the design for manufacturing context, the REX program elicited the need for the sourcing of engineering knowledge and eventually for an efficient tool that would make knowledge explicit to future engineers.

Indeed, the tools for REX had a double folded interest on tackling specific issues existing in traditional engineering design practice at aerospace organisations. On one hand, REX aimed to address the prevalent lack of knowledge at early design stages, intrinsically related with the complexity of the products. This lack of knowledge at the conceptual engineering design stage has negative side effects and may derive into engineers being forced to make decisions based on low fidelity models. On the other hand, the implementation of novel aircraft programs (e.g. A30X) may as well use novel manufacturing technologies which have not previously been extensively used and may still be under development.

These two issues represent real difficulties when it comes to establish design requirements and design solutions. For example, in this situation, some of the parameters of the models used to estimate the manufacturing time and costs have a short life and need to be constantly updated, thus substantiating the need of an efficient knowledge sourcing from the very early stages.

In summary, the REX program is the ideal platform to establish best practices on how to capture and exploit knowledge through KBE methods and tools. In parallel, the complexity present in the definition and application of these tools in this early design context raises some challenges associated with the need to increase the efficiency of the knowledge sourcing practices. However, given the high impact of the design choices made at early stages, there is significant interest from aerospace companies to increase the knowledge sourcing for early knowledge exploitation. This thesis aims to contribute in this direction by building further understanding on the role of knowledge sourcing in the KBE research literature

1.3. Research aim and objectives

The aim of this research is “to develop a new and more efficient Knowledge Sourcing Framework (KSF) that supports KBE development and enables a better link between knowledge sourcing and KBE tools”. A list of specific objectives for the research which are required to fulfil this main aim is shown below:

1. To identify the practice of knowledge sourcing in the industry.

2. To help engineers in identifying AI tools that are most appropriate for their particular problem.
3. To enable efficient knowledge acquisition by exploiting AI tools capable of generating automated KBE rules from data assets.
4. To enhance KBE reliability by the assessment of KBE rules, allowing the designers to identify the quality of the results obtained.
5. To manage engineering knowledge separately from KBE applications by systematically storing KBE rules into a human readable format.
6. To adopt a methodology to systematically source engineering knowledge.
7. To validate the proposed KSF using two use cases shown in section 1.5 ("Research methodology").

On one hand, the aim of developing a new and more efficient knowledge sourcing approach will be achieved by adequately performing the first three objectives listed above. On the other hand, carrying out objectives number 4, 5 and 6 will provide a better link between knowledge sourcing and KBE. The last objective encompasses the realisation of two case studies and it addresses the validation of the research targets.

In conjunction, all the objectives aim to tackle important aspects represented by the research problems presented in previous sections of this chapter, thus pointing towards a more convenient solution for KBE in the context of the aerospace industry.

The achievement of the research objectives will confirm the hypothesis formulated in this work. This hypothesis assumes that a generic and extended KBE system, which delivers a more efficient sourcing of engineering knowledge, can be achieved by integrating Artificial Intelligence (AI) tools and methodological support (EKM techniques). The combination of these two elements will allow to systematically manage the knowledge (using the KM methods and tools) efficiently captured and modelled (employing AI algorithms and expert involvement).

1.4. Research approach

According to the literature, two main research design strategies are widely used: fixed and flexible designs [13]; also known as quantitative and qualitative research designs [14].

A fixed or quantitative design strategy uses mainly data in a numerical format. In this type of approach the details of the study such as the procedure used to collect and analyse the data, and the details of each case are defined prior to the data collection process. In contrast, a flexible or qualitative design uses data in the form of words [13] usually collected by interviewing experts or by observation. In this kind of approach the research questions and assumptions can evolve as data is being captured. Some additional characteristics of both approaches which were significant in the selection of the research design strategy are presented in Table 1 which are supported by [13],[14].

After analysing the characteristics of the two different research design strategies considered for this thesis, three main aspects of this research were determined to adopt the flexible approach as the research strategy:

- **Nature of the research objectives.** As commonly performed in researches adopting the flexible approach, this study explores in depth particular situations of interest for this work such as knowledge sourcing in the aerospace industry and artificial intelligence techniques for knowledge sourcing (objectives 1–3).
- **Use of a feasibility study to state the research hypothesis**. The realisation of a feasibility study –presented as the first use case of this thesis– enabled the identification of the research gap and supported the formulation of the research hypothesis using collected data. This characteristic is often found in researches following a flexible strategy.
- **The type of data to be collected**. In this research, the knowledge captured is provided by two different sources: experts and machine learning algorithms.

Consequently, it is important to be able to explore the interactions between experts and the advanced algorithms using qualitative analyses. Only the flexible approach provides a platform for such a natural combination of subjective and objective data.

Table 1. Characteristics of fixed and flexible research design strategies

Research design strategies	
Fixed or Quantitative	**Flexible or Qualitative**
It is based on precision (quantitative analysis) and control (sampling of experimental data). It is more reliable and easy to validate than the flexible approach.	Significant events or situations related to the particular research are studied in details
There is no discussion about the veracity of the facts proposed.	The veracity of the facts can be dependent of the context.
Hypothesis is stated before the data is collected.	The hypothesis is stated using collected and analysed data (using a qualitative analysis). It often employs a feasibility study to define the initial research hypothesis.
Researchers using this approach use only a few variables and many cases.	Researchers using this approach use many variables and only a few cases.

1.5. Research methodology

The proposed research methodology was defined according to the design strategy selected. It is divided in six stages to accomplish the research outcome of developing a new knowledge sourcing approach. The research methodology is reproduced in Figure 1; it shows the relationship between some of the research steps and objectives. Considering the characteristics of this research and following the appropriate guidelines, the steps followed in the research methodology are: problem description, literature search, research gap analysis, framework development, use case validation, reporting conclusions and discussion.

- **Problem Description:** In this stage the research problem is defined. Additionally, to support further investigation in the subject corresponding to this work a feasibility study was carried out. The feasibility study, which became the first case study of this work (see chapter 6), encouraged the author and the “Airbus Group Innovations” research team to proceed with the PhD project.
- **Literature Search**: This stage is associated with the literature review process in which related research concerning the topic of this work is further investigated (see chapter 2 and chapter 3). In this case three different areas are related to the research topic: EKM, AI and engineering design. The analysis of the literature associated to the research areas enabled to narrow down the research and define two main research streams. The evaluation of papers classified within these main research streams permitted the identification of the practice of knowledge sourcing in the industry (see chapter 3) and the identification of a set of research trends which were essential in the research gap analysis phase.
- **Research Gap Analysis**: The literature search process provides evidence of the main research issues that need to be tackled leading to the identification of the research gap. Within the literature review stage, the research gap is analysed prior to develop the methods and tools required to achieve the research outcome. An expert assessment process was carried out to further support the research gap previously identified. The assessment of a group of experts combined with the highlights obtained from the literature search enhanced the identification of research trends, opportunities and priorities, thus leading to the confirmation of a research gap on knowledge sourcing (further explained in section 2.4).
- **Framework Development**: At this stage, methods and tools required to achieve the research aim were developed. Prior the development of the knowledge sourcing framework, a study aiming the detailed review of industrial examples of methodologies focused on sourcing knowledge in the context of the aerospace

industry was carried out. This study presented in chapter 3 enabled the identification of a set of research limitations representing the key barriers stopping KBE of being a widely used methodology in the aerospace sector, thus helping the author to develop KBE methodology considering the characteristics of the industrial context. This was followed by the realisation of a Knowledge Sourcing Framework (KSF) –described in chapter 4– which is divided in five main tasks. These tasks are directly related to the six first research objectives of this work as shown in Figure 1 . The proposed KSF employs state-of-the-art machine learning algorithms to provide new knowledge about knowledge-intensive problems. It also uses a established methodology to develop a more robust KBE system capable of managing the knowledge life cycle efficiently and source of engineering knowledge in a systematic manner.

- **Use Case Validation**: Two use case studies (chapter 6) were used in this phase to validate the knowledge sourcing framework and prove the foundations of this research. These case studies are: MFG time estimation of aircraft wing cover designs and estimation of lightning strike effects (nut cap pressure prediction). The validation process is achieved by realising quantitative and qualitative analyses. Qualitative evaluations using the intervention of experts in charge of evaluating the framework and the AI model generated, and a quantitative analysis in charge of studying the results (time or pressure predictions) provided by the AI rules.
- **Reporting Conclusions and Discussion**: Once the use cases were validated, reporting conclusions and discussion processes were carried out (presented in chapter 7). This stage discusses the key findings of this research including their evaluation regarding their quality, generality and applicability. Moreover, limitations and future work associated to the developed methodology are described. Finally, a summary of the performed framework and the evaluation of the objectives achievement are presented.

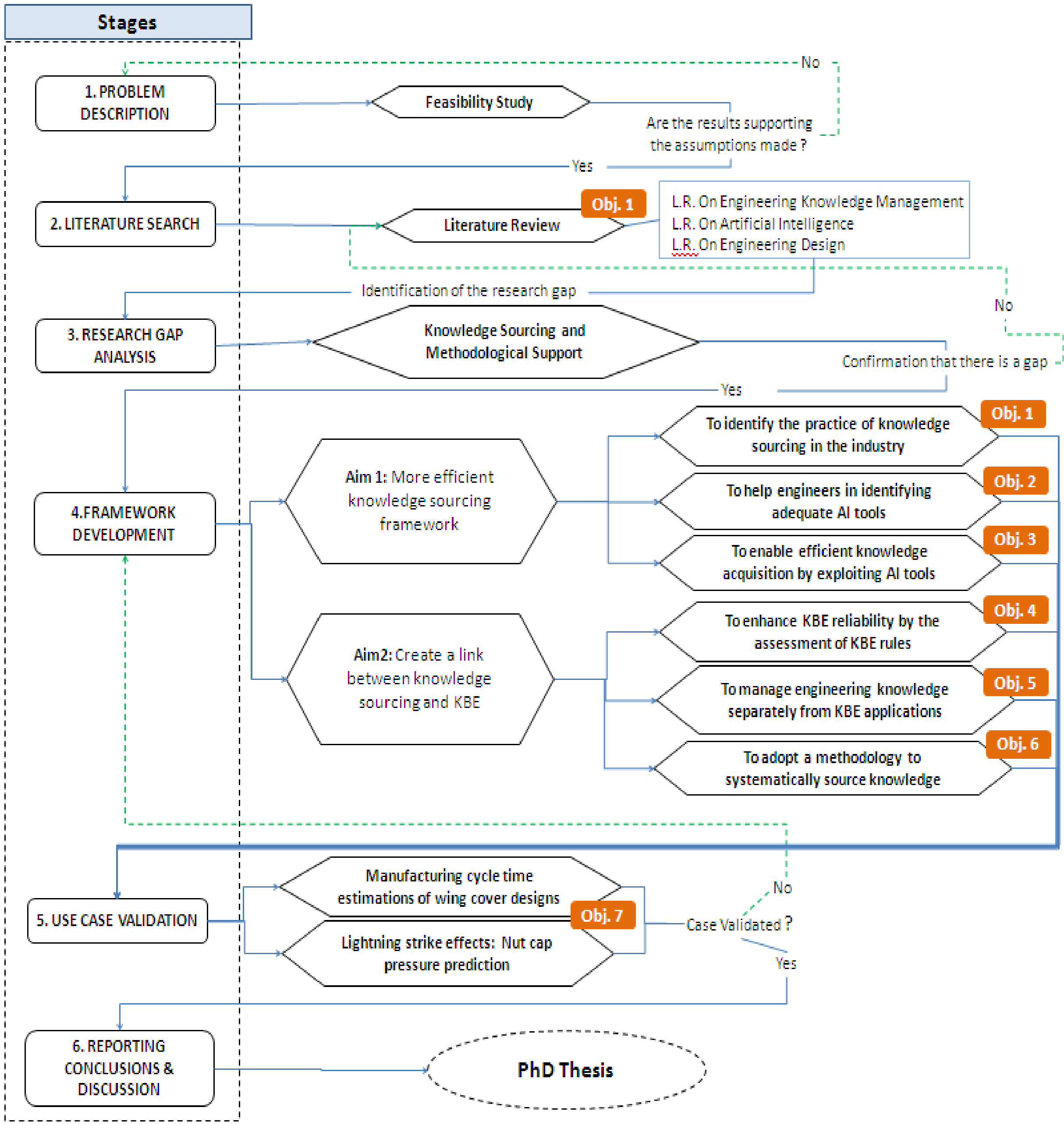


Figure 1. Research Methodology

1.6. Contribution to Knowledge

The outcome of this research is the creation of a KSF which supports KBE development. Two elements of the KSF are considered as the main contributions of this study. The first contribution is the use of AI tools and expert knowledge which provides a faster knowledge acquisition approach and increases the AI method reliability. It is fast by factoring in data-oriented sourcing and reliable by enabling expert intervention in different phases of the methodology. The second one is the embedment of AI applications within a robust KBE system to efficiently manage the complete knowledge life cycle.

In this study, Knowledge sourcing responds to the following three elements of the extended KBE development process suggested by the research community [6],[9]:

1. **Creating new knowledge:** Enable knowledge sourcing not only from experts but as well from company data assets in which often relevant knowledge is hidden.
2. **Retaining the knowledge:** Enable the sharing and the modification of the knowledge.
3. **Making the knowledge ready to be exploited into KBE applications:** Capturing knowledge from experts and representing this knowledge in a human- and computer-readable language. It allows the exploitation of knowledge across different engineering problems.

The realisation of the KBE elements highlighted above by integrating AI techniques and the required methodological support will enable the exploitation of automated reasoning methods to generate KBE rules.

To achieve the extended KBE development process, a fourth element in charge of creating a better link between knowledge sourcing and KBE is needed. This element encompasses methods and tools representing the methodological support that effectively integrates the components of the KBE framework and manages the engineering knowledge captured in a systematic manner. Therefore, the supporting methodology will contribute to realise the fourth element of the extended KBE development process:

4. **Integrating KBE applications into engineering workflows.**

As a result of realising this research a more efficient extended KBE system was developed allowing engineering design teams at industry to:

- **Reduce costs of the knowledge capture process:** Capturing knowledge not only by interviewing experts but also by automating its extraction from data sources to generate KBE rules that otherwise:
    - ➢ Would have been extracted only from experts in high cost knowledge capture sessions.
    - ➢ Would have missed correlations from large amounts of data that are often missed by experts.
- **Solve knowledge intensive problems:** Finding correlations in data in which knowledge is hidden, delivering a solution to those complex problems in which knowledge is limited.
- **Automate repetitive tasks:** Automating engineering design processes in traditional KBE but based on robust provenance of the knowledge.

1.7. Thesis structure

This section presents a summary of each chapter with the aim of familiarising the reader with the contents of this research. In addition, Figure 2 shows the relationship between the thesis chapters and the main steps followed in this research providing the reader a better understanding of the thesis.

- Chapter 1 starts with a description of existing challenges in the context of engineering design that led to the formulation of this research hypothesis. This is followed by the highlight of the industrial aspects that motivated the realisation of this work. After that, research scope, objectives and research approach are introduced. Finally, the chapter ends with the specification of the methodology followed to achieve the defined objectives, and the contribution to knowledge of the research.
- Chapter 2 explores relevant literature in this study, describing the process followed to identify and confirm a research gap on knowledge sourcing.

- Chapter 3 provides two industrial examples of knowledge sourcing capabilities presenting the research limitations previously identified which represent the key barriers stopping KBE of being a widely used methodology in the aerospace sector. The study presented in this chapter further motivated the author to carry out this PhD thesis and helped in the development of an extended KBE considering the characteristics of the aerospace context.
- Chapter 4 describes the procedure followed to develop the knowledge sourcing framework.
- Chapter 5 describes the achievement of the research objectives related to the knowledge sourcing framework implementation. More precisely, each section of this chapter shows clearly how the research target under study has been reached. Moreover, to provide the reader with a better understanding of the methodology proposed in this research a simplified use case using non real data has been implemented within the knowledge sourcing framework developed.
- Chapters 6 and 7 introduce two case studies instantiating and further explaining the methodology presented in chapter 4.
- Finally, chapter 8 discusses the research findings outlining their quality, applicability and generality. This is followed by a description of the limitations of the methodology proposed and future work recommendations. Finally, the research conclusions are underlined.

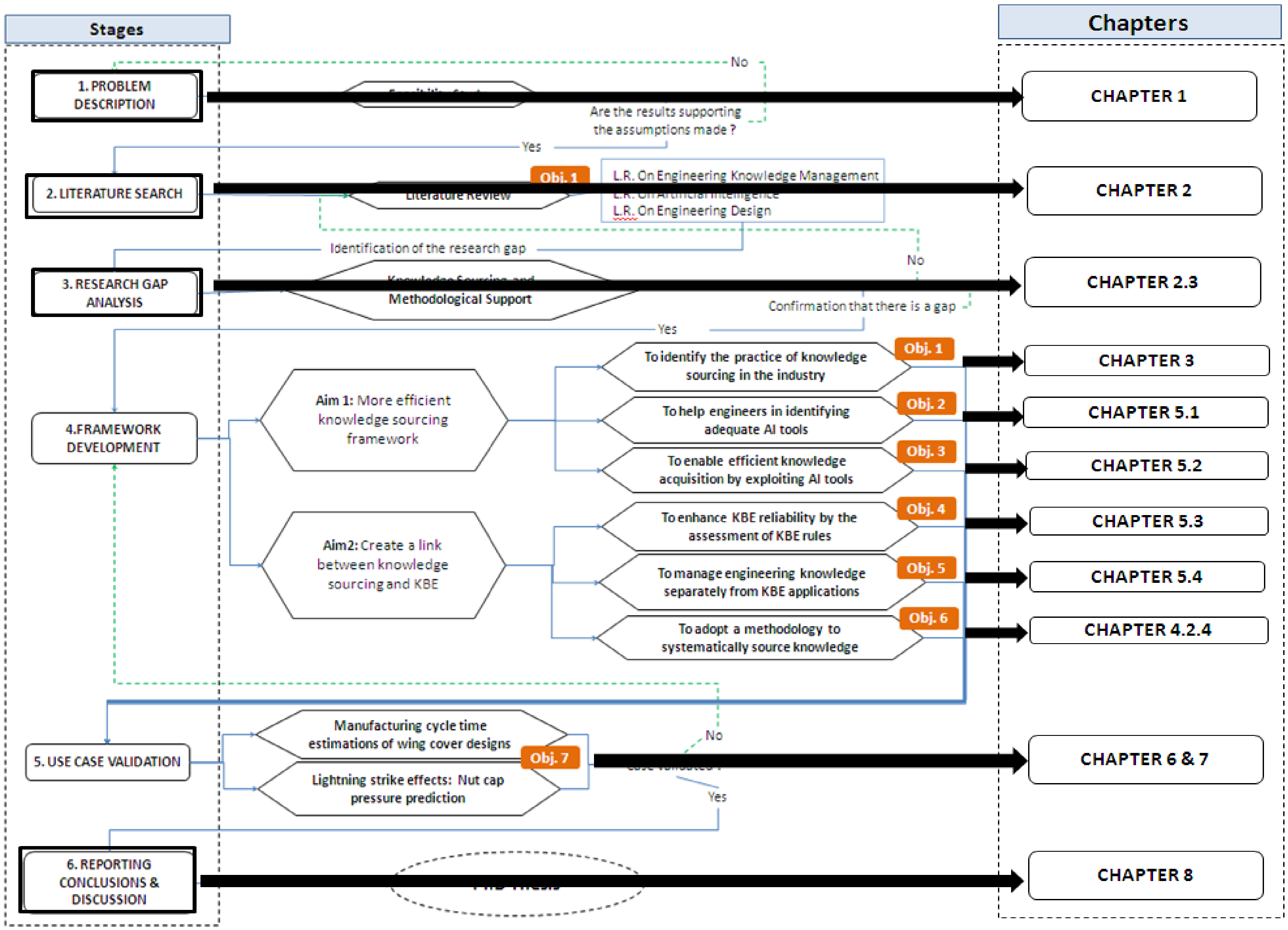


Figure 2. Relation between thesis chapters and main steps of the research methodology

2. Literature review

This chapter reports the literature survey which is relevant for the understanding of knowledge sourcing. Special attention is paid to knowledge-based engineering systems supporting the knowledge capture, retention and reuse. Research interest is also in AI technology for knowledge capture and its integration with knowledge management methods.

The purpose of the literature review is double folded. On one hand the author wants to obtain some knowledge about the domains which are significant to this research. On the other hand the literature search is focus on identifying and analysing other works related to this study.

The main outcomes of the literature review achieved through the collection and analysis of research studies are: (i) a basis for classifying and positioning this research among other industrial examples; (ii) identification of key research trends which analysis led to the validation of a research gap on knowledge sourcing.

This chapter is structured as follows. First, a brief description of the survey approach followed in this research is presented in section 2.1. This is followed in section 2.2 by the description of the theoretical foundations of this study. Section 2.3 provides a clear view of how the findings are obtained through the detail analysis of the literature lead to the realisation of this research work. The identification of key research trends and the analysis of the research gap are shown in section 2.4. This is followed by the description of the research scope in section 2.5. Finally, a summary of the work described in this chapter is presented in section 2.6.

2.1. Survey Approach.

Identified challenges to deliver an effective and a more efficient knowledge sourcing capability led the author to the definition of a research scope. The research scope is focus on two main topics: (i) the sourcing of engineering knowledge through the use of automated reasoning tools and; (ii) the methodological support to manage engineering knowledge and making it available to KBE. These topics are related to the areas of knowledge management and artificial intelligence. Moreover, engineering design is an area of interest for this research as it is the context where this study is applied.

The classification and analysis of researches associated to the research scope resulted in the definition of two research streams which represent the background (foundations) of this study. The background of this research is on:

- Knowledge management in engineering design.
- Role of artificial intelligence for knowledge management in engineering design.

The analysis of literature associated to the research background in form of tables and charts led to the identification of key research trends. Finally, the study of the key research trends in the research gap analysis stage ended with the confirmation of a research gap in knowledge sourcing (Figure 3).

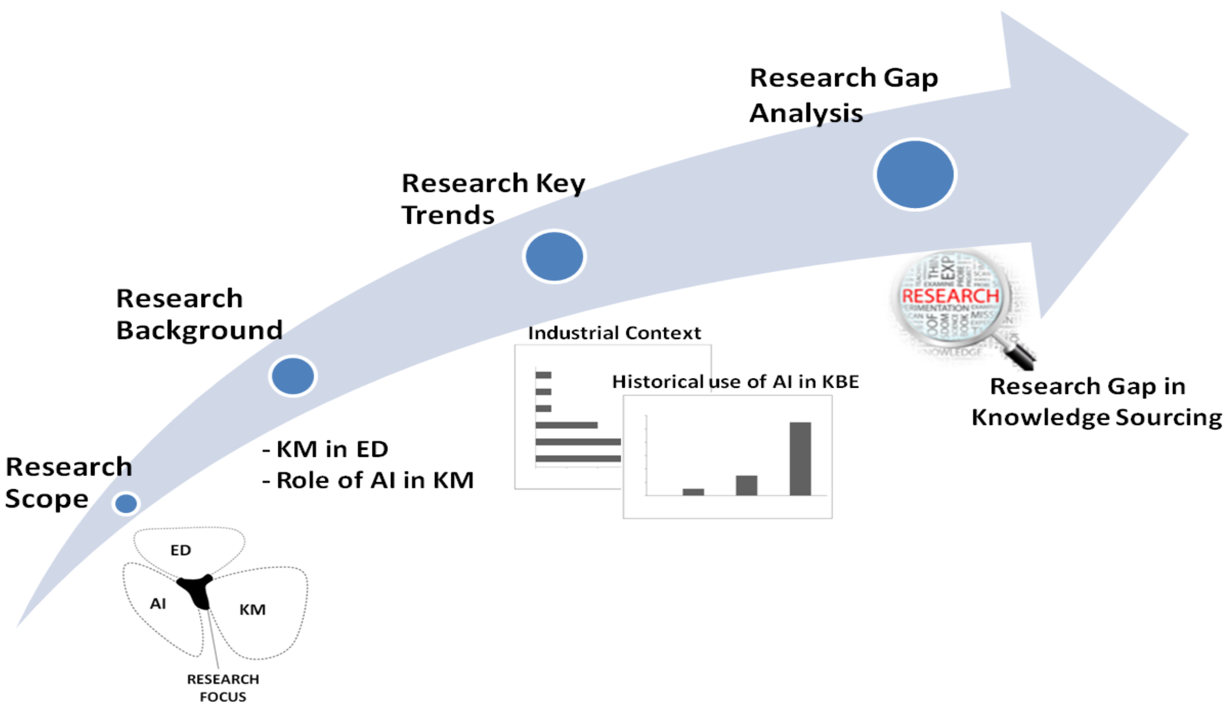


Figure 3.Survey approach

2.2. Background

This section presents a review of the literature associated to the main research streams of this study. These were identified through the analysis of researches related to the research scope topics. More precisely, literature search was realised on:

- Knowledge management in engineering design.
- Role of artificial intelligence for knowledge management in engineering design.

2.2.1. Knowledge management in engineering design

Engineering design is usually defined as a systematic process where customer needs are the performance specifications and functions used to obtain optimised design solutions [15]–[20]. In the engineering design process there are four main steps [15],[21] (Figure 4):

1. **Clarification phase:** Information captured encompassing design needs and constraints to be included as part of the solution.
2. **Conceptual design:** Definition of the functional structures, searching for suitable a solution principle or concept. This is the most important step of the engineering design process caused by the fact that around 80-90% of the production costs are determined at this stage [22].
3. **Embodiment design:** From the principle solution, the design is realised taking into consideration the technique, economic requirements and constraints.
4. **Detail design:** At this stage further information is specified (such as the surface properties or part material) in order to complete the design description. Moreover, the technical and economical feasibility are re-checked.

It was believed that 100% of the designer's time was spent in the engineering design process steps [15]. However, later studies have proved that part of the designer's time is used in non-technical tasks such as looking for information and attending meetings.

More precisely, by observing 27 designers –while they were working– it was concluded that 45% of the designers' time is dedicated to perform the steps of the engineering design process [23]. Moreover, it was observed that 21% of designer's time was spent in tasks related with the management of engineering knowledge [23]. These facts support some author's studies defining the engineering design process as a set of knowledge intense activities [24]. Consequently, more recent studies have been focused in capture the time spent by designers in knowledge management activities [22]. For instance, [25] concluded by observing 13 engineers work that 33% of their time was spent in capturing knowledge and distributing it, including 9% on meetings.

Each stage in the engineering design process requires technical knowledge and experience to be effectively captured, modelled and reused. This enhances the quality of the product development procedure while reducing time and costs of the production process [24]. A potential solution well known by the research community to address the effective elicitation (capture), model and reuse of engineering knowledge is the use of knowledge management methods and tools [26]–[28]. To facilitate the understanding of the literature review followed in this thesis, concepts at the interface between knowledge management and engineering design were analysed in details.

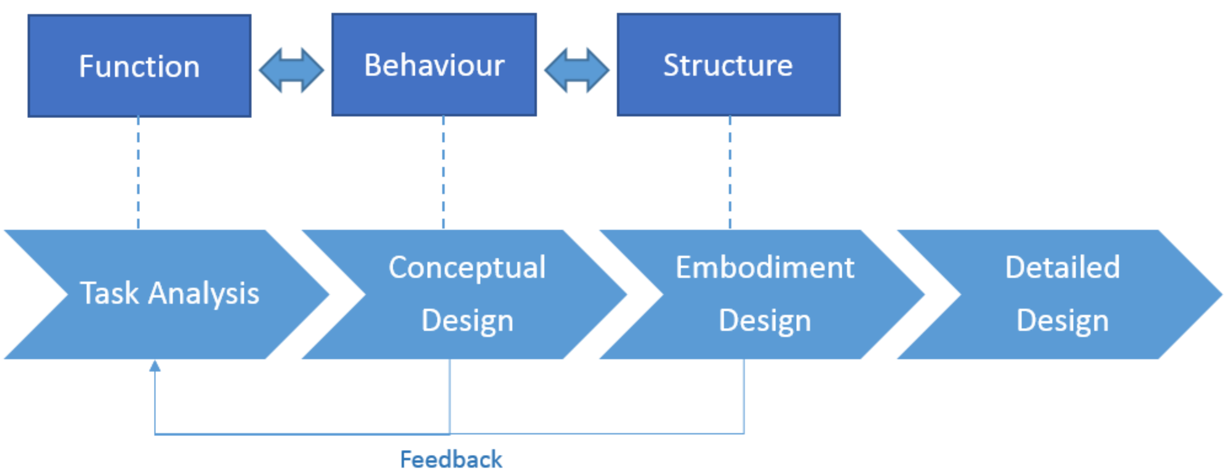


Figure 4. Engineering design phases based on [29].

Concepts between knowledge management and engineering design

a) Knowledge capture

In the knowledge capture process, engineering rules are generated by the extraction of knowledge from human expertise [30]. These rules are usually stored in a database to facilitate its later representation and reuse. The knowledge capture process is often carried out using well known method such as interviews, observing but also using other techniques such as card sorting, protocol analysis and laddering [31],[32].

- **Card sorting:** This method requires experts to write down on a set of cards objects, experiences or rules. Then they are asked to group them and explain to the knowledge manager the reasons or similarities between cards that take them to classify them in such a manner. The use of this technique enables experts to structure their ideas and therefore elicit expert knowledge.

  Card sorting has been successfully applied to explore attitudes towards the use of expert systems for aids diagnosis [33] and to facilitate in-depth exploration of teens' experiences of asthma [34].

- **Protocol analysis:** To implement this technique, experts are think-aloud when solving an established task. All what the experts say is recorded and transcript to be further analysed using a specific procedure. The way the information is transcript must facilitate its access and retrieve by the participants. Moreover, in some cases the knowledge manager in charge of the activity writes down the physical movement of the experts. This is known as 'motor protocols' and can be useful to capture some additional expertise. After all the information is correctly stored and formatted experts are asked to go through this material to develop further on the key points that they considering more relevant to solve the proposed task.

  This type of technique has been successfully implemented in the development of expert systems [35] and providing a better understanding about the influence of social learning on student nurses' acquisition of Information and Communication Technology (ICT) knowledge and skills [36]. However, using protocol analysis can difficult experts trying to solve complex task to think-aloud about something he/she is not completely aware or don't know how to express [35].

- **Laddering:** This technique represents a way of gathering engineer's view of a specific problem or even life in general. It is a semi-standardized interview where the relationship between attributes of products or services, the consequences this attributes represent to engineers and the values which are accomplished by the consequences are shown. In this method the knowledge manager asks experts to answer if an attribute is relevant and why in case of affirmation. Answers are used by the knowledge manager as starting point for further questions. Finally a Hierarchical Value Map (HVM) is created by synthetizing the cognitive concepts captured in the laddering interview. Laddering provides an easy and systematic way of mapping the beliefs of the participants. It has been widely used in the sales management [37] and services marketing [38]. For instance, [39] used laddering to explain choice option attractiveness by beliefs about the choice option whereas [40] combined laddering with an neural network technique as a strategy for acquiring customer requirement patterns.

Knowledge capture methods can be divided into natural and contrived [32]. Natural methods are those in which the experts express their knowledge in a common and familiar way. Examples of natural techniques are interviews and observing. In parallel, contrived methods are defined as non-expert familiar techniques such as laddering or card sorting. All these methods require expert intervention in order to obtain a model that defines a particular engineering problem. However, other methods can be used to derive rules automatically [41],[42] such as automated learning algorithms.

Traditionally, Rule-Based Reasoning (RBR) has been widely used in the context of KBE to extract expert knowledge in the form of rules [43]. These rules were manually created by experts in the domain of the problem to be tackled. This method involved the requirement of experts in intensive time consuming tasks. In this context, research has been undertaken to minimise the impact on experts to free them and enable them to use their time in added-value activities. An example of this is the use of AI methods and more particularly the implementation of Case Based Reasoning (CBR). CBR uses past experiences in the form of "cases" to predict the behaviour of new cases. For example, [44] integrates an ontology with a CBR tool to provide better accuracy and reliability of FEM analysis results whereas [45] supports engineers to define an appropriate production scheduling by implementing CBR within a content management systems containing successful past experiences.

Although CBR is considered as a breakthrough in the knowledge acquisition problem acknowledged as the main bottleneck for the development of KBE systems, it stills requires a considerable amount of time of domain experts to express their ideas and weight the features considered in problem to be solved [46].

Additionally to the described techniques aiming the elicitation of expert knowledge, new methods and tools belonging to the entertaining context are gaining interest to elicit knowledge [47]. In this direction, Serious Games (SG) have been applied for knowledge acquisition areas such as production and logistic systems [48] and [49], manufacturing [50] and resource planning [51]. SG are games which primary objective has been mainly for learning but also for educational purposes and knowledge elicitation [47]. Regarding the use of SG for knowledge elicitation, [52] proposes a new approach for validating knowledge from a PLM using SG where experts are initially asked to provide an explicit assessment of the problem. The explicit assessment is later enriched by the users of the SG which provide implicit feedback rating. Although existing research on SG claim to provide an effective approach for knowledge elicitation there is a lack of rigorous evidence of its effectiveness [47].

b) Knowledge modelling

The main goals of KBE are first to automate repetitive tasks and second to retain the knowledge of a specific task for solving future problems in the same domain [31],[32]. To reach these objectives, knowledge must be properly modelled in order to ensure the quality of the rules extracted. These rules represent the required static knowledge (explicit representations of a domain) to be used within a reasoning process.

Knowledge modelling is a highly researched topic where many methods have been developed to adequately represent engineering design knowledge. Frame-based, logical formalisms, semantic networks, production systems and object oriented representations are some of the existent knowledge representation schemes. The last two schemas are known for being the most used into KBE systems [32]:

- **Production systems:** It refers to those techniques where IF...THEN rules are used to represent the model.

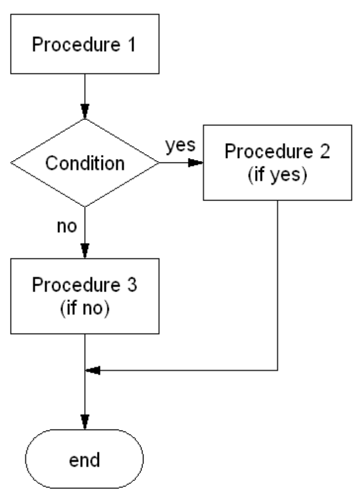


Figure 5. Procedural model.

- **Object-oriented representations**: It considers a problem as a set of associated objects that interact with each other [53]. An example of this is the UML (Unified Modelling Language) where a group of models and notations permit the suitable visualisation, specification, documentation and communication of objects that define the information system.

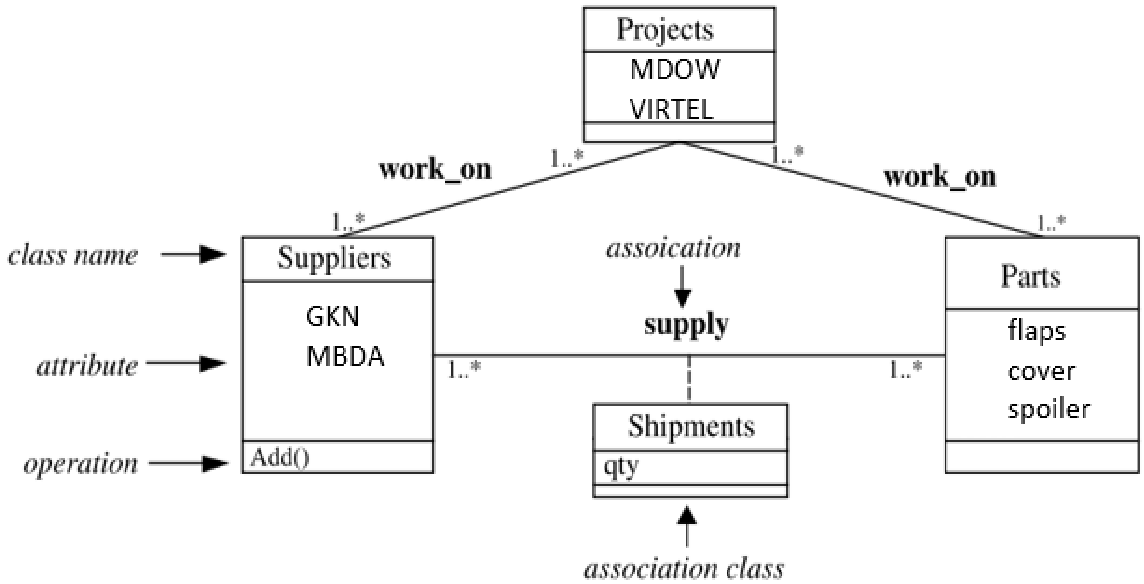


Figure 6.Object oriented model based on [53].

In short, production and object-oriented schemes represent two valid techniques delivering a methodology capable of generating a descriptive model of a design presenting its requirements and constrains.

c) Knowledge reuse

Knowledge reuse is defined as the process of sharing best practices including processes information and documentation aiming to support designers in solving knowledge intensive engineering problems efficiently [41],[42]. Some of the most common knowledge reuse techniques are [54]:

- **Computer Aided Design (CAD) and Computer Aided Engineering (CAE):** Parameter-driven applications usually based on geometric models. These tools have embedded engineering rules enabling the user to reuse previous knowledge models. Well known commercial implementations are: CATIA, ICAD and Pro-Engineer. These software implementations are widely used when a high level of optimisation is needed, thus delivering a solution to repetitive engineering problems.

- **Functional techniques:** Methods where knowledge structure is based on functional decomposition. In these types of techniques, flexible rules are used to define domain engineering problems. These rules are stored and managed in a knowledge based system as described in [11].
- **Matching methods:** Focused on the classification of new cases based on characteristics of preceding cases previously indexed. The most common technique is Case Based Reasoning (CBR) where a set of initial cases already known are indexed and, having a database of known cases or past experiences, a specific technique –such as the ones belonging to the artificial intelligence field– is used to find similar cases helping designers to make more informed decisions [46],[55],[56].

The knowledge reuse benefits are [57],[58]**:**

- **Duplication of efforts due to lack of knowledge retained**: When a new problem to be developed is similar to previous cases already performed, organisations are forced to spend efforts on gathering information that could have been easily transferred.
- **Limited resources:** The capture of new knowledge is acknowledged to be the major bottleneck of the knowledge sourcing process, thus an effective method is needed to reduce the development time of new products. This time could be reduced by efficiently reusing knowledge captured in past studies.

2.2.2. Role of AI for KM in ED

Artificial intelligence is usually defined by the research community as the art belonging to the field of computer science aiming to mimic human thinking, delivering a solution to problems such as decision making and problem optimisation [59]–[63]. The ultimate goal of artificial intelligence is to achieve with machines what if realised by humans the use of intelligence is needed [64].

In this context, Knowledge Management combines concepts from multiple disciplines, such as information technology and artificial intelligence to deliver a more efficient engineering design process [10]. It is acknowledged that AI provided tools are well suited for process optimisation but this can also be used to source expert knowledge [10],[65]. The role of AI in the knowledge management field can be summarised by the following two main contributions [10]:

- **Knowledge sharing:** knowledge is captured and stored in a database enabling the knowledge to be exploited across different engineering problems within the company.
- **Knowledge discovery**: AI-based methods can be used to source knowledge by looking for patterns in company datasets, thus obtaining some trends that could have been missed by experts creating new knowledge as a result.

A potential solution to tackle the engineering design challenges identified in this research and provide engineers with a solution to manage engineering knowledge more efficiently, is the use of knowledge management methods combined with tools belonging to the AI [46],[66]–[70].

2.3. Knowledge Sourcing Role in the KBE context

This section provides the reader with a better understanding of the work undertaken to source engineering knowledge effectively and more efficiently in the context of KBE and engineering design.

2.3.1. Background on Knowledge Sourcing

The complexity of the products in the engineering design context leads engineers to face with complex problems which requires knowledge sourced from different disciplines in order to provide an effective solution. Distributed and heterogeneous teams are required to work interactively as cross-functional teams with access to tools enabling intensive knowledge exchange [71] and [72].

In this direction, wikis started to be used within the engineering design context enabling designers to find the right people with the required skills to solve a problem. An example of this are the expert yellow pages which enable to quickly create a multidisciplinary team with the required skills to solve a specific task. Engineering wikis also enable to represent product or service functional requirements, geometry, constraints, rationale, etc. [73] proposes a Design Rationale editor (DRed) to facilitate the transfer of knowledge and the understanding of the “know how”. DRed enables the sourcing of the decision rationale at the time it was generated and it is currently used by an aerospace company. The users of the DRed software to create a graph of predefined nodes (issue, answer and argument) with directed arcs.

In addition, methods and tools were created in the engineering design area to facilitate the capture and reuse of the functional behaviour and requirements of a product. An example of an approach for functional modelling is the presented in [74] where primary/carrier flow relationship is presented using functional basis design. The creation of functional models is key for retaining knowledge while permitting the reuse of this knowledge across different engineering problems within an organisation [75].

Many research studies have been undertaken regarding the storage and management of geometrical data in engineering design [76]. In this regard, [77],[78] propose an object oriented model aiming at sourcing knowledge from CAD models for automated process planning which is capable of classifying and modifying geometrical features to cavities that can be manufactured.

Knowledge sourcing in the KBE context has been also carried out through existing knowledge elicitation techniques. These methods are described in section 2.2.1 and the knowledge extracted by their implementation has been traditionally hard coded within KBE applications also described in many cases known as expert systems [6].

[79] presents a KBE system using encoded rules (non-human readable) stored in a knowledge base to estimate MFG cost of CAD moulded parts whereas [80] also developed a KBE application considered as a “black box” (hard coded knowledge) using

CATIA V5 with the aim of estimate MFG cost and weight of a composite structure at the conceptual design.

KBE allows the automation of repetitive tasks therefore enabling experts to use their time in added-value tasks. To avoid knowledge leaks due to the complexity of the hard coded knowledge within the KBE applications, newer research studies have stored the knowledge or rules used by the KBE application in a human and machine readable format. For instance, [9] describes the development of a KBE application as a software service where knowledge is elicited through interviewing is stored in an external knowledge repository used by the inference tool. [81] Presents a KBE system that automatizes the creation of the Finite Element (FE) model for the efficient design creation of automotive body structures where the knowledge is kept in a human readable format within a knowledge repository easily accessed by the users of the tool. Keeping the knowledge separated from its application and stored within formal and informal models adequately managed in a knowledge repository enhances the knowledge retain and reuse.

However, these researches highlight as main limitation to implement KBE applications the knowledge elicitation phase which still represents a highly time consuming activity for the domain experts. In this direction, [46] discuss about the use of AI methods to considerably reduce the time employed in the expert knowledge extraction. Relevant papers for this PhD thesis aiming at using AI methods for a more efficient knowledge sourcing are described in details in the next section.

2.3.2. Review of Knowledge Sourcing for KBE

KBE is considered by many authors as much more than a mere system capable of automating some repetitive engineering tasks. [5],[6] review existent KBE applications highlighting the potential to use KBE to effectively capture and reuse engineering knowledge while automating many tasks in the engineering design process which are currently manually performed. Both researches point the main limitations stopping KBE to be often used where to tackle some of them it is required a more effective and efficient knowledge sourcing process. These limitations are: the need for a faster way of capturing engineering knowledge reducing the time required by experts in the domain and the need for a transparent system enabling traceability of the knowledge.

From this perspective, [82] presents a KBE system for generative model creation using CATIA and adopting MOKA methodology to manage the knowledge life cycle and therefore enhance the transparency and traceability of the knowledge. For that purpose, the work implements each step of the MOKA methodology focusing on the use of formal and informal models to facilitate the transfer and reuse of the knowledge captured using traditional knowledge elicitation techniques such as interviews and observation. With the same objective, [83] uses CommonKADS (Common Knowledge Acquisition and Documentation Structuring) methodology to formalize a knowledge repository where the knowledge is structured and stored facilitating its reuse across different engineering problems. The research highlights the flexibility of the methodology due to, a difference to other methodologies to develop KBE systems like MOKA, it uses predefined templates to annotate knowledge. Apart from the benefits claimed by [83] and [82], [84] presents an easier to implement methodology named as KNOMAD which enables the creation of KBE systems with improved knowledge maintenance and reuse due to the integration of an ontology. These research studies focus on the use of the required methodological support to manage the complete knowledge life cycle but are not focus on the efficient knowledge capture.

In this direction, several studies have aimed at implementing methods from the AI area to source engineering knowledge, delivering a solution to knowledge intensive problems. AI methods in engineering have traditionally used for optimization. In this context, [85] uses AI algorithms for logistics optimization whereas [86] describes the use of algorithms to solve a supply chain optimization problem by looking for patterns in data. Both research studies behave as “black boxes” having the knowledge used for the optimization process hard coded within the application. Moreover, these two researches mention the difficulties to select an optimal algorithm to solve a specific problem due to they are context dependant. To provide a solution to that they both propose an algorithm portfolio which enables running different algorithms in parallel or interleaving.

In parallel, [73] and [74] provide several examples using different algorithms to provide the reader with a better insight about the applicability of each type of algorithms.

More recently, AI algorithms and more particularly Machine Learning (ML) algorithms have been used for knowledge elicitation purposes providing a faster knowledge extraction process from company data assets. [70] proposes a rule-based knowledge system for retaining wall selection where the rules are automatically generated from data using Rule Induction (RI) method. RI method is a machine learning technique which enables the creation of explicit rules from data. The research integrates RI within a database management system to allow automated wall selection and generate a feasible solution for new retaining wall problems. The approach enables the knowledge automatically extracted to be ready to use by any KBE system. This is achieve thanks to the storage of the knowledge in an external knowledge repository. [89] presents a KBE system for advanced structural design using an AI algorithm which also helps in the intelligent interpretation of the results provided by the platform. The platform uses a knowledge base to facilitate the access to the knowledge and its reuse.

[90] also develops a KBE system using AI for automated knowledge generation focusing on the analysis task and how meaningful knowledge was elicited. He remarks that the data pre-processing is the most time consuming and challenging activity due to it is essential to perform several data mining tasks to ensure the quality of the data.

[69] apart from automatically generate a set of rules from data using AI techniques, proposes a methodology to manage the life cycle of the knowledge captured. The research delivers a framework capitalizing on past experiences by formalising domain knowledge and representing it using conceptual graphs; and extracting new knowledge using a machine learning algorithm. [68] developed a KBE system for raising awareness of the priorities and needs of the members of a team. This study uses data mining techniques for automated knowledge extraction and it highlights the importance of the data preparation phase and the need for an expert or group of experts to analyse the meaningfulness and coherence of the knowledge extracted. Moreover, the research concludes that it is required that each rules has a small number of features to keep its meaningfulness and expert guidance is often required to discover patterns in data. In

this direction, [91] presenting a KBE application for intelligent failure mechanism detection and [92] describing a KBE system for medical diagnosis propose two different KBE systems where expert intervention is required to review, modify and validate the rules automatically extracted by the AI algorithms. In addition, both frameworks provide the user with an interface to easily modify the AI rules.

In summary, the research studies discussed in this point show existing challenges between knowledge sourcing and kbe and provide some examples of how researchers are facing them. This discussion also supports the hypothesis of a research gap on knowledge sourcing in the KBE context. In order validate this assumed research gap a set of functional roles between KBE and knowledge sourcing have been defined and used to classify papers relevant to this research. These functional roles are described in details in the following points.

2.3.3. Definition of functional roles between knowledge sourcing and KBE

To facilitate the confirmation of the research gap, a review and analysis of tools aiming the efficient source of knowledge and the creation of a better link between knowledge sourcing and KBE was carried out. This activity led to the identification of a set of functional roles which are closely linked to the research challenges of this study. The defined functional roles are listed and described as follows:

- **Capture of Expert Knowledge:** Use of methods and tools to elicit knowledge from experts via interviews and data forms.
- **Access to knowledge by KBE tool:** It refers to implementations enabling the retrieval of information from an external data source which allows the knowledge to be exploited by a KBE application.
- **Knowledge lifecycle management:** Presence of methods and tools aiming the methodological manage of engineering knowledge along its complete life cycle.

- **Automated knowledge extraction from data:** Employment of tools able to reason over large datasets extracting engineering rules, constraints or other logical correlations from data.
- **AI extracted knowledge ready for reuse:** Use of tools to automatically codify rules from knowledge in a way that is directly reusable by KBE applications (knowledge is in a computable format) rather than only being employed by humans.
- **Advice on AI tool suitability:** It refers to the use of applications to advice practitioners on the selection of AI tools to perform automated reasoning tasks.

The first three functional roles listed above were identified through the review of the literature and considered as common KBE practices [5],[9],[28],[46],[66],[93]–[96]. In parallel, the fourth and fifth functional role were identified as common procedures when using AI knowledge-based tools to source engineering knowledge [67]–[70],[89]–[91],[97],[98].The last functional role defined was identified by aerospace experts as a useful activity that would considerably increase the reliability of the results delivered by an AI implementation, thus supporting the efficient knowledge sourcing process.

2.4. Key Trends and Research Gap Analysis

This section presents the trends in the literature that contributed towards the research gap validation. To ensure the achievement of meaningful and reliable research trends, a systematic data collection process was required. In this regard, this section starts with the description of the methodology followed from the initial definition of the research scope to the identification of research trends and research gap confirmation (sub-section 2.4.1). This is followed by the detailed description of the key research trends and the research gap analysis (presented in sub-sections 2.4.2 and 2.4.3 respectively). Finally, a discussion of the key findings obtained from the literature survey is presented in sub-section 2.4.4.

2.4.1. Data collection process followed

Two paths were followed to realise the data collection process. First, a structured literature search focused on the use of external databases such as Scopus, ScienceDirect, SpringerLink, IEE Xplore and Taylor&Francis was carried out. Moreover, Google web and Scholar were used in order to complement the literature search.

In addition to the systematic approach, an unstructured search of the literature was also performed. In this direction, references identified from the analysis of relevant papers were also reviewed. This procedure provided the author with a better understanding and consolidation of the main concepts and findings associated to this PhD thesis.

The starting point of the literature search was the definition of a set of keywords which are in agreement with the research scope. Therefore, this set of keywords –used in the initial phase of the literature search– was used to compose searches such as "Knowledge management & engineering design" and "Artificial intelligence & Knowledge management". A preliminary literature search in the areas of interest defined by the research scope led to the definition of the background topics. In order to reduce the number of research documents displayed on the browser –when carrying out an online search– and facilitate the finding of significant papers, several searches were realised using combinations of the keywords identified associated to the research scope.

For instance, when searching in Scopus for "Knowledge Management AND Engineering Design" more than 3.000 articles appear on the screen. Thus it became apparent the need of filtering the files browsed. As a general rule applied in this work, the content type was constrained to articles from books, journals and conferences.

When manually screening the results (ordered by their relevance) it was common that in most of the cases only documents in the first six pages (300 documents) were relevant for this research. After going through the first 300 documents on the search list, abstracts of those documents with titles concerning with the search performed were revised. Moreover, only those documents with relevant abstracts were stored and classified in a comparison matrix (described in sub-section 2.4.1.1).

The analysis of the papers obtained using keywords associated to the research scope led the author to the definition of the background of this study (research foundations).

The same procedure of manually screening documents obtained through queries linked to the research scope was also carried with keywords associated to the research background. In addition, due to the huge number of documents obtained in the search process, the implementation of filters was required. In this regard, the number of results was often considerably reduced when excluding documents containing “mathematics”, “business management” and “economics” as subject areas.

In this context, Table 2 presents the keywords (related to the research scope and background) which were used in the data collection process.

Table 3 shows an example of the data collection process followed using Scopus search engine.

Table 2. Keywords used in the literature search

Keywords	
Engineering Design (ED)	Knowledge Discovery (KD) / Elicitation (KE) / Capture (KC)
Knowledge Management (KM)	Knowledge Reuse (KR)
Artificial Intelligence (AI)	Knowledge Modelling (KMO)
Knowledge based engineering (KBE)	Expert Collaboration (EC) /Involvement (EI) / Interaction (EIN)
Knowledge Sourcing (KS)	Machine Learning (ML)

Table 3. Data collection using Scopus search engine

Keywords used in Scopus search	Initial number of documents	Number of documents after filtering	Areas excluded in the filtering process	Relevant papers to be further analysed
KM + ED	3.685	N/A		143
AI + ED	17.589	N/A		171
KM + AI	40.251	N/A		96
KM + AI + ED	1.578	1.093	Mathematics, Economics, Business Management	46
KS + ED	19	N/A		4

KS + AI	28	N/A		7
KS + AI + ED	3	N/A		1
KBE	1476	N/A		178
KBE + AI	555	442	Mathematics, Business Management	156
KBE + AI + KD	39	N/A		21
KBE + ML	78	N/A		16
KBE + KC + AI	25	N/A		13
KBE + KC + ML	3	N/A		3
KM + KR	1430	1.027	Mathematics, Economics, Business Management	55
KM + KMO	1304	918	Mathematics, Economics, Business Management	66
KM + KC	1547	1.080	Mathematics, Economics, Business Management	81
AI + (KD or KC or KE)	69.143	9.261	Mathematics, Social Sciences, Business Management, Computer Sciences	133
AI + (EC or EIN or EI)	103	N/A		45
ML + (EC or EIN or EI)	425	237	Mathematics, Social Sciences, Business Management	11
ML + (KD or KC or KE)	56.458	9371	Mathematics, Social Sciences, Business Management, Computer Sciences	60

Most of the relevant papers found using keywords derived from research scope and background topics were obtained using Scopus, Science Direct and Springer Link external databases.

After reading in details those papers classified in

Table 3 as “Relevant papers to be further analysed” more than 150 papers have been identified as important for this research. From those 150 papers including a description of a framework containing features closely related to the areas of KBE and knowledge sourcing where further analysed using a classification matrix (APPENDIX E) as described in section 2.4.1.1 which were used in the identification of research key trends and research gap analysis.

Documents used to identify the research background topics and key research trends were identified and classified as described in section 2.4.1.1 ("Data classification"). Together with the key trends identified, a list of functional roles associated to KBE and knowledge sourcing were employed in the research gap analysis in order to classify papers containing KBE applications.

The use of the defined functional roles in the classification and analysis of the literature facilitated the identification of research trends and opportunities. Moreover, the research priority levels of each functional role were considered and evaluated, thus increasing the reliability of the research opportunities and trends. The classification and analysis of relevant papers was realised by the combination of the following instruments:

- **Literature analysis and classification**: 63 research articles reporting on KBE implementations were analysed and classified within the functional roles defined by the author.

Expert assessment: 6 experts from various domains working in different organisations were selected to assess the importance of the functional roles identified in this study. The objective of the expert assessment activity was to support the research opportunities obtained through the analysis of the literature

2.4.1.1. Data classification

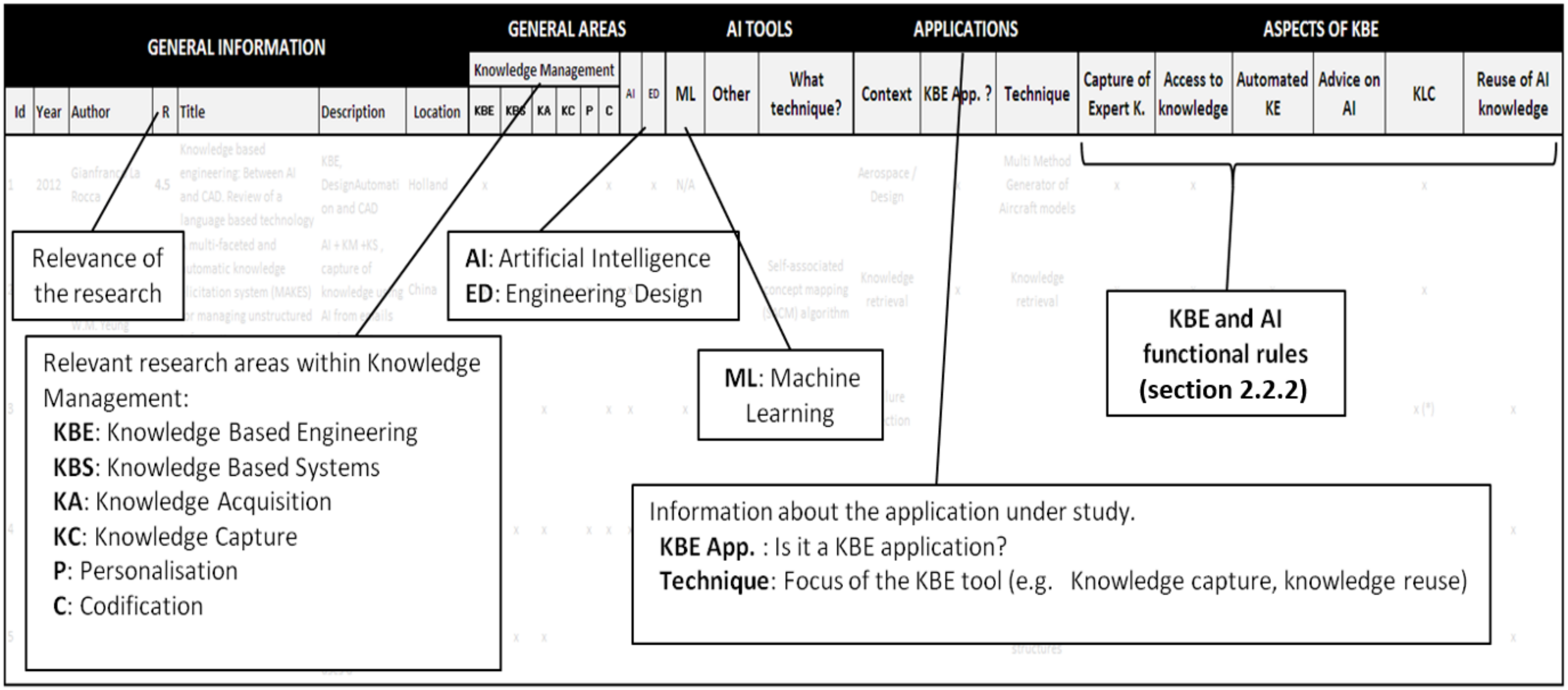
In order to support the assumptions made in this work, relevant papers found in the literature search process have been classified in a comparison matrix including information about the following aspects:

- **General information**: Generic characteristics describing a paper such as title, author and location.
- **General area:** Aspect encompassing information related to main areas covered in the research scope of this study.
- **Artificial intelligence tool exploited:** AI method used in the research under analysis.

- **Type of application implemented:** Type of tool applied in the research under study (special interest on KBE applications).
- **Aspects of KBE applications**: Set of functional roles aiming the efficient source of knowledge and the effective integration of a knowledge sourcing approach within a KBE framework.

A comparison matrix (Figure 7 and APPENDIX E) was created with the purpose of enhancing the data analysis task by allowing the author to easily create graphs that facilitate the identification of the research background topics and key research trends.

Figure 7. Comparison matrix



2.4.2. Key Trends

To support and confirm the interest by researchers on the background of this study, four research questions were formulated:

A. **Is the context of this research appropriate to the proposed approach?**
B. **Is the use of tools integrating KBE and AI methods of interest to the research community?**
C. **What are the expected benefits to be delivered by the proposed extended KBE development process? How are these benefits compared to the benefits brought by existent KBE applications?**
D. **Is the collaboration between experts and AI implementations of interest to the research community?**

A set of tables and charts –obtained as a result of the classification of papers found in the literature review– have been analysed aiming at answering the research questions defined above. The evaluation of these visual representations is described in the following points.

A. Industrial context of KBE implementations

Historically the use of KBE implementations has been limited to large companies such as aerospace and automotive [95]. Indeed, the applicability of this work is constrained to aerospace composite wing structures. Therefore, to validate the research gap it was important to identify studies where KBE applications are applied in the context of the aeronautic industry. In this direction, Figure 8 and Table 4 shows a classification of the KBE papers used in this research regarding their industrial context.

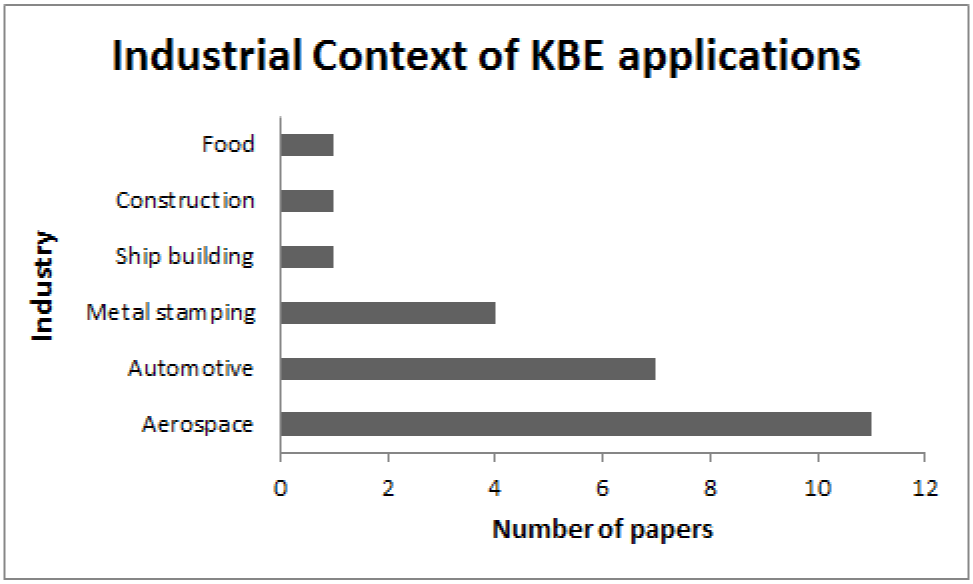


Figure 8. KBE application papers and their industrial context.

Table 4. KBE application papers and their industrial context

Industrial Context	References
Food Industry	[99]
Construction	[70]
Ship building	[100]
Metal stamping	[6],[98],[101],[102]
Automotive	[28],[81],[96],[103]–[105]
Aerospace	[9],[82],[93],[104],[106]–[112]

After analysing papers containing KBE implementations it became apparent that the use of KBE tools is mostly employed by large companies, the aerospace organisations being the ones with more interest in this type of technology. Consequently, the selection of the research scope –constrained to the aerospace industry– is supported. Due to KBE is a relatively new research area where methodological and technological considerations are constantly evolving it is mainly implemented by large companies [6].

B. Historical use of AI in KBE implementations

A few years after Knowledge Management was created, AI started to be used to tackle knowledge intensive problems efficiently [65]. However, the use of AI tools into KBE implementations has not been so evident. In fact, when searching for research papers combining KBE and AI not many are found. For instance, when looking in “ScienceDirect” search engine for KBE and AI keywords 92 articles were displayed on the screen, only 8 papers being relevant to this specific search. Figure 9 and Table 5 show existent papers where AI is used in a KBE application since the appearance of KBE systems.

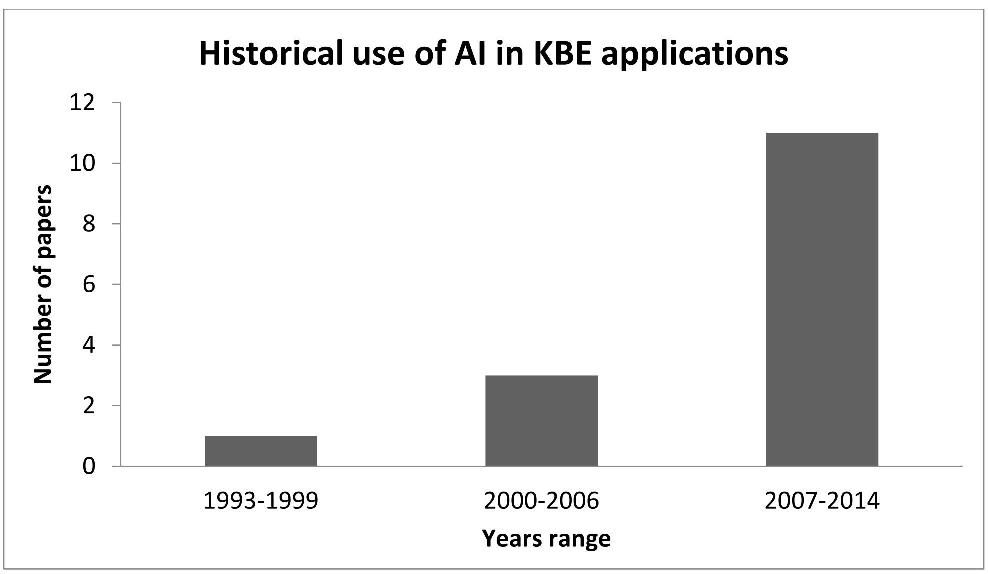


Figure 9. Historical use of AI in KBE applications.

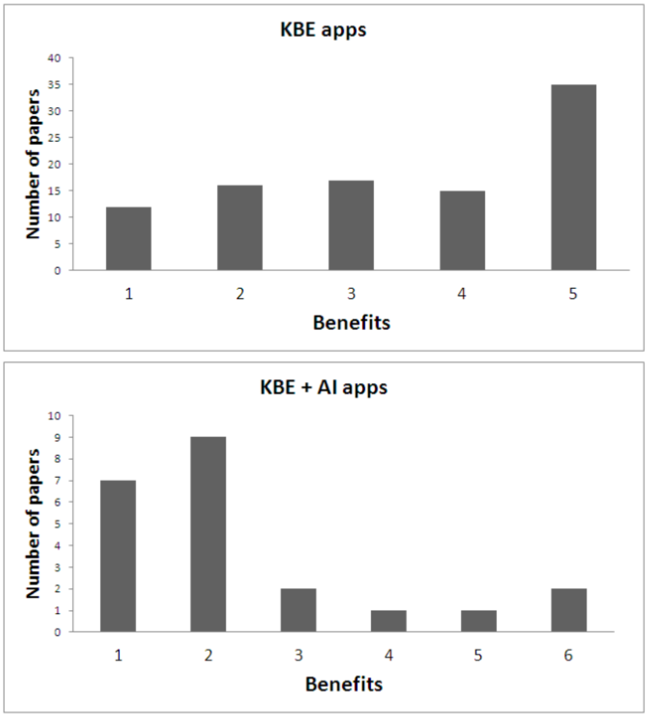
Table 5. KBE application papers and their industrial context

Years range	References
2000-2008	[70],[97],[104],[113]
2009-2016	[46],[66],[67],[89],[94],[107],[114]

Observing the figure above, it became apparent to the author that the use of AI within KBE implementations is gaining interest, thus it supports the development of a framework integrating AI methods within a KBE system. However, further analysis is required to know if those KBE applications using AI techniques aim to deliver a more efficient source of engineering knowledge. This additional research needed is provided in the following point by analysing the benefits of the KBE applications using AI.

C. Benefits of KBE implementations

Airbus's journey into "learning by doing" and "return of experience" is intended to bring some benefits. The benefits expected to be obtained by carrying out this study are described in section 1.6. To show which of the benefits planned to be achieved by this work are already accounted by current KBE applications, Figure 10 has been created. The charts below show the number of research articles providing a particular benefit. The top chart in figure below was performed taking into account all KBE papers found in the literature. In contrast, the chart at the bottom of the illustration was made using only those papers combining KBE and AI methods.



	Benefits
1	Reduce costs of the knowledge capture process
2	Solution to knowledge intensive problems
3	Fosters Knowledge Sourcing
4	Design optimisation
5	Solution to multi disciplinary problems
6	Automate repetitive tasks

Figure 10. Benefits of KBE applications.

When considering articles containing any type of KBE application, there was a higher number of papers focus on the design optimisation and automation of repetitive tasks. In contrast, taking into account only those KBE implementations that use AI methods, two benefits fostering the knowledge sourcing process (“Reducing costs correspondent to the knowledge capture process” and “Delivering a solution to knowledge intensive problems”) were the most repeated. By analysing the charts and tables corresponding to this point, two major remarks were identified:

- Based on Figure 10 it can be deducted that there is interest in obtaining the benefits provided by the knowledge sourcing approach presented in this work.
- Table 6 and
- Table 7 show the lack of researches delivering a solution which integrates all the benefits provided by the research reported here.

Table 6. Benefits of KBE applications

Benefit	Associated references
Reduce costs of the knowledge capture process	[66]–[70],[89]–[91],[98],[115]–[118]
Solution to knowledge intensive problems	[9],[46],[66]–[70],[89]–[91],[97],[114],[117]–[120]
Foster knowledge sourcing	[28],[68],[69],[82],[94],[100],[102],[104],[111],[114],[115],[117],[118],[120]–[128]
Design optimisation	[9],[46],[89],[93],[98],[103]–[105],[107],[109],[112]–[114],[117],[119],[122]–[125],[128]–[139]
Solution to multi-disciplinary problems	[5],[9],[91],[94],[96],[113],[117],[120],[122],[123],[126],[131]–[133],[136],[140]
Automate repetitive tasks	[5],[28],[68],[69],[82],[93],[96],[98],[102],[103],[111],[112],[115],[119]–[144]

Table 7. KBE applications using AI tools: benefits classification.

Benefit	Associated references
Reduce costs of the knowledge capture process	[66]–[70],[89],[91],[98],[116]
Solution to knowledge intensive problems	[66]–[70],[89],[91],[97],[116]
Foster knowledge sourcing	[68],[69],[104]
Design optimisation	[89],[98],[103],[104],[107],[113]
Solution to multi-disciplinary problems	[91],[113]
Automate repetitive tasks	[68],[69],[98],[103]

D. Collaboration between experts and AI implementations

Knowledge capture is acknowledged by researchers as the main bottleneck in the development of KBE systems. Hence, the use of AI techniques to obtain more efficient elicitation tools has attracted more attention in the last decades [68],[145]. However, to effectively realise the knowledge sourcing using AI applications it is advised the intervention of an expert or a group of experts [92].

From this perspective, one of the key features of this work is the use of intelligent machine learning algorithms that allow the creation of new knowledge from non-processed data. In the machine learning area there is a considerable amount of methods which can be used to automate the knowledge elicitation process. In this context, methods capable of creating explicit models or rules from raw data have gained engineers attention due to their higher reliability compared to other methods known as “black box” applications. “Black box methods” refers to the techniques where rules that drive the process are hard coded and not accessible by the user.

Table 8 shows the classification of research papers integrating AI and KBE regarding their context and the interpretative information provided by the AI algorithm.

Table 8. Characteristics of automated reasoning methods used in KBE applications.

Id	Rules can be managed by the user	“Black box” method	Use of explicit rules easy to understand	Context
[66]		x		Electronics
[67]	x		x	Customer Satisfaction
[68]	x		x	Services Moderator
[69]	x		x	Maintenance
[70]			x	Construction Design
[91]	x		x	Failure Detection
[89]			x	Structural Design
[97]		x		Mould Design
[98]	x		x	Design of dies

[90]		x		Chemical structures

2.4.3. Research Gap Analysis

The identification of a probable research gap on knowledge sourcing aspect of KBE led the author to the definition of a research hypothesis assuming that an extended KBE system, delivering the knowledge sourcing process requested by the industry, can be achieved by integrating methodological support (EKM methods) with AI tools. Moreover, the intervention of an expert or group of experts in the problem domain is also an element of the integrated solution enabling the knowledge sourcing framework to deliver more accurate and reliable predictions.

To support the research gap presumed at the beginning of this research, a set of functional roles associated to KBE and knowledge sourcing tools were defined. The aim behind the definition of these functional roles was to facilitate the identification of significant research trends and opportunities (apart from the ones identified in section 2.4.2). Moreover, the reliability of the trends and opportunities found was increased by asking six experts in domains related to this research (

Table 9) to prioritise each of the functional roles (Table 10).

Table 9. Experts' description.

Expert id	Area of expertise	Academia/ Industry	Experience in KBE / AI	Years of experience
Exp 1	Through life services, concurrent engineering, rule based approaches and numeric data analysis	Academia	KBE + AI	10
Exp 2	Sustainable design and manufacturing	Academia	KBE	20
Exp 3	Advanced knowledge capture systems	Academia	KBE + AI	7
Exp 4	KBE systems in design for manufacturing and automated CAD tools	Industry	KBE	17
Exp5	KBE systems in design for manufacturing and manufacturing systems	Industry	KBE	15

Exp 6	CAD applications and aerospace manufacturing optimisation	Industry	KBE	11

Table 10. Experts' assessment

Functional Role	**Expert id**					
	Expert 1	Expert 2	Expert 3	Expert 4	Expert 5	Expert 6
Capture of expert knowledge	Low/Medium	Low	High	Medium	Low	Low
Access to knowledge by KBE tool	Medium	Low	Low	High	Medium	High
Automated knowledge extraction from data	High	Medium/High	High	Medium	High	High
AI extracted knowledge ready for reuse	High	Medium/High	High	High	High	High
Advice on AI tool suitability	High	High	Medium	High	High	High
Knowledge lifecycle management	High	Medium	Low	Medium	Medium	Low

The level of research priority was defined by the combination of expert assessment together with the analysis of papers reporting KBE implementations. The expert assessment activity was realised in a focused session where experts were asked to score the priority of the tools using High/Medium/Low scores (Table 10). The "KBE expert assessment" column in Table 11 provides the averaged values of the outputs from the interview process described in Table 10. In parallel, the values of the column entitled as "Literature" in Table 11 represent the number of articles using each of the functional roles.

The analysis of the results provided by Table 11 led to the definition of three main groups of researches. The first group is closely related to the knowledge management area and it includes three functional roles entitled as "Capture of expert knowledge", "Access to knowledge by KBE tool" and "Knowledge lifecycle management". Special attention was paid by the author to some of the papers included in this group such as [5],[9],[28] where the use of the three functional roles mentioned is clearly visible. More precisely, in [9] it is well described how knowledge captured from experts is systematically stored separately from the application. This knowledge is placed in an external knowledge repository within computable knowledge models, thus enabling the knowledge to be easily retrieved by any KBE application.

The second group identified is related to the use of AI within knowledge-based tools. It includes two functional roles named as “Automated knowledge extraction from data” and “AI extracted knowledge ready for reuse”. The integration of AI within KBE applications is a key feature of this work and, –as observed in the previous table– although experts are highly interested in this topic there are only a few papers present in the literature concerning this type of implementations. [66],[70],[90],[91] are the most significant papers found in which the two functional roles are identified. From this perspective, [67]–[69],[89],[91] apart from using a machine learning approach –to automatically extract knowledge from data and store it in a database–, include an extra functionality that allows the management of the rules extracted (e.g. Figure 11) .This extra functionality is of high interest to this research as it shows an example where machine and expert knowledge are combined.

Table 11. Research priorities analysis.

Functional roles	P	KBE experts assessment	N of Papers	Literature
Capture of Expert Knowledge:	Low	Not a particularly well understood tool in the context of KBE. However, often its use results on a costly process so automated approach would beneficial.	62	[5],[9],[28],[46],[66]–[70],[81],[82],[89],[90],[93]–[98],[100],[102]–[105],[107],[109]–[115],[117]–[144],[146],[147]
Access to knowledge by KBE tool:	Med	Well understood tool at mainstream software development as interoperability. However, its realisation in KBE is limited.	57	[5],[9],[28],[46],[66],[68],[69],[81],[82],[89]–[91],[93]–[98],[100],[102]–[105],[107],[109]–[114],[117]–[144],[146],[147]
Automated knowledge extraction from data:	High	A common tool in environments where large amount of data exist but no formal knowledge models are available, (i.e. maintenance, in service support). In engineering where	7	[46],[66]–[70],[120]

		knowledge models exist there is a high interest in transforming data into usable (computable) knowledge models.		
Advice on AI tool suitability:	High	The particular choice of AI algorithms to cope with engineering problems is based on the experience of experts. Tools providing advice on algorithm selection are of high interest and low availability. A sensible approach may be exposing to users a portfolio of problems and the algorithms successfully used to provide a solution.	0	
Knowledge lifecycle management:	Med	Changes to the knowledge used by KBE applications are common. The ability to modify, select and deselect engineering rules on an offline process has interest. This capability is progressively being introduced into current applications.	60	[5],[9],[28],[46],[66],[68]–[70],[81],[82],[89]–[91],[93]–[98],[100],[102]–[105],[107],[109]–[115],[117]–[125],[127]–[137],[139]–[144],[146],[147]
AI extracted knowledge ready for reuse:	High	Assuming that usable engineering knowledge can be extracted from data, still there is a challenge on how to utilise it. Research is needed to understand how the outputs from AI-based knowledge extraction tools can be leveraged in KBE tools.	9	[67]–[70],[89]–[91],[97],[98]

Finally, the last group is related to the functional role termed as “Advice on AI tool suitability”. Despite the high interest of the experts in this kind of applications, there is no research advising on AI algorithm selection which considers the knowledge management aspect. However, there are many papers presenting AI algorithm portfolios –constrained to a specific problem– that guide the method selection such as [85],[86]. Moreover, some research studies present applicable examples of a wide variety of AI algorithms, thus facilitating the task of selecting a particular technique [87],[88].

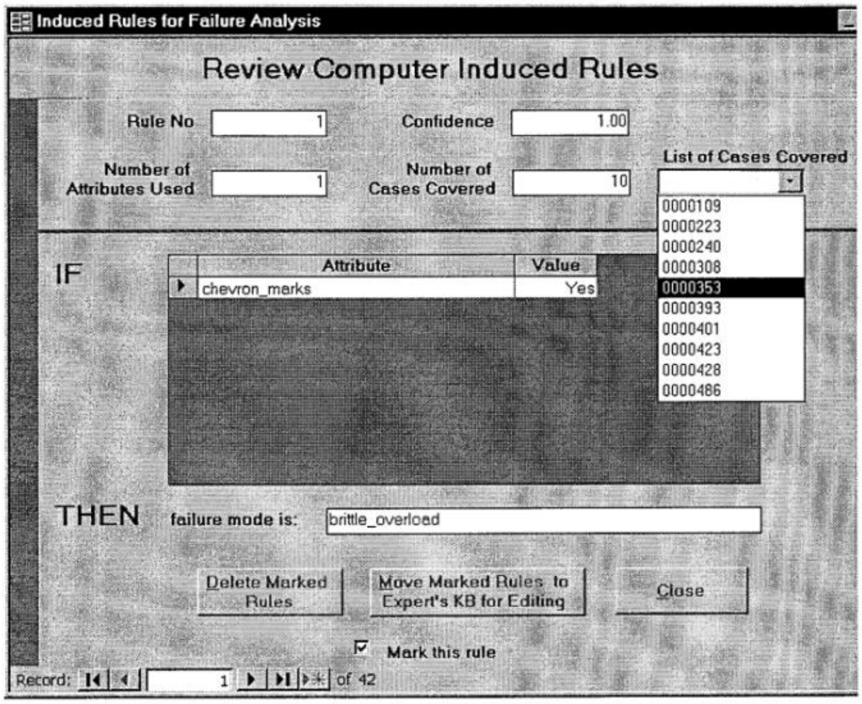


Figure 11.Rules' management user interface [91]

After analysing Table 11 in more detail, it became apparent to the author the interest on tools enhancing the knowledge management and knowledge sourcing aspects of KBE. This is supported by the fact that:

- KBE applications presenting a methodology to manage knowledge life cycle (papers classified within "access to knowledge by KBE tool" and "knowledge lifecycle management" functional roles) don't use AI tools to perform the knowledge sourcing process.
- KBE papers focused on delivering a more efficient knowledge sourcing process using AI tools (classified within the functional roles entitled as "automated knowledge extraction from data" and "advice on AI tool suitability" and "AI extracted knowledge ready for reuse") don't present a generic methodology in charge of managing the engineering knowledge.

In summary, the review of KBE papers classified using the defined functional roles confirmed the need for a more efficient knowledge sourcing approach, and the creation of a better link between KBE and knowledge sourcing.

2.4.4. Discussion

The analysis of the existing literature enabled the definition of the research background and the identification of meaningful research tendencies. Besides, the analysis of the research trends found allowed the author to validate the research gap.

Initially, a research gap on knowledge sourcing was identified as a response to the current limitations of KBE systems (listed in section 1.1). The research gap found represented an opportunity to extend the capabilities of current KBE systems often used only as inference implementations. To validate the research gap, research trends obtained from the literature analysis process were further analysed. The research trends identified, which are relevant to this study, are listed in Table 12.

Table 12. Research trends.

Research Trend	Related Section and Visualisations
Increasing interest of KBE within the Aerospace Industry	2.4.2. A & (Figure 9)
Increasing interest of using AI within KBE systems	2.4.2.B & (Figure 10)
Interest of integrating AI and KBE to enhance the source of engineering knowledge	2.4.2.C & D (Table 7 & Table 8)
High interest on tools enhancing knowledge life cycle and knowledge sourcing aspects of KBE.	2.4.3 & (Table 11)

The evaluation of the research trends led to the conclusion that there is high interest on KBE and more precisely in integrating AI tools and KBE methods within the context of the aerospace industry. This trend together with the inexistence of researches presenting a generic methodology –using advanced computational tools to source engineering knowledge– confirmed the existence of a gap on knowledge sourcing and represented an opportunity to develop a framework aiming the validation of the research hypothesis.

2.5. Research scope

The detailed literature review presented in this thesis focused on identifying and analysing published work contributing to:

- **The sourcing of engineering knowledge through the use of automated reasoning tools.** In particular, AI is a field of research delivering promising tools to acquire knowledge efficiently. In fact, the use of AI tools seems to enhance the extraction of knowledge finding correlations within large amount of data [10],[65]. Furthermore, it reduces costs related to the process of knowledge extraction from experts. In this thesis, only existing AI implementations will be applied (e.g. Weka and Scikit-learn) rather than developing new algorithms.
- **The methodological support to manage engineering knowledge and making it available to KBE**. The route proposed to integrate the knowledge sourced is the combined use of existing Engineering Knowledge Management (EKM) tools.

Literature on these topics is both associated to Artificial Intelligence and the Engineering Knowledge Management research communities. In addition to these specific topics, the application of this research in the engineering design domain will be surveyed in detail in order to build the foundations of the research cases presented. In summary, the fields of study related to the contribution to knowledge are: engineering design, artificial intelligence and Engineering Knowledge Management (EKM).

Moreover, aiming at providing the reader with a better understanding of this research an explanation of the KBE, EKM and knowledge sourcing concepts and their relationships is provided in APPENDIX A.

2.6. Concluding remarks

This chapter presented the methodology followed in order to search, classify and analyse literature relevant to this work. Special attention has been paid to researches aiming at enhancing the engineering design process through knowledge-based engineering capabilities. Two main conclusions have been extracted from the work described in this chapter:

- KBE technology delivers a framework for automating design repetitive tasks but also supports the knowledge retention and reuse. One of the key features of KBE is the ability to manage engineering knowledge out of the application. This minimises the risk for knowledge loss and fostering its reuse across different engineering problems. Although KBE is mainly use for automation it has been also successfully applied to enhance knowledge sourcing. Despite the benefits provided by KBE, it still fails to provide an efficient approach for expert knowledge capture.
- AI methods are well suited for knowledge capture which is an essential component of knowledge sourcing. However, they are often used as “black box” applications which don’t provide explicit knowledge. Moreover, when using AI techniques delivering interpretative information they are independently exploited. Although AI technology is capable of delivering a more efficient knowledge sourcing process, there is still a lack of integration of AI algorithms within a methodology supporting the management of the knowledge generated.

These two concluding remarks highlight a research gap on knowledge sourcing in the engineering design context. Based on this, an extended KBE development process also defined in this research as a knowledge sourcing methodology – encompassing the adoption of a generic methodology for managing the complete knowledge life cycle and the use of AI techniques for knowledge capture– was proposed. Prior the development of the knowledge sourcing capability a set of research questions had to be answered. These research questions sought to identify if there is a real interest by the research community and the industry in the knowledge sourcing methodology proposed.

To respond to these questions a detailed analysis of researches aiming at improving knowledge sourcing was undertaken. This review highlights the increasing interest in using KBE and AI in the context of aerospace industry for knowledge sourcing. While the literature survey supports the development of the knowledge sourcing proposed, the next chapter seeks to support the realisation of this work from an industrial perspective. This is done by identifying and further analysing existing challenges in actual knowledge sourcing capabilities implemented in the context of the engineering design in aerospace.

3. Knowledge Sourcing: Industrial applications

The purpose of this chapter is to present previous industrial research studies aiming at delivering an effective knowledge sourcing approach. Research studies including the description of industrial application belonging to various research contexts have been identified (APPENDIX E). Relevant studies to this PhD thesis (with special attention to applications developed in the aerospace industry) have been described being two key researches selected to be studied in more details. These researches selected to be further analysed that address the source of engineering knowledge are presented in this chapter, highlighting their benefits and limitations.

The structure of this chapter starts highlighting the characteristics of existing KBE systems which consider knowledge sourcing as a priority with special attention on existing researches in the aerospace industry in section 3.1. This is followed in section 3.2 by the description of the criterion followed to select the knowledge sourcing capabilities to be further studied. In section 3.3 the methodology followed regarding to analyse the capabilities selected for in depth review is described whereas their review is detailed in section 3.4. Finally, conclusions derived from the analysis of the industrial examples are presented in section 3.5.

3.1. Introduction

The analysis of the literature led to the identification of research trends denoting an increasing interest in KBE tools focused on the management aspect. Indeed, KBE is currently considered by the research community as a potential solution to enable the capturing, retaining and reusing of engineering knowledge.

From an industrial perspective, this research has paid attention to KBE researches focus in delivering an effective knowledge sourcing approach. Two different strategies followed by existing KBE applications aiming at improving knowledge sourcing have been identified:

1. **Encoding tacit knowledge into knowledge-based system**. The objective of this type of approach is to retain expert knowledge within the organisation minimising the impact of knowledge loss caused by experts leaving or retiring. To do that, expert knowledge is captured and coded into an inference tool. In this context, [148] presents a framework for the management of Knowledge-Based Engineering applications as software services enabling the automation of repetitive tasks in the design for manufacturing context. It systematically manages the engineering knowledge life cycle management by using a content management system which facilitates the accessibility, traceability, review and reuse of the knowledge stored. The knowledge used in this capability was captured by interviewing experts in the domain. In this direction [135] also focuses on the automatization of repetitive and non-creative engineering tasks by proposing a KBE tool supporting engineering design application development. This KBE application was successfully applied for the design and manufacturing of aircraft wiring harnesses reducing the recurring time of the assignment process by 80%. [9] presents an effective and faster KBE cost estimation tool encoding expert knowledge within a knowledge repository as independent pieces (knowledge elements). The knowledge elements selected by the user are later on used for the automated estimation of the cost of composite wing structures. [141] also presents a KBE tool for weight and cost estimation of composite airplane structures including structural analysis where knowledge is externally stored within a knowledge based separately from its inference tool to facilitate the reuse of the knowledge across different engineering problems. [96] and [109] developed KBE applications focussed on design automation using decoded knowledge which is stored in an external knowledge base. [96] automates the design activity using flexible CAD models where the parameters used to iteratively create the objects that compose a design are encoded in a knowledge repository. In parallel, the KBE application developed by [109], automating the design task, is built on top of two well-established methodologies for the effective

knowledge life cycle management: CommonKADS and MOKA which are further described in section 4.2.

2. **Decoding knowledge from company data assets**. The loss of relevant knowledge is often caused by the reduction of experts' availability together and the existence of tools or large data lakes which are not being effectively analysed. In this direction [149] developed a methodology to extract knowledge from company data assets generated from an aerospace application that estimates the wing structures. This approach proposes the use of Design of Experiments (DoE) to generate the required dataset and Response Surface Method (RSM) to create the explicit model of the problem. [89] developed a KBE system where a ML algorithm is used to create the rules modelling a finite element analysis problem for structural design. In the same direction [69] uses a machine learning algorithm for the automated decoding of expert knowledge while implementing a methodology for the systematic capture and reuse of engineering knowledge. However, this application does not support the systematic maintenance and update of the knowledge repository. Therefore, it does not supports the complete management of the knowledge life cycle. [91] apart from presenting a framework for the automated elicitation of engineering knowledge for the automated detection of a failure in a mechanism it enables the review, modification and validation of the knowledge decoded by the ML algorithm.

Recent KBE implementations developed either to encode or decode expert knowledge produced promising results regarding knowledge retention and retrieval. However, some key barriers still remain constraining the use of KBE to source engineering knowledge [9]. In this context, the detailed analysis of two industrial cases carefully selected (as explained in section 3.4) enhances the understanding of the existing constrains limiting the use of KBE for sourcing engineering knowledge.

3.2. Selection of knowledge sourcing capabilities

The selection of the industrial researches to be further analysed in this chapter was based on two characteristics: a) the industrial context where the researches were applied and; b) the availability of the experts who carried out the studies.

On one hand, the detailed evaluation of two case studies applied in the same context of this PhD thesis was realised. This pursued the identification of current practices targeting the identified KBE challenges in the context of the aerospace industry. In doing so, the framework proposed in this research is developed considering the relevant KBE aspects associated to the aerospace context.

On the other hand, the researches under evaluation were carried out in the same company where the author of this work is based. This enabled a better access to the experts who developed the case studies, thus a more detailed analysis of the researches was permitted.

3.3. Method followed to analyse the selected KBE capabilities

A method was defined in this research to effectively select and analyse KBE capabilities belonging to the domain problems identified in section 3.1. The steps followed in this method are listed and described as follows (Figure 12):

1. **Identification of existing studies in the defined problem domains.** Researches aiming at improving the knowledge sourcing process including the description of a use case applied in a real industrial scenario were identified and reviewed.
2. **Search and analysis of publish material related to the selected capabilities**. In this direction a journal paper describing a case where expert knowledge is encode was identified [148] whereas in the case of decoding knowledge from company assets a master thesis describing the capability was found [149].

3. **Search and analysis of internal material.** Documents such as "statement of work" or documentation created for the project deliverables were studied in details. The KBE selected for both domain problems were part of an industrial project therefore it was possible to gather useful information mainly from the documentation created as part of the requested project deliverables. More specifically, this documentation included detailed technical information regarding the architecture and functionality of the developed capabilities.
4. **Analysis of the existing capability.** This was performed from the user point of view in order to identify the benefits and drawbacks of using such a framework.
5. **Informal interviews with the domain experts.** After gathering some knowledge and understanding of the architecture and functionality of the capability it was useful to discuss about the identified advantages and disadvantages with the experts in charge of developing such a framework in order to get a better understanding of the real limitations of the tool. After each informal interview it was useful to go back to the analysis of the capability in order to gain a better understanding of what is stopping a wider application of KBE systems in the industry.

Figure 12. Method followed to analyse the selected KBE capabilities

In summary, the structured approach followed to review the selected industrial researches provided a more thorough understanding of the current practices to source engineering knowledge in the aerospace industry. This lead the author to the development of a more complete and adequate knowledge sourcing framework.

3.4. Examples of knowledge sourcing in the aerospace context

A common background and purpose is shared by the two industrial cases described in this section. More precisely, both cases were developed in the context of an aerospace organisation aiming at the improvement of an existing cost modelling capability by realising a more efficient source of expert knowledge.

The cost modelling tool was built to provide AGI with improved manufacturing systems supporting the conceptual design evaluation of wing designs: in short, manufacturing processes and products were optimised considering a specific driving parameter such as cost, time, etc. Initially, the number of parameters and operations considered in the cost model was small enough to allow the tacit knowledge –included in the tool– to be easily understood. However, the increasing growth of the cost model derived in more than 300 driving parameters and more than 1000 intermediate calculations to be accounted, thus resulting in too many combinations and possible iterations of knowledge elements. As a consequence of the cost model complexity, tacit knowledge became impossible to understand by those people not involved in the creation of the tool. Therefore, the costing capability became a "black box" application making the knowledge transfer and reuse difficult to achieve.

Challenges linked to the research problem –caused by the cost model evolution– motivated engineering teams within AGI to propose solutions to decompose the knowledge encompassed in the costing application (Figure 13); changing implicit knowledge into explicit allowing the knowledge to be retained and transferred. In this direction, two different approaches addressing the knowledge sourcing process were undertaken as described below.

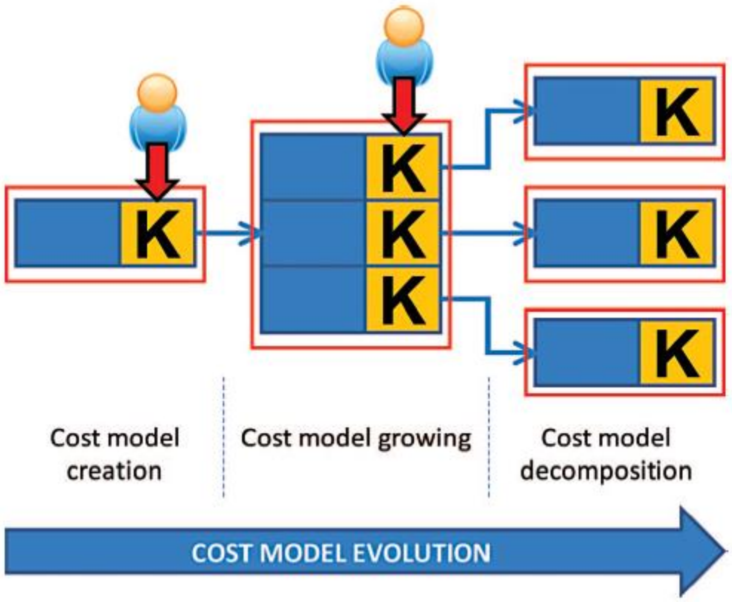


Figure 13. Cost model evolution

3.4.1. Industrial case 1: Encoding expert knowledge

The first case aiming an improved cost modelling tool was focused on transforming tacit into explicit knowledge [148]. The proposal tackled the inefficient sourcing of knowledge by adopting three main EKM foundations: knowledge life cycle management, knowledge visibility and knowledge exploitation. Moreover, to adopt the listed knowledge management foundations within a successful framework, three major elements are considered in the solution architecture: *Knowledge Repository, Knowledge Reuse Engine and Knowledge Reuse Spreadsheet.* The way this use case addressed these EKM principles and their relation with the architecture solution elements is described as follows [148]:

- **Knowledge visibility.** The usability of the knowledge is increased by:
    - Storing knowledge independently from its application within an external knowledge Repository.
    - Decomposing cost model into knowledge elements or building blocks, enabling the creation of unique cost models having a specific purpose.

The decomposition of the cost model is a manual task realised by experts in the domain in collaboration with knowledge managers.

- **Knowledge life cycle management.** It refers to the update of knowledge elements and storage of tailored cost models, to respectively: (i) facilitate the maintenance of the knowledge captured; (ii) allow the review of previous studies –listed in all the cost models performed. These benefits are achieved implementing:

    - Knowledge Reuse Engine using VBA scripts and Web Services to save specific cost model versions within the Knowledge Repository.
    - Content management system as a Knowledge Repository to update the knowledge elements used to build the cost model.

- **Knowledge exploitation**. It encompasses the automated extraction of reliable and updated knowledge –from an external database– and the generation of user-defined cost models. The adequate exploitation of the cost model was achieved using the following architecture elements: (i) *Knowledge Reuse Engine* in charge of driving the queries –defined by the user– to retrieve the cost model building blocks from the knowledge repository; (ii) *Knowledge Reuse Spreadsheet* in control of leveraging the selected building blocks and creating the corresponding cost model, including user defined values and options selected. Additionally, the solution also allows the execution of existing cost models.

Overall, the resulting framework proved –by the realisation of a proof of concept– to be a capability able to realise the cost modelling exercise more efficiently. More precisely, the benefits provided by the developed solution are:

- Less time required to compose a cost model.
- Management of the model complexity.
- Delivery of a more user friendly application.
- Fosters reuse and transfer of relevant knowledge.

However, despite the benefits provided by the solution, the semiautomatic nature of the capability presenting non-value added and expert time consuming activities has limited its use in the industry. Experts were only required to take part at the beginning of the methodology in a time consuming task to decompose the knowledge model into building blocks. This means that any mistake made by an expert would remain in all the following steps of the procedure.

In addition, the limited expert´s availability provoked the developed framework to be just partially completed; only taking into account a small fraction of all the existing processes. As a consequence, a research work emerged in AGI addressing the expert availability problem; decoding knowledge from the existing cost model using Design of Experiments (DoE) and data mining methods.

3.4.2. Industrial case 2: Decoding knowledge from company data assets

In this industrial case, the framework proposed is focused on the automated extraction (decoding) of knowledge from the costing tool, considering the cost model as a large database [149]. To do that, DoE techniques were employed to reduce the number of parameters whereas data mining methods were used to generate mathematical functions of the procedures contained in the cost model.

The architecture corresponding to this study contains three elements: DoE engine, data mining application and knowledge reuse template. These elements were integrated within a semiautomatic methodology summarised in three main steps:

- **Parameter extraction and reduction**. Prior the realisation of the DoE study, an analysis of the data and processes contained by the cost model was performed. The outcome of this evaluation process was the delivery of a list of parameters which experts believe are relevant for a specific process. The DoE procedure consisted of automated process in charge of reducing the number of parameters existing in the list provided by the experts. To do this, the impact caused by the parameters on their associated processes (when modifying the parameters values) was calculated, provoking the parameters with low impact values. As a result a reduced list of parameters driving the cost modelling processes was generated.

- **Rule generation.** Before using a specific algorithm to create the rules describing the cost modelling processes, a set of data mining methods were evaluated with Response Surface Method (RSM) being the selected technique. RSM was considered as the more appropriate after realising an analysis based on method´s usability, processing time, accuracy and design limits (maximum population allowed in the analysis). The list of parameters provided by the DoE is employed by RSM to create a set of rules describing the costing tool behaviour. More precisely, for each process in the cost model a mathematical function was generated by RSM.

- **Knowledge reuse.** To allow the reutilisation of the rules previously generated, a VBA script integrated within an excel template was created. The purpose of the excel template is to store and enable the automated execution of the mathematical equations extracted by RSM. The knowledge reuse template represents a user friendly platform where users can generate MFG time estimations of new design configurations by modifying the values corresponding to the parameters used by the cost model capability.

The capability developed in this industrial case demonstrated to be an effective tool to simplify the complexity of the cost model and enable the user to trace back the results obtained. However, the methodology followed contains several constraints limiting its use:

- **Lack of knowledge life cycle management**. The systematic capture, retain and reuse of the knowledge is not supported.

- **No consideration of rule complexity in the data mining evaluation process.** An identified limitation to the wide use of this capability was the difficulty found by experts to rely the complex mathematical functions modelling each of the manufacturing processes contained in the costing tool.

- **Non expert intervention in the rule validation process.** The validation of the data mining rules generated without requiring expert involvement permits the existence of not meaningful rules from an engineering point of view, thus causing the lack of reliability on the results provided by the framework.

3.5. Concluding remarks

This chapter has presented previous work addressing the efficient sourcing of engineering knowledge. In this regard, two industrial cases developed under the same context have been analysed and validated employing the same cost model capability. The access to the tools created in each use case, and to the experts involved in their development facilitated the understanding of the limitations representing the key barriers stopping KBE from being a widely used methodology to capture, retain and reuse knowledge. Indeed, the limitations associated with the industrial cases studied support the KBE research challenges identified in [9] as shown in Table 13.

Table 13. Case 1 and case 2 limitations regarding the identified KBE challenges

<table>
<tr><th>KBE challenge</th><th>Limitations associated with Case 1 (Encoding expert knowledge)</th><th>Limitations associated with Case 2 (Decoding knowledge from a company assets)</th></tr>
<tr><td rowspan="2">Need for a more generic and reliable methodology</td><td colspan="2">Lack of an established methodology –such as MOKA (Methodology and Software tools Oriented to Knowledge-based engineering Applications) or KNOMAD (Knowledge Nurture for Optimal Multidisciplinary Analysis and Design)– to systematically source knowledge</td></tr>
<tr><td></td><td>No expert involvement in the validation process. This characteristic is the main cause of the capability’s low reliability.</td></tr>
</table>

A more efficient way of sourcing knowledge (and especially capturing expert knowledge) is required.	Highly time consuming approach: it encompasses the use of an EKM methodology to realise the knowledge retention and reuse aspects of knowledge sourcing more efficiently, and the use of an inference tool to automate repetitive tasks –integrating KBE applications into engineering workflows. However it becomes an unaffordable solution when considering expert availability.	

The table above presents the current limitations of two engineering capabilities realising the knowledge sourcing process in the context of the aerospace industry. It highlights the need for a knowledge sourcing approach delivering an efficient knowledge capture process and the need of a generic methodology capable of managing engineering knowledge in a systematic manner. Therefore it becomes apparent the requirement of a framework delivering a faster and generic knowledge sourcing process.

4. Knowledge Sourcing Framework Structure

The literature survey performed in chapter 2 highlights a research gap in knowledge sourcing whereas chapter 3 remarks the limitations of industrial knowledge sourcing capabilities that support the conclusions gathered from the literature review. Based on this, this research takes the challenge of developing a knowledge sourcing framework.

The outcome of this chapter is the architecture definition of the knowledge sourcing framework. The framework architecture is structured in three main blocks. These main blocks represent the required actions to accomplish the objectives of this research. In this regard, the actions to be taken together with their related research objectives are presented in Table 14.

Table 14. Framework Structure: main blocks.

Elements of the framework architecture	Related research objectives
Search, analysis and exploitation of machine learning methods to enhance knowledge capture.	2 – 3
Adoption of an existing methodology enabling the systematic knowledge life cycle management.	4 – 6
Development of a platform where experts and machine learning algorithms can interact, create, review and validate new knowledge. The platform also allows the access of relevant data, and the advanced analysis of information.	2 – 6

4.1. Search, analysis and exploitation of machine learning methods

The first task realised in the development of the proposed framework was to identify machine learning methods capable of generating an explicit model supporting the sourcing of engineering knowledge. To do that, a study of common ML methods widely used by the research community was carried out using as main resources books [87],[88],[150]–[152] and acknowledged open-source machine learning tools (Weka and scikit-learn).

ML methods are generally classified in two main groups: supervised and unsupervised learning [150]. Algorithms belonging to the supervised learning group contain target values or labelled data. This means that these methods, in order to create a set of rules, use input data samples containing the sought output values in the learning process, thus the explicit model can be easily evaluated (e.g. using Cross-Validation method or a “Test Set”). Moreover, supervised methods assume that some observations (inputs) are the cause of the other observations (outputs). An example of this type of methods is presented when a supervised learning algorithm is used to predict next month weather forecast based on historical data.

The second type of machine learning methods is known as unsupervised learning. These algorithms aim finding hidden structure within unlabelled data in cases where you don’t really know what you want to predict. These methods are often used to find commonalities in data, delivering as a result a set of clusters [151]. A typical example of this is the definition of shirt sizes or clusters (small, medium, large, etc.). The use of this technique helps to select the optimum number of clusters that would make the clients feel more satisfied, saving time and costs.

In this research only supervised methods were used due to the characteristics of the problems faced in the 2 case studies implemented. These properties are: a) the nature of the data capture where observations are caused by other observations and; b) the data available was labelled. Supervised learning encompasses two types of techniques: regression and classification [150]. Regression algorithms aim at the prediction of continuous values. In manufacturing, an example of regression methods is the prediction of the time required to perform a specific activity. In contrast, in classification the target is to identify the category of a particular observation. The diagnosis and detection of cancer is considered as one of the most common examples of classification techniques [153]. The categories are ‘YES’ or ‘NO’ which refer to the two possible symbolic values in the cancer prediction process.

Within the supervised methods, many machine learning methods behave as ‘black box’s algorithms [154]. This means that the user introduces input data and only retrieves the predicted results without any interpretative information. This leads to the impossibility of tracing back the results generated. However, there are some methods which, on top of the prediction values, also provide an explicit model; thus the user is able to better understand how the results were obtained. In this research, the delivery of interpretative information provided by ML algorithms (in the learning phase) is a key feature necessary to achieve the established objectives for the reliability of the methodology.

The explicit model or rules provided by some ML methods are usually found in the format of “IF THEN” equations or tree representation. Table 15 shows a classification of acknowledged and widely used supervised learning methods [120]–[122] which are embedded within web mining applications such as [155] and [156]. The classification of the techniques is based on the type of method that they can deal with and the level of understanding of the interpretative information provided. Most of the information of the table below was gathered from the researches specified in “References” column. The table was completed with knowledge captured from the experts involved in the development of the use cases implemented in this research (mainly knowledge related to the level of rule’s understanding).

From the analysis of the table shown below it is observed that explicit rules can be only extracted from a few techniques. The existence of some methods providing the user with relevant information –about the problem-logic or the equations driving the target variable– represent an opportunity to carry out the knowledge elicitation process more efficiently (what is aligned with some of the objectives established in this study).

Finally, the methods presented in Table 15 –containing explicit rules– have been integrated within the platform developed (see section 4.3) to allow the improved knowledge capture. Moreover, a description of these selected methods is described below.

Table 15. Classification of common supervised machine learning methods.

Method	**Solve Classification Problems**	**Solve Regression Problems**	**Explicit Rules can be directly extracted**	**Level of rule's understanding**	**References**
Decision trees (REPTree)	Yes	Yes	Yes	Low / Medium	[157],[158], [159]
Random Forest	Yes	Yes	No	-	[160]
Naive Bayes	Yes	No	No	-	[161],[162]
Logistic Regression	Yes	No	No	-	[163],[164]
Support Vector Machines	Yes	Yes	No	-	[165],[166]
Linear Regression	Yes	Yes	Yes	It depends on the problem's complexity	[167]
Gaussian Processes	No	Yes	No	-	[168]
Rule methods (M5R)	No	Yes	Yes	Medium / High	[169]–[171]
Neural Networks	Yes	Yes	No	-	[166],[172]
Fuzzy Rule-Based Systems	Yes	Yes	Yes	Medium/High	[173],[174]

A. Linear regression

This statistical method is one of the oldest and it finds the best-fitting straight line by modelling the relationship between a dependant variable y and one or more explanatory variables x with their specific slope coefficient [175],[176].

There are different types of linear regression depending on their parameters:

- **Simple linear regression**: It only uses one independent variable counting with two parameters expressed as follows:

  $$Y = Bo + BiXi + \varepsilon\mathrm{i}$$

  Where Ɛi is the error associated to the Xi parameter and Bo and B1 are the coefficients of the equation.

- **Multiple linear regression:** It considers the response variable as a linear function of the parameters of the model where there are no independent variables in the model. The equation of a multiple linear regression model is represented as follows:

  $$Y = Bo + B1X1 + \cdots + BnXn + \varepsilon\mathrm{i}$$

In short, simple linear regression investigates the linear relationship between one dependant variable and one explanatory or independent variable whereas multiple linear regression is focused on the relationship between one dependant variable and two or more independent variables. Due to the greater number of independent variables considered by multiple linear regression it also difficulties the achievement of a highly accurate model due to certain effects such as collinearity and detection of outliers [176].

B. Fuzzy Rule-Based Systems

Fuzzy Rule-Based Systems (FRBSs) are an extension of production systems which are classified as classical rule-based systems. FRBSs are based on fuzzy set theory introduced by [177]. Fuzzy set theory enables engineers the adequate representation of their knowledge (concepts, variables and rules) in a more natural language. To do that, a model in the format of IF-THEN rules in fuzzy set theory. This is in the form of “IF X THEN Y” where X and Y are fuzzy sets. The use of fuzzy sets shifting from only using members of the model with only two values (0 or 1) to use as many values from 0 to 1 as desired. In fact, the number of values of a specific fuzzy set is determined by the membership function.

For instance, having two parts of a rule X and Y known as antecedent and consequent respectively and assuming we are trying to solve a problem of calculating the distance to cover running we need features such as speed of the runner and time while he/she was running with linguistic values defined as follows:

- Distance: short, medium, long.
- Speed: low, medium, high.
- Time: short, medium, long.

Based on the specific conditions of the problem it is possible to define an IF-THEN fuzzy rule as: IF the speed in the runner speed is high but in a short period of time THEN the distance covered will be medium.

The use of FRBSs has been used for solving problems regarding uncertainty, uncertainty, non-linearity and imprecision in wide range of areas such as pattern recognition [178], data mining [179] and control engineering [180] among others. The advantage of using fuzzy rules is their facility to understand them by experts due to the use of the fuzzy sets expressed in a format closer to the human language.

C. Decision Trees

This type of method generates a model in the format of a tree which maps observations used to predict the value of a target value. It uses Divide-And-Conquer technique which consists on dividing the samples in a set of subsets and then it uses each subset to create a rule or set of rules. The rules created using each subset are finally merged. In the tree representation leaves target values whereas branches represent conjunctions of features that lead to the target values. The decision trees methods using numeric target values are known as regression trees. Some examples of this type of method are ADTree, Random Tree and REPTree.

- **ADTree**: It applies boosting procedures to generate the classifiers which are in the form of majority vote over a number of classifiers what delivers smaller an easier to understand classification rules.
- **RandomTree**: This algorithm creates a model considering a number of randomly selected attributes at each node and it does not perform any type of pruning.
- **REPTree**: It constructs the predictive model by using information gain/variance and it prunes the tree by considering reduced-error with back-fitting. This method only sorts for numeric features once at the initial run and then uses the sorted list to decide the splits.

The advantage of decision trees is that they are easy to interpret after a brief explanation and don’t need a large dataset to create a meaningful tree. It also enables the easy integration of new scenarios and its integration within other decision techniques such as Random Forest [157],[158], [159].

D. **M5rules**

This method constructs a decision list for regression problems using Divide-And-Conquer technique. It builds a model using M5 decision tree in each iteration and makes a leaf out of the best rule generated. M5 is a robust and efficient method [170] used for instance in the water sector [181],[182]. The “IF THEN” format of the explicit rules generated by M5rules method often enables the quick understanding of the predictive model generated [169]–[171].

This algorithm uses an estimate of the expected error at each node using test data. This error is calculated using training data which leads to underestimating the real error that would exist when using testing data. To account for that a penalty is added by multiplying the error by a factor. Moreover, M5rules generates a linear model for each node of the unpruned tree. This model is calculated using standard regression and it is simplified by dropping terms until error stops decreasing. After that, the model is pruned from the leaves as long as the estimated error decreases.

4.2. Adoption of an existing methodology to manage the knowledge life cycle

Two key features are considered as the foundations of this work: the use of AI knowledge-based applications to source knowledge, and the use of methodological support to adequately manage the knowledge generated. The adopted methodology allows the integration of KBE applications into engineering workflows, thus facilitating the adoption of the proposed framework in the industry.

In the previous decades, several methodologies supporting the development and maintenance of knowledge based systems (KBS) have been realised [183]. In this context, three methodologies were analysed and one of them was selected as the most suitable to be adopted in this project. The three methodologies studied are: MOKA (Methodology and Software tools Oriented to Knowledge-based engineering Applications), CommonKADS (Common Knowledge Acquisition and Documentation Structuring) and KNOMAD (Knowledge Nurture for Optimal Multidisciplinary Analysis and Design). MOKA methodology is based on eight KBE life cycles steps which are classified within 3 different stages [82] as shown in Table 16.

Table 16. MOKA methodology

Stages	Life cycle steps	Description
Stage 1	Identify	Knowledge identification and management approval to continue with the process.
	Justify	
Stage 2	Capture	Knowledge collection using elicitation methods and representation of the knowledge captured in a formal, consistent and standard manner.
	Formalise	
Stage 3	Package	Design and implementation of the KBE system.
	Distribute	
	Introduce	
	Use	

CommonKADS presents many similarities with the MOKA approach and it is also divided in three main layers [83]. The first layer, which is related to the organisation task and agent models, is similar to the first stage of MOKA where the opportunities to develop KBS are identified. The second layer is the most widely used component of CommonKADS. In fact, the expert model is created in the process linked to this layer. The expert model contains three different types of knowledge: task, inference and domain knowledge. The creation of the expert model is similar to the second stage of MOKA [183]. In this step, knowledge is captured and formalised to enable its later reuse. Finally, the creation of the design model and its posterior implementation are carried out in the third and last layer.

MOKA and CommonKADS have been acknowledged by the research community as the most relevant methodologies when talking about KBE development [183]. However there are some challenges remaining related to their main drawbacks as described in [183]. In this context, CommonKADS provides the user with a set of guidelines and templates to complete the tasks improving repeatability. Nevertheless, the use of these guidelines and templates decreases its flexibility. Users also complain that there are not enough templates to cover most of their current tasks. The main drawback of MOKA is identified in [183] where it is argued that this methodology is more focused on supporting knowledge engineers than the end user. Moreover, a common limitation to both approaches –MOKA and CommonKADS– is the incapacity to deal with knowledge change, not accounting for its origin and the repercussions of its change.

Due to the limitations described, it is observed that a set of challenges in terms of developing KBE systems still remain open. A potential solution adopted in this study to overcome these challenges is the use of an updated version of KNOMAD methodology. A description of the KNOMAD methodology together with an explanation of why it was chosen to be used in this thesis is described below.

4.2.1. KNOMAD Methodology

As MOKA and CommonKADS, KNOMAD (Knowledge Nurture for Optimal Multidisciplinary Analysis and Design) is a methodology created to support the development of KBE systems [184]. This methodology encompasses a set of steps that might be repeated in order to achieve a proper management of the knowledge life cycle. It consists of the following steps [183]:

- **Knowledge capture.** Scope, objectives and assumptions of the project are identified and knowledge is extracted from two different sources: explicit and tacit sources. To capture engineering knowledge efficiently this methodology advises the use of existing acquisition methods. On one hand, the use of natural techniques such as interviews and group meetings is convenient when trying to capture explicit knowledge. On the other hand, if tacit knowledge needs to acquired card sorting or three card trick are usually the methods employed.
- **Normalisation.** After the knowledge is captured, it goes through quality control and normalisation processes where this knowledge is broken down into a set of relations removing data redundancies and ensuring data consistency. In the normalisation stage data is stored within informal and formal models defined by

the knowledge manager. The use of informal model enables the easy storage of unstructured data containing various data type format. In parallel, the use of formal models permits the use of knowledge elements within automated processes (e.g. storage of machine learning rules stored in formal models which are later on retrieved to automatically generate the requested predictions of new data samples). This task must be carried out to comply with the necessary requirements to proceed with the next stages of the methodology.

- **Organisation.** In this step, the data structure is defined improving knowledge storage and accessibility. By defining the data structure, the knowledge becomes computable and independent from its application. This means that it is possible to retrieve relevant bits of this knowledge stored and use them in other problems across the organisation. At this stage, the use of ontology is advised in order to increase knowledge transparency and applicability (see 4.2.2).
- **Modelling.** To model the knowledge, informal and formal models belonging to MOKA and CommonKADS are used at this point of the methodology. In this stage of the KNOMAD methodology three Enterprise Knowledge Resources (EKR's) have been created: "Knowledge", "Applications" and "Case Reports". The EKR's defined enhance a better management of the knowledge life cycle by allowing to independently store the elements of the proposed framework into knowledge packages managed in a central Knowledge Repository. Once the knowledge has been modelled, this phase is followed by the implementation of these models.
- **Analysis.** The analysis of the methodology is a crucial task in order to achieve a suitable KBE system. Qualitative and quantitative analysis of the created EKR's are realised in order to ensure the proper functionality of the system. This analysis is crucial to validate the case studies implemented in the knowledge sourcing framework developed in this study –see section 6.3 for more details about the validation process.

- **Delivery.** Finally, once the platform integrating the corresponding methods and tools is developed, and after the system is reviewed and validated by experts in the domain, the knowledge sourcing capability is delivered to the stakeholders.

A detailed analysis classifying the three methodologies previously mentioned used to develop KBE systems are shown in

Table 17. The decision criteria followed in the table below is based on the extensive analysis of MOKA, CommonKADS and KNOMAD carried out by [32],[184],[185] and specially by [183].

The outcome of the study presented in

Table 17 shows the advantage of using KNOMAD compared to its competitors. This is due to KNOMAD integrates the best practices employed by MOKA and CommonKADS. Moreover, KNOMAD uses an ontology to properly account for knowledge change, improving its applicability and transparency.

Table 17. Classification of methodologies supporting KBE development.

<table>
<tr><th rowspan="2">Property</th><th colspan="6">Methodology</th></tr>
<tr><th colspan="2">MOKA</th><th colspan="2">CommonKADS</th><th colspan="2">KNOMAD</th></tr>
<tr><td>Promotes Knowledge Reuse</td><td>↑</td><td>Knowledge capture and formalisation stages of the methodology enable knowledge reuse. However it doesn’t use guidelines to support users.</td><td>↑↑</td><td>Use of engineering templates supporting the realisation of certain tasks, enhancing knowledge reuse.</td><td>↑↑</td><td>It uses guidelines adopting a similar approach to the one employed by CommonKADS to capture, formalise and maintain knowledge.</td></tr>
<tr><td>Promotes Knowledge Change</td><td>↓</td><td colspan="3">Not clear what happen when knowledge changes (no guidelines).

Lack of investigation of the knowledge change consequences.

A costly solution proposed to respond against knowledge update is to go through all the steps of the methodology</td><td>↑↑</td><td>Use of Knowledge Lifecycle Model to cope with knowledge change. Implementation of a dynamic approach to face knowledge changes.</td></tr>
</table>

		again. Static approach to face knowledge changes.				
Flexibility	↑	Use of a graphical object-oriented technique (MOKA Modelling Language) to model the knowledge and display it at user level, thus enhancing the flexibility of the methodology.	↓	The use of guidelines improves replicability but decreases flexibility.	↑	Use of an ontology to deliver a more flexible approach.
Applicability	↓	It doesn't focus on user support.	↑↑	Use of guidelines to solve complex problems employing easy steps.	↑↑	Use of guidelines to solve complex problems employing easy steps. The use of ontology enables its wide applicability across different engineering problems.
Transparency	↓	They are not focus on knowledge transparency and accessibility.			↑	The use of an ontology-based EKR approach which users can make annotations enhancing the knowledge to be accessed, traced and maintained at any time.
Accessibility	↓					

↑ Optimal; ↑↑Highly optimal; ↓Property not perform in an effective manner

The conclusions derived from

Table 17 together with two key characteristics of KNOMAD are the main reasons why this methodology was selected to be adopted in this research. These relevant features provided by the KNOMAD methodology are:

- **Applicability:** the methodology is simple enough to promote its use across different engineering problems within the organisation, thus enabling the transfer and reuse of engineering knowledge. Moreover, the use of an ontology enables a wider applicability of the developed capability across different engineering problems.

- **Supports knowledge change:** In the context of this work, this is the most important feature of this methodology. For instance, new data is often generated causing changes in the rules modelling a problem. Therefore, in order to allow the machine learning algorithm to create better quality rules, the effective management of knowledge update is a must. It is acknowledged by researchers that the knowledge created by machine learning methods is only as good as the data used in the learning process. Therefore, keeping knowledge updated enable machine learning algorithms the delivery of more accurate predictions.

4.2.2. KNOMAD Methodology: KLC Ontology

To improve knowledge transparency and support knowledge change and traceability, conceptual models were used to create a KLC ontology. This ontology was created in [183] as part of the updated KNOMAD methodology and it has been adopted in this research. The reason to re-use this ontology is due to the close link between the researches context.

The language selected is UML which uses a set of classes representing KLC concepts. UML also contains certain aspects such as class attributes and relationships. The main classes of the UML diagram of the KLC ontology are (Figure 14):

- **Enterprise Knowledge Resource**. This class is the backbone of the existent KLC ontology. It represents the ontology functionalities and encompasses three main classes: EKR Knowledge, EKR Process and EKR Case.
- **EKR Knowledge.** It stores the knowledge that can be used by any other EKR element. For instance, –in the first case study of this research described in chapter 6– a list of wing design descriptors collected and stored within the knowledge class are used by the process element when the application is executed.
- **EKR Process.** It contains automated processes or applications. It is here where the user executes a specific inference tool such as the tool in charge of creating a new ML analysis report or the one that creates a new case prediction report.

- **EKR Case**. It collects data generated when the application placed in the EKR process element is run.
- **Product, Process and Resource:** These three classes are used to annotate EKR’s and their subclasses.

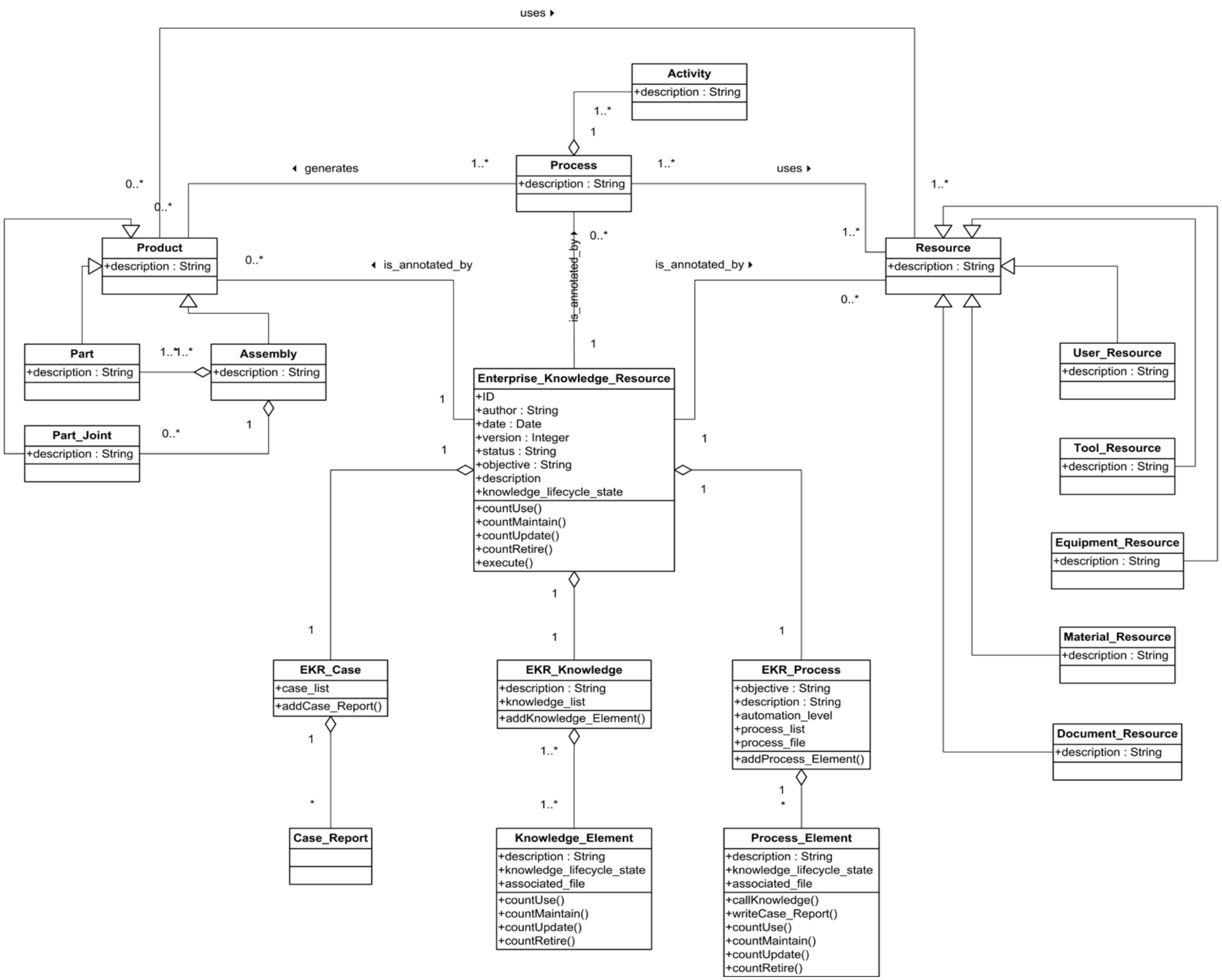


Figure 14.UML class diagram of KLC ontology [183]

4.3. Knowledge sourcing platform

A knowledge sourcing platform has been realised with the aim of achieving an efficient knowledge acquisition process while enabling the systematic creation, capture and reuse of engineering knowledge. The main elements of the KSF architecture integrated within the platform are:

- **Content repository**. It is a web-based content management system built on top of a database allowing the storage, access and modification of the knowledge captured. As described in section 4.3, it contains two different interfaces: administrator and user interface. The administrator interface permits the development of the platform functionalities and webpage composition within this content repository. In parallel, the user interface enables the interaction between the expert and the content stored, thus permitting the execution of any inference tool embedded within the platform.
- **Machine learning libraries**. Two different ML libraries have been considered in this research: Weka and scikit-learn. The most common methods belonging to these libraries have been analysed in section 4.1. The selection of the methods belonging to these libraries is based on:

    - The accuracy of the results obtained in the learning process.
    - The meaningfulness and understanding of the rules provided which will be later used by the algorithm to generate the target class predictions.

  After the analysis is performed, the selected methods are encoded within the platform to allow their automated execution.

- **Scripts.** Most of the scripts were developed with the aim of automating ML libraries. In this regard, the use of java and python scripts was needed to automate Weka and scikit-learn libraries respectively. The rest of the scripts were written in JavaScript and PHP to allow the expert and web-based platform interaction. These scripts are linked to the functionalities of the content repository such as searching, filtering and storing data.

The framework proposed was developed using a web-based content management system –open source– built on top of a database. This system contains two different interfaces named as “administrator interface” and “user interface”. In the administrator interface, the data structure and a set of html files describing how the data is displayed to the end user are created. In the web-based user interface, information is displayed to the end user as it is often done in a website (Figure 15).

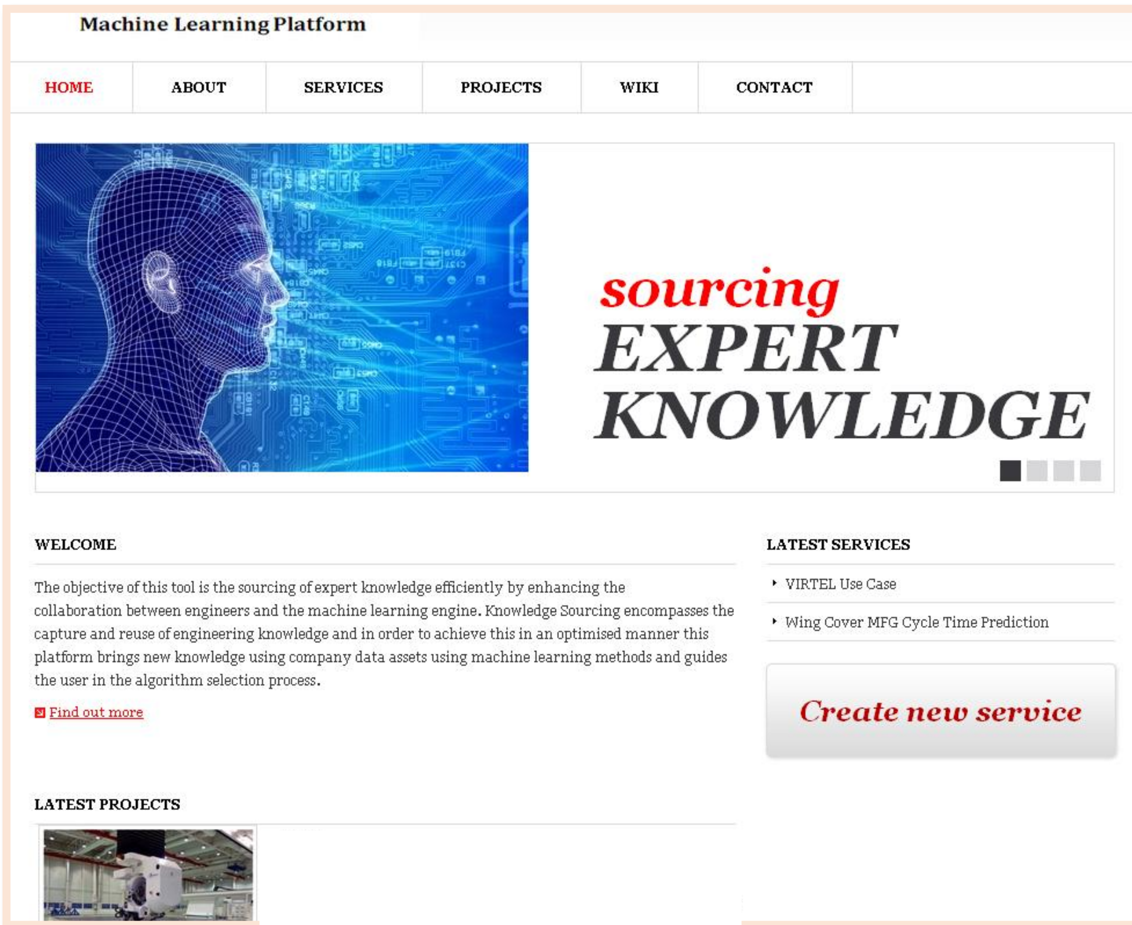


Figure 15. Content management system: User interface

The procedure to be followed in order to create a new service and obtain the predicted results is divided in two main phases: offline and online stages that are explained in details below. In this work, a service is considered as any process containing three EKR elements: knowledge, tools and case reports. The aim of a service is to predict a particular event by combining machine learning algorithms and expert knowledge. For instance, the first use case (see chapter 6) –based on a service that predicts the manufacturing cycle time to produce a wing cover– is composed of three EKR’s where the knowledge related to the service, the inference tool and reports generated are independently stored.

4.3.1. Offline Phase

The steps involved in this stage represent the framework backbone as they enable the creation of the EKR’s elements for a specific service. This phase encompasses four steps which are directly related to the KNOMAD methodology implemented in this platform. These four methodology steps are:

1. **Knowledge capture.** Initially, there is a need to define the problem objective or event that the user requires to be predicted, and identify the parameters that drive that particular event. In this process, experts and knowledge managers are usually involved. More precisely, knowledge managers use acquisition techniques to extract tacit and explicit knowledge from experts. The knowledge acquired permits the definition of an initial set of parameters that drive the event to be predicted. The list of parameters defined –affecting the analysed event– constitutes an initial model of the problem. This model will be used by the ML algorithm to build a more complex one –describing the problem behaviour– that will be employed in the prediction process.
2. **Normalisation and organisation of the knowledge captured.** Initially, the data capture goes through an extensive quality control. After it is reviewed by experts and knowledge managers, the specific ontology and data structure are created. The data structure is manually created within the content management system by using the administrator interface. The creation of the ontology and data structure facilitates the access to the data constituting the semantic backbone of the knowledge base implementation.
3. **Modelling of the knowledge.** In this step the knowledge models correspondent to the EKR components (knowledge, tools and case reports) are created. The main models are:

- Knowledge articles containing relevant knowledge associated to the service.
- Models including user interface functionalities (e.g. data upload)
- Case report articles containing the results generated by the machine learning algorithm.

The creation of these models enables the easy representation, storage and retrieval of the data captured and generated within the scope of a service.

4. **Implementation.** This step encompasses the development of a set of applications which enable the online execution of the service. The tools created in this step are common for any service apart from the feature recognition tool which is in charge of automatically capture input design data.

4.3.2. Online Phase

The online phase encompasses all the steps carried out by the user utilising the user interface excluding the automated tasks running in the background. In the first place, the user uploads a file (“Training Set”) containing the data that will be used in the automated learning process (Figure 16). After that, it is time to proceed with the main tasks of the online process which are the selection of the machine learning method and the posterior management of the model automatically generated by the machine learning algorithm.

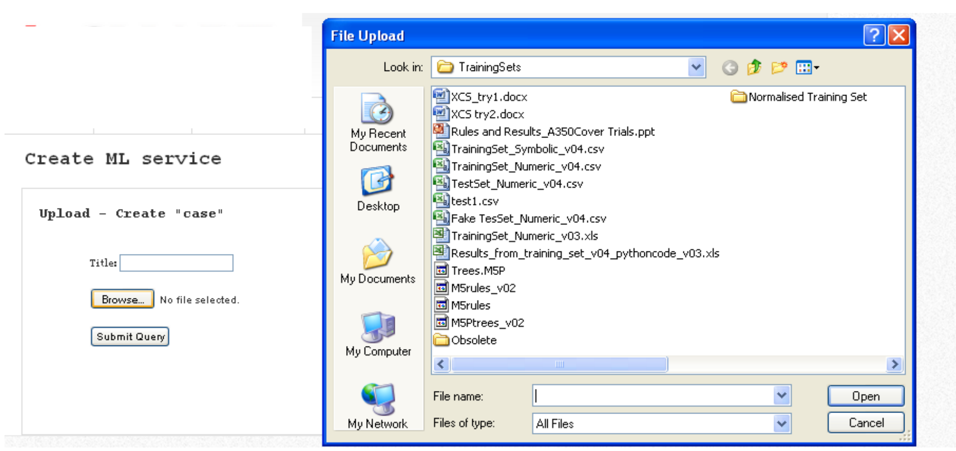


Figure 16. User interface: Upload input data set.

To select the adequate ML algorithm, the intervention of an expert in the domain is required. The expert will analyse the results obtained by the machine learning algorithm and will select the most suitable method (Figure 17).

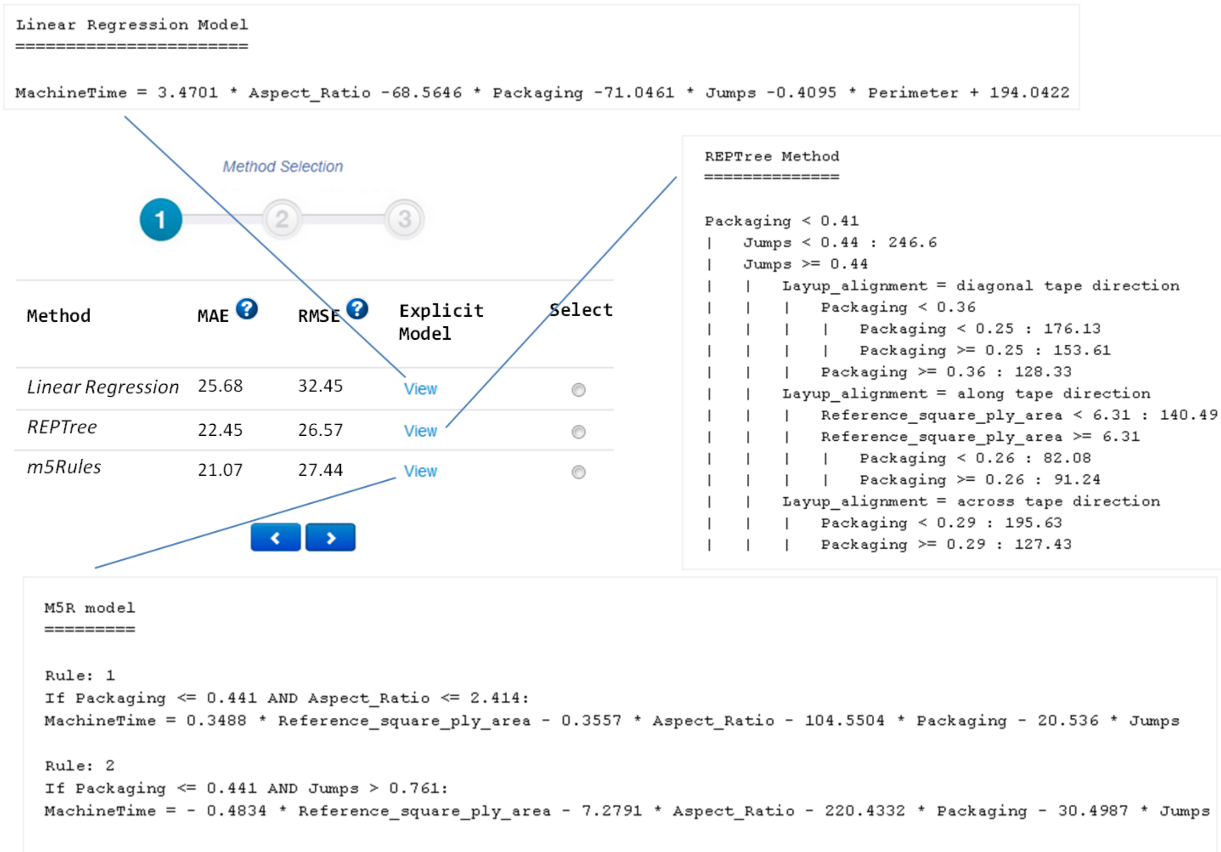


Figure 17. User interface: Method Selection

The expert criteria used in the selection process is based on the level of understanding of the rules automatically generated and the accuracy of the results generated in the prediction process. The accuracy of the algorithms is assessed through the analysis of two machine learning scoring values: Root Mean Squared Error (RMSE) and Mean Absolute Error (MAE). Once the method has been selected, the expert makes use of a set of data analytics tools to identify incoherencies in data and modify the rules in response (Figure 18). The review of the machine learning rules is an iterative process that ends with the rules' pre-validation step that is achieved when an expert or group of experts are satisfied with the results obtained.

To achieve a reliable validation of the rules, the model generated –previously pre-validated– is employed to predict the samples contained in the “Test Set” file. To avoid over fitting, the data contained in the “Test Set” file was not used in the learning process (not used in the creation of the rules). The validation of the explicit model is completed after the experts accept the accuracy of the results (predictions corresponding to the “Test Set” samples generated by the pre-validated rules) as acceptable from an engineering point of view.

After the model is validated, in case the user wants to obtain the predictions of a new design configuration, it is necessary first to upload an input data file containing information regarding design descriptors of the new design case. After the file containing design data is uploaded into to system the prediction process is automatically triggered and as a result the predictions are displayed on the screen. When a file is uploaded for prediction, a python script uses the validated ML rules to generate the corresponding predictions. Next to these predictions, a set of tables are displayed presenting the input data (corresponding to the design descriptors), predicted values, rule applied to predict each sample and the list of validated rules employed in the prediction process (Figure 19).

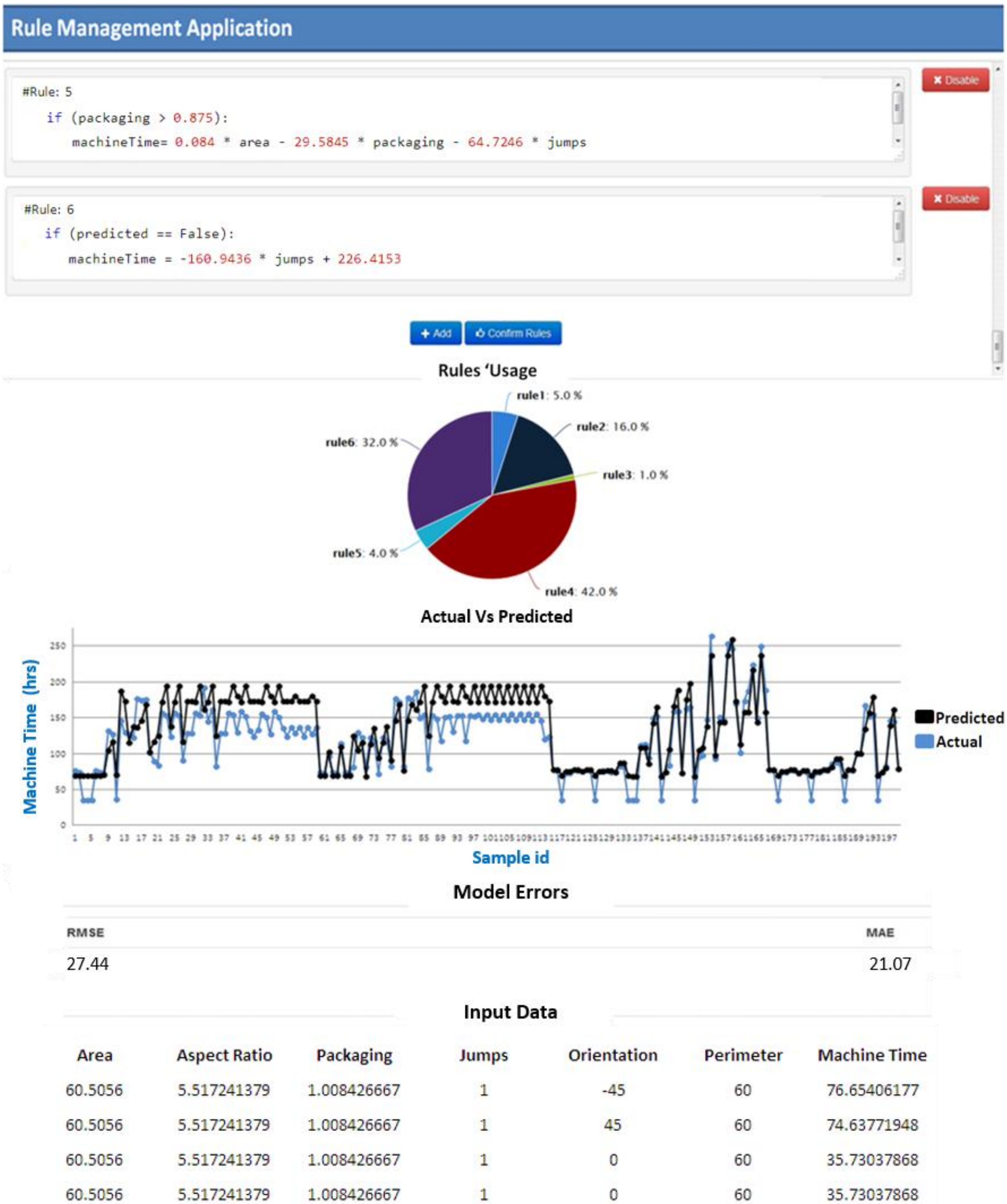


Figure 18. User interface: Rules' management Application

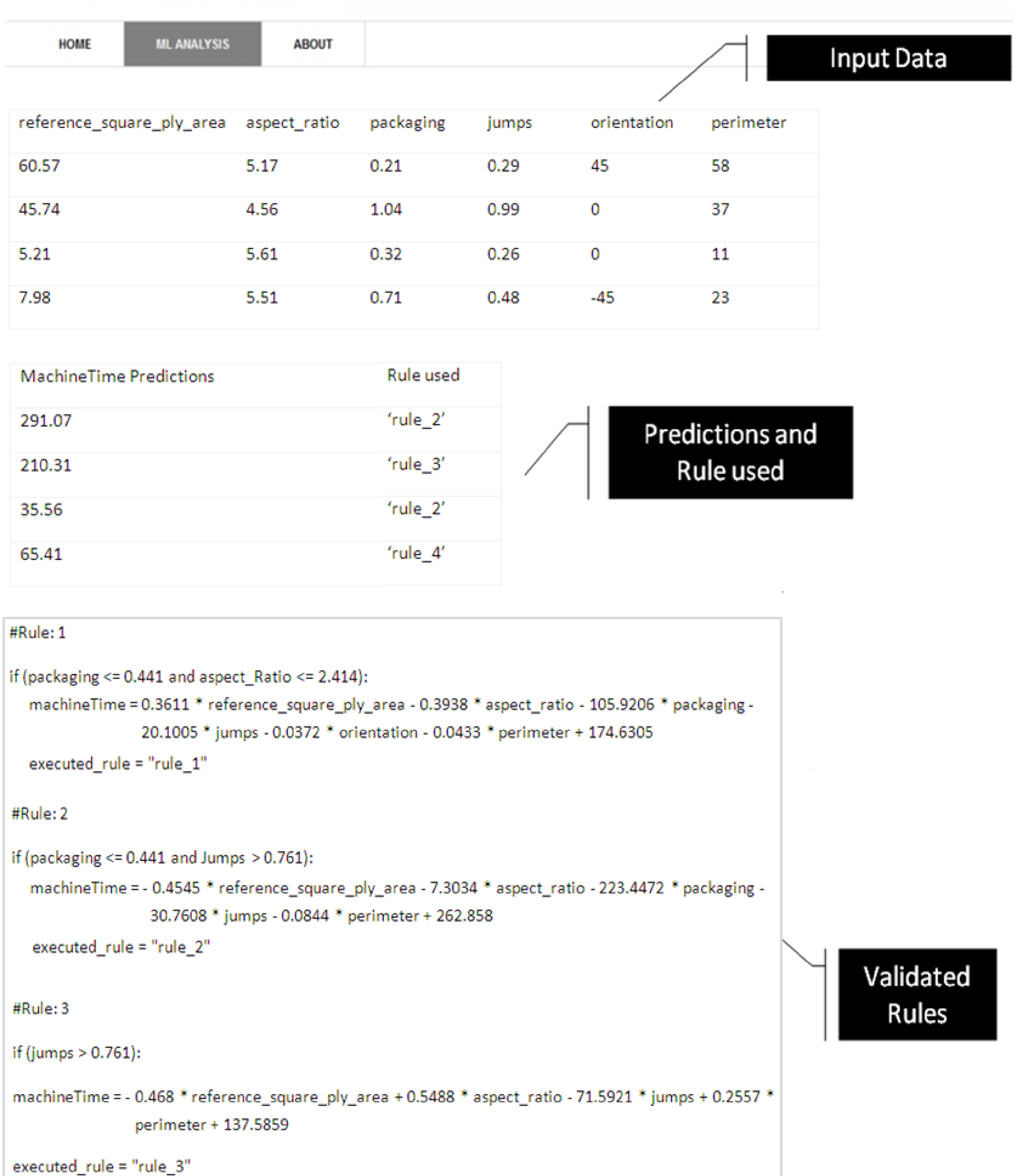


Figure 19. Case report generated when executing the prediction process of a new case.

4.3.3. Offline and online phases overview

In summary, the capture and pre-process of the data together with the creation of the EKR's models are classified as offline processes. The automated generation of AI rules, the review and validation of the rules, and automated prediction of new design configurations are included in the online phase. Figure 20 shows the procedure followed to create and execute a new service using the methodology proposed in this work.

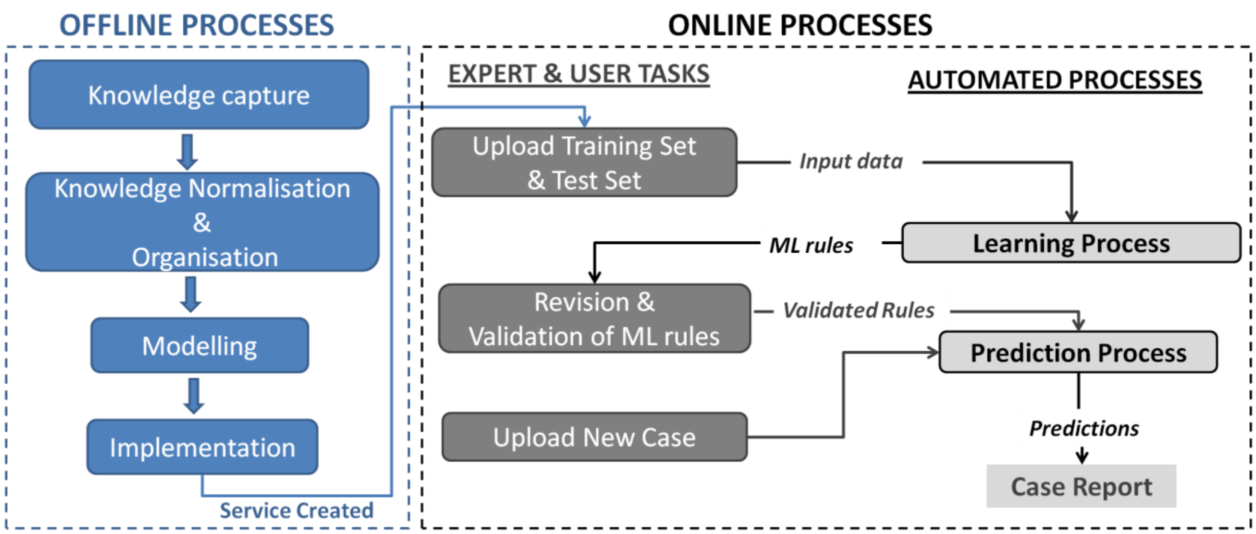


Figure 20. Service creation within KBE Platform: Process flow.

All the tasks involved in the development of a new service within the KBE platform are described in more details in the following two chapters where two different case studies are presented.

4.4. Contribution and novelty of the framework

The presented methodology is composed by two main features. The first feature implies the adoption of a generic methodology which enables the adequate knowledge management through the systematic source of engineering knowledge. The adoption of a methodology permitting the management of the complete knowledge life cycle allows the effective knowledge capture, retention, reuse and update. In doing so, the risk of knowledge loss and the time for carrying out the engineering design process are reduced.

The second feature is the integration of AI techniques and experts. The use of AI algorithms reduces the time to capture engineering knowledge whereas the review and validation of the knowledge generated by the AI method increases the reliability of the predictions. Therefore, combining AI methods and experts delivers a more efficient and reliable knowledge sourcing process.

The analysis of researches related to the first feature published in the last two decades denotes an increasing number of implementations adopting well established methodologies. In this regard, existing studies achieved the effective knowledge management through the implementation of established methodologies such as MOKA or CommonKADS [82],[109]. Literature related to the feature of the approach integrating AI methods and experts highlights the use of AI methods to create new knowledge [10],[65]. However, only a few researches consider the expert intervention as an essential feature [69],[70],[91].

Although some researches aim to tackle the first or the second feature of the methodology proposed, studies targeting both features have not been found in the literature. Therefore, the integration of both elements into a single framework enabling the effective management of the knowledge life cycle and the efficient source of engineering knowledge is considered by the author as the novelty of this PhD thesis.

The consecution of the integrated knowledge sourcing framework led to a set of research contributions described as follows:

- **Efficient creation of new knowledge**. The use of AI algorithms, providing the user with explicit rules which model a problem, enable not only capturing knowledge from experts but also from company data assets. Using these advanced techniques time required to capture expert knowledge is reduced.
- **Support engineers to use the most appropriate AI tool.** The knowledge sourcing framework encompasses a method which displays relevant information

about the executed AI algorithms such as explicit rules modelling the problem and some scoring values evaluating the accuracy of the results obtained by each algorithm. This interpretative provided facilitates the selection of a suitable AI technique for a specific problem.

- **Enhance KBE reliability.** The reliability of the KBE application is achieved through the use of a methodology which enables experts to review, modify and validate the explicit model generated by the AI tool. To do so, the rule management application embedded within the knowledge sourcing platform provides experts with visual analytical tools that facilitate the data analysis process.
- **Effective knowledge retention**. This is achieved through the adequate implementation of the formalisation KNOMAD step where knowledge is stored into informal and formal models.
- **Effective knowledge reuse**. It is permitted by the use of domain ontology in the creation of the knowledge architecture (organisation stage), and the data storage into knowledge packages (EKR's). On one hand, the use of a domain specific ontology enhances knowledge accessibility and traceability. On the other hand, the use of EKR's allows to independently store the elements of the proposed framework into knowledge packages which are managed in a central knowledge repository where knowledge is ready to be used by any KBE application.
- **Integration of KBE tools into engineering workflows.** KBE applications are embedded in a content management system set to follow KNOMAD steps. In doing so, KBE tools are ready to be used in workflows across different engineering problems.

4.5. Concluding remarks

The definition of the knowledge sourcing framework architecture has been reported in this chapter. Each element of the architecture has been design to fulfil the research objectives 2–6. The architecture created is divided in three elements:

1. Search, analysis and exploitation of AI methods for knowledge capture.
2. Adoption of a well-established methodology for generic and effective management of engineering knowledge.
3. Creation of a knowledge sourcing platform.

The first architecture element aims the identification, classification, and analysis of methods and tools for capturing engineering knowledge more efficiently through the use of AI algorithms. In this regard, some of the most widely used machine learning algorithms [88],[150],[151] were classified and analysed based on their ability to solve different type of problems (classification or regression problems), and generate an explicit model. Moreover, two well-known libraries for machine learning are integrated within the knowledge sourcing platform for automated knowledge capture. The research outcome related to this architecture item is the delivery of a method for fast knowledge capture reducing the need for expert availability.

The objective of the second element of the architecture is to ensure the correctness and validity of the methodology used to manage the complete knowledge life cycle, thus delivering a framework that realises the effective capture, retain, reuse, and update of relevant knowledge.

Finally, the objective of the third element of the architecture is to define the structure of the knowledge sourcing platform. The definition of the platform structure provides a baseline for the systematic and fast implementation of the case studies. In doing so, the appropriate application of the use cases of this research is ensured.

To provide the reader with a better understanding of how the established research objectives are achieved, technical contributions of the knowledge sourcing framework presented in this work are described in the following chapter.

5. Knowledge sourcing framework: Technical contributions

This chapter describes in more details the technical work carried out to achieve the research objectives related to the implementation stage of the KNOMAD methodology (objectives 2-5) listed in the table below.

Table 18. Description of the research objectives related to KNOMAD implementation.

Research Objective	Section	Description
5	5.2	To manage engineering knowledge separately from KBE applications by systematically storing KBE rules into a human readable format
2	5.3	To help engineers in identifying AI tools that are most appropriate for their particular problem
3	5.5	To enable efficient knowledge acquisition by exploiting AI tools capable of generating automated KBE rules from data assets
4	5.4	To enhance KBE reliability by the assessment of KBE rules, allowing the designers to identify the quality of the results obtained

Although the work performed towards the execution of these research targets is described in chapter 4, a more detail explanation focused on technical aspects is required to prove their achievement. Therefore, the outcome of this chapter is to show the achievement of the research objectives 2–5 from a technical perspective. To do that, the functionality of the framework is demonstrated by using simplified engineering problem, thus making the accomplishment of the targets more comprehensive.

This chapter is structured as follows. First a brief description of the engineering problem used in the demonstration of the knowledge sourcing framework is presented in section 5.1. Sections 5.2, 5.3,5.4 and 5.5 describe, from a technical point of view, how the research objectives are achieved. Finally, in section 5.6, a brief summary highlighting the achievements described in this chapter is presented.

5.1. Description of the engineering problem

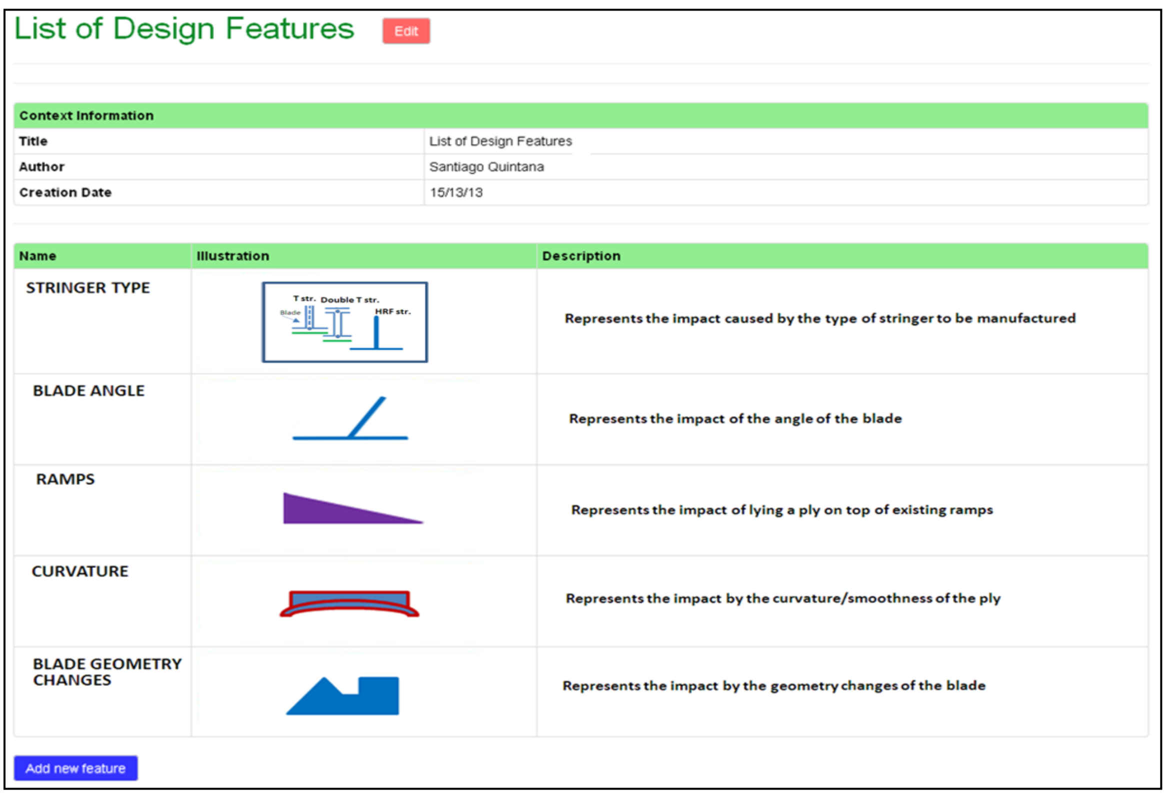
This chapter describes in details the implementation of the KSF developed in this research to solve an engineering problem. The context of this engineering problem is the aerospace industry. More precisely, the context is the quality estimation of Carbon Fibre Reinforced Plastic (CFRP) wing stringers. The values of the output class (quality class) belonging to the dataset –used by the machine learning application in this engineering problem to create a model of the problem– where provided by domain experts. These values corresponding to the quality class where defined by experts in the wing stringers manufacturing context by visually analysing each of the design samples. Certain characteristics of the wing stringers such as wrinkles on the surface and delamination issues where taken into account by the domain experts to define the quality value for each stringer sample.

These stringers are manufactured using Advanced Fibre Placement (AFP) technology. AFP is considered by the engineering community as an immature technology requiring further research in order to identify the design features of a wing stringer that drive part quality. In this direction, the use of the knowledge sourcing capability developed in this research provides a solution capable of facilitating the understanding of what design parameters drive stringer quality.

5.2. Management of engineering knowledge separately from its application

The effective management of the knowledge created within the KSF is considered an essential feature of this work. To do that, an established methodology to develop a KBE system has been adopted (KNOMAD). The use of this methodology enables the storage of the knowledge captured within formal and informal models.

In the engineering case reported in this chapter (aiming at fast estimation of wing stringer quality) the list of the design descriptors experts believe have an impact on stringer quality are stored within informal models as shown in Figure 21. In parallel, the rules automatically generated by the AI tool and later review and validated by experts are stored within formal models in which knowledge stored is in a computer and human readable format (Figure 22). In doing so, the rules modelling the problem become more comprehensive and they can be directly executed by any KBE application.


List of Design Features Edit

Context Information	
Title	List of Design Features
Author	Santiago Quintana
Creation Date	15/13/13

Name	Illustration	Description
STRINGER TYPE	T str. Double T str. Blade HRF str.	Represents the impact caused by the type of stringer to be manufactured
BLADE ANGLE		Represents the impact of the angle of the blade
RAMPS		Represents the impact of lying a ply on top of existing ramps
CURVATURE		Represents the impact by the curvature/smoothness of the ply
BLADE GEOMETRY CHANGES		Represents the impact by the geometry changes of the blade

Add new feature

Figure 21. Informal model of stringer quality prediction service.

Figure 22. Formal model of stringer quality prediction service.

These informal and formal models are contained within an Enterprise Knowledge Resource (EKR) class named as “Knowledge”. EKR’s are knowledge packages that enable the knowledge to be stored separately from its application, thus fostering knowledge reuse and facilitating knowledge maintenance. Including the “Knowledge” class, the KSF platform created contains two more EKR’s named as “Applications” and “Case Reports” as shown in Figure 23. “Applications” class contains the models associated to the process integrating machine and expert knowledge in which new knowledge in the form of rules is generated, reviewed and validated. In parallel, “Case Reports” class contains models capturing the data used and generated in the machine learning process.

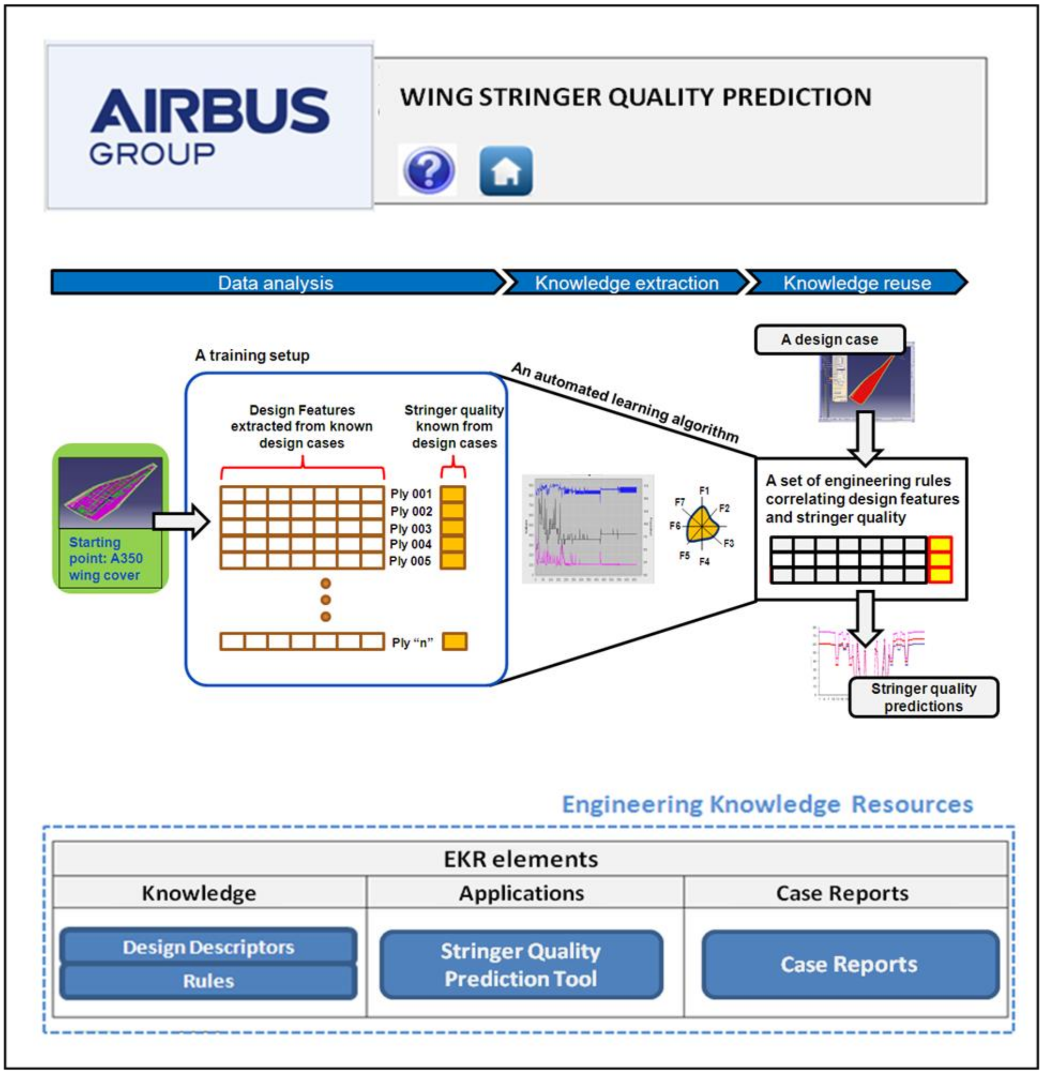


Figure 23. Front page KSF service aiming at predicting stringer quality.

5.3. Support in the identification of a suitable AI algorithm

A key feature of this work is the use of artificial intelligence algorithms which enable the creation of new knowledge from non-processed data. More precisely, the algorithms selected in this research belong to the machine learning (ML) field. In the ML area there is a considerable number of methods which can be used to automate the knowledge elicitation process as described in section 2.3.2 where a set of examples are described. However, techniques creating explicit models or rules from raw data have gained the engineers attention due to their higher reliability compared to the ones classified as "black box" applications [85] [86].

In this context, only the algorithms delivering an explicit model were employed in this work. The selection of these methods is described in section 6.4.2 where Table 15 classifies some of the most common supervised learning methods based on the type of method that they can deal with and the level of understanding of the interpretative information provided. The methods evaluated in Table 15 delivering an explicit model were embedded within the KSF.

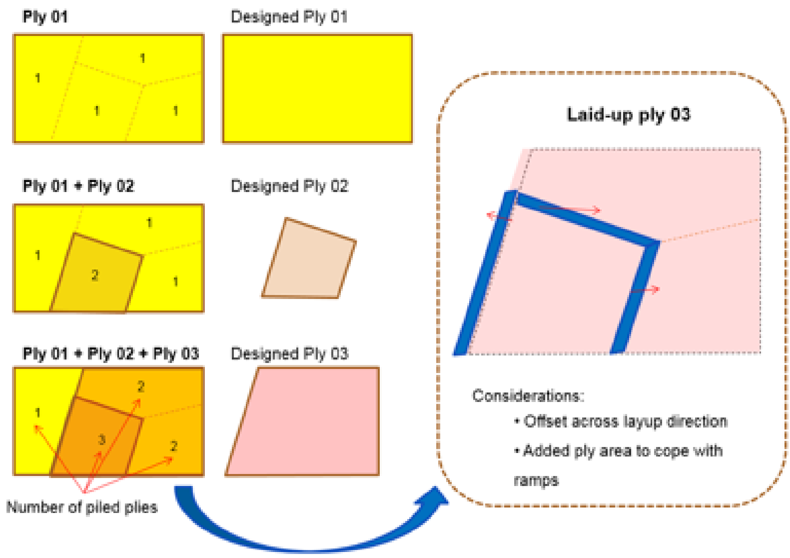
The KSF platform developed encompasses a method where the user is guided in the selection of a suitable AI method and the review and validation of the model automatically generated by the AI technique. In the AI algorithm selection step, the user is asked to upload an input file containing design data samples and their corresponding stringer quality (output class values). In the case of the stringer quality prediction, the user uploads a file including 75% of all the samples gathered in the capture process which is composed of the values corresponding to the design descriptors (stringer type, blade angle, ramps, curvature and blade geometry changes) and the output values (stringer quality) as illustrated in Figure 24.

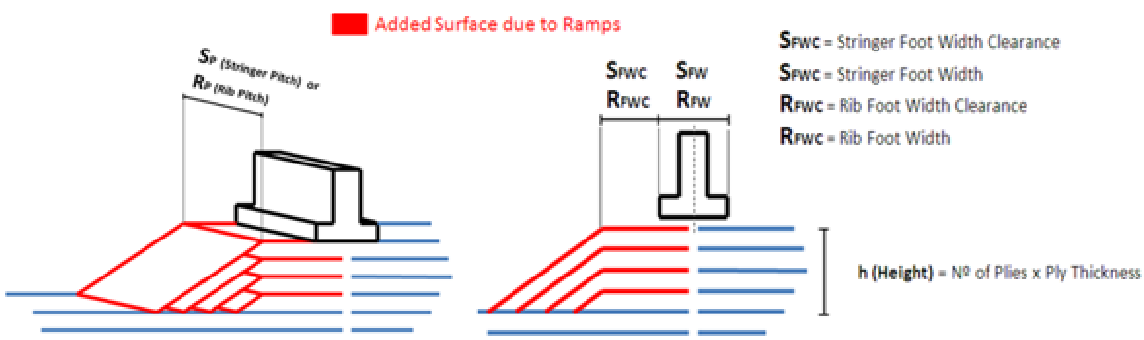
Design Descriptors or Input Data **Output data**

Id	**Stringer type** *t → T stringer* *dt → Double t stringer* *hrf → Hot roll forming stringer*	**Blade angle** *0 → Angle = 90°* *1 → Angle = 90° ±10 °* *2 → 100° > Angle < 80°*	**Blade geometry changes** *0 → None* *1 → Low* *2 → Medium* *3 → High*	**Ramps** *0 → None* *1 → Low* *2 → Medium* *3 → High*	**Curvature** *0 → None* *1 → Low* *2 → Medium* *3 → High*	**Stringer quality**
1	t	1	0	0	1	4.75
2	t	1	2	1	2	3.5
3	t	2	1	1	1	4
4	hrf	0	1	3	1	2
5	hrf	2	3	1	3	1
6	hrf	1	1	2	1	4
97	hrf	1	3	2	3	2.5
98	dt	1	1	2	2	3.75
99	dt	2	2	2	3	2.75
100	dt	2	1	2	3	3

Figure 24.Prediction of stringer quality service: Training set.

The design descriptors were automatically captured through the use of a feature recognition tool whereas the stringer quality predictions were generated by experts in the domain who scored from 0 (lowest quality) to 5 (best quality) each of the samples. The feature recognition capability developed in this case is in charge of extracting the design descriptors data that characterise a design. The feature recognition output is a file including the information displayed in the figure below. This tool consists of an algorithm developed in this research composed of a set of Python scripts. These Python scripts have the objective of opening an excel file containing the designs of the wing stringers and extracting the information corresponding to a set of design descriptors initially defined by the domain experts (stringer type, blade angle, geometry changes, curvature, width and length of the stringer elements, etc.). More particularly, to gather the design descriptors data a python script goes iteratively through each line of the excel file containing the designs and retrieves the information in each spreadsheet cell to compute each of the design descriptors. For instance, the calculation of the ramps is made overlying each layer on top of the other and identifying the areas where there is a change in the number of plies below (identification of slopes in the material) as shown in Figure 25.





Ramp Ratio (Assumptions for Covers)

- If Ramp between Stringers:

  Ratio= $\Delta y/\Delta x = 1/10$

- If Ramp between Ribs:

  Ratio= $\Delta y/\Delta x = 1/20$

Height = Ratio x Base

Ramp Area = ½ x Base x Height = ½ x Base² x Ratio

Ramp between Stringers:

$$\boxed{\textbf{Added Surface per Ply } (A_{SP}) = \text{Ramp Area} + S_{FWC} + \tfrac{1}{2} S_{FW}}$$

Figure 25.Feature recognition tool: Ramp calculation

Once the file is uploaded into the system, a set of algorithms are displayed in a table which contains relevant information of each algorithm (Figure 26). The information shown to the user includes a link to the explicit model created, the values of two ML scoring parameters (RMSE and MAE) and a link to a page containing information regarding how RMSE and MAE are calculated. Below it is described how RMSE and MAE are calculated in this research.

Mean Absolute Error (MAE): Mean absolute error is the average of the absolute difference between predicted and actual value of the samples [186] (Equation 1).

Equation 1. Mean Absolute Error

$$\mathrm{MAE} = \frac{|a_1 - c_1| + |a_2 - c_2| + \cdots + |a_n - c_n|}{n}$$

Root Mean-Squared Error (RMSE): Root Mean Square Error, RMSE is the squared root of the average of the squared difference between predicted and actual value [186] (Equation 2).

Equation 2. Root Mean Squared Error

$$\mathrm{RMSE} = \sqrt{\frac{(a_1 - c_1)^2 + (a_2 - c_2)^2 + \cdots + (a_n - c_n)^2}{n}}$$

In the above equations "a" is the actual value of the output, "c" is the predicted value of the output. For MAE and RMSE a lower value means a more precise model where a value of 0 would be a perfect model.

Using the information provided by the platform, the user selects a ML algorithm by evaluating the format and meaningfulness of the rules together with the accuracy of the predictions (RMSE and MAE). In the stringer quality prediction case, three algorithms were displayed on the screen together with their respective machine learning score values and models generated. After evaluation of the models and accuracy values, the experts selected "m5Rules" algorithm as the most suitable for delivering the required predictions. This decision was made due to the values for MAE and RMSE were the lowest compared to the other algorithms and the model generated by “m5Rules” method was the more coherent and meaningful from the domain expert point of view.

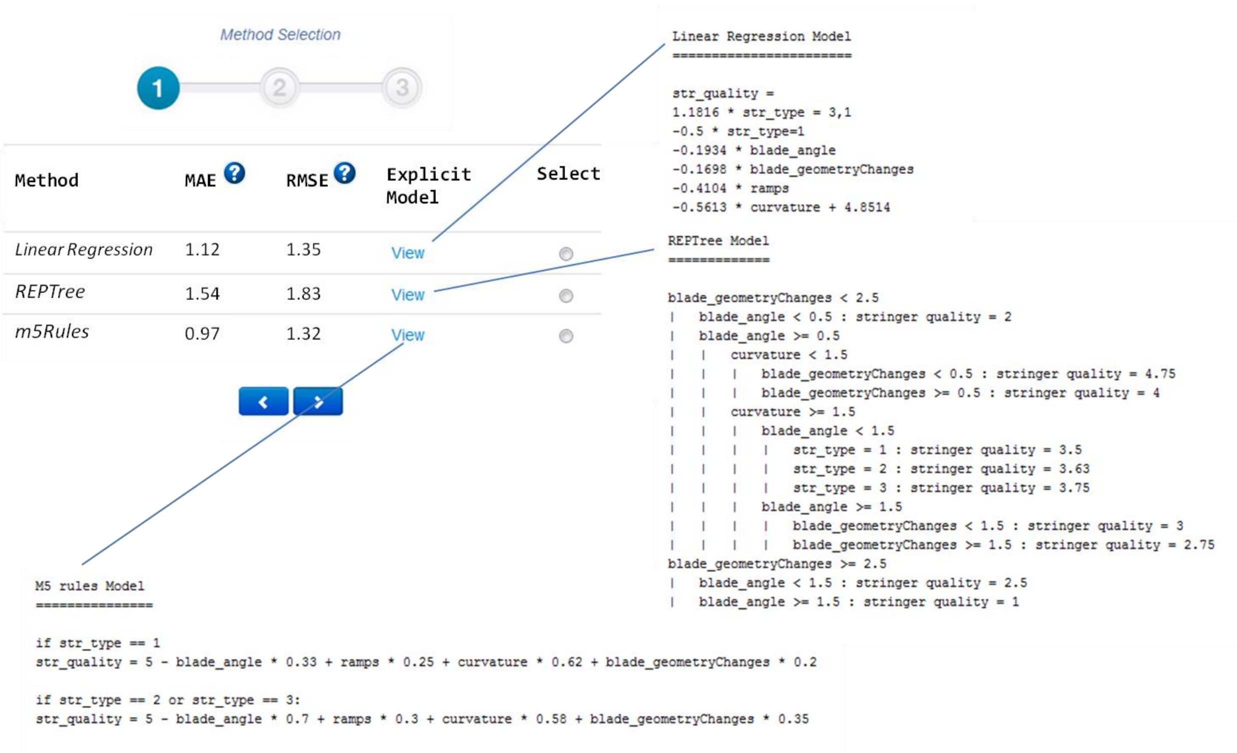


Figure 26. Explicit models displayed by the KSF platform

5.4. Enhancement of KBE reliability

The extended KBE development process proposed in this research encompasses the use of ML methods to capture engineering knowledge more efficiently. The analysis of some of the most common ML algorithms made apparent the lack of interpretative information (Table 15). These methods are known by the research community as “black box” applications and the use of these techniques in the engineering community has caused some reliability issues due to the inability to trace back the results delivered by the AI tools. However, there are a few ML techniques which provide engineers with a set of rules describing the problem, not only allowing users to trace back the results but also enabling them to gain a better understanding of the problem.

To enhance the reliability of the extended KBE methodology, this research proposes a combined approach integrating ML methods which deliver an explicit model, and expert knowledge in the form of expert intervention. A key feature of this work is the use of experts in the domain to carry out the analysis of the rules generated by the AI method. This task is realised using a web-based capability which enables the systematic review and validation of the rules. Figure 27 shows for the stringer quality prediction case the Rule Management Application (RMA) created and embedded within the KSF platform. This application permits the analysis of the results before and after the user creates a new rule or the user modifies or deletes any of the existing rules.

In this example, the errors (RMSE and MAE) initially obtained were over 1 unit and experts considered as acceptable a model delivering accuracy values lower than 1. Therefore, it was required to reduce MAE and RMA scoring values. RMA supports experts in reviewing the rules by displaying a set of charts and tables with interpretative information. This helps users to understand how the changes made in rules affect the estimations generated. Moreover, the machine learning scoring values and visualisations are updated after the user modifies the rules, thus helping the user to rapidly identify if the changes made are giving the expected results.

In the stringer quality prediction problem, experts observed in the RMA visualisations (hovering over the data points in the line chart) that the samples with higher deviation (observed when comparing the predicted values Vs actual values) were generated by the rule number 2 and belonged to the “hot roll forming” stringer type (“hrf” in the figure below). In addition, experts realised that the deviations were mainly caused by the “blade_geometryChanges” parameter which causes a higher impact on “hot roll forming” stringers. In this context, to reduce the deviations caused by rule number 2, the experts went through an iterative process of modifying this rule and evaluating how the changes made were affecting to the machine learning scoring values (RMSE and MAE) and to the data points in the line chart. Finally, experts achieved the reduction of the errors by modifying rule number 2 (to only be applied for “double T” stringers) and adding a new rule (only applied for “hot roll forming” stringers). This can be observed comparing Figure 27 to Figure 28.

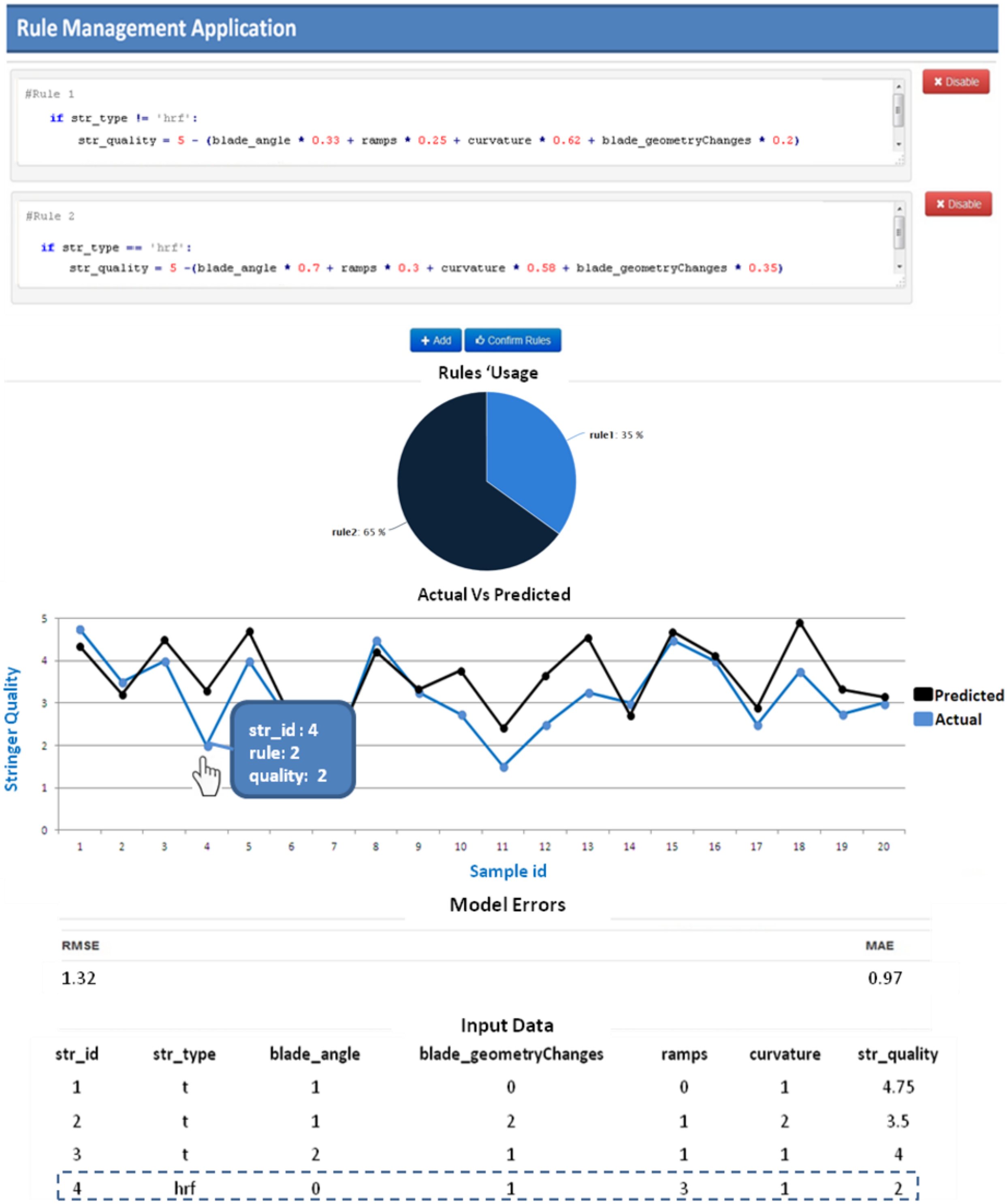


Figure 27. RMA: Rules initially created by the machine learning algorithms

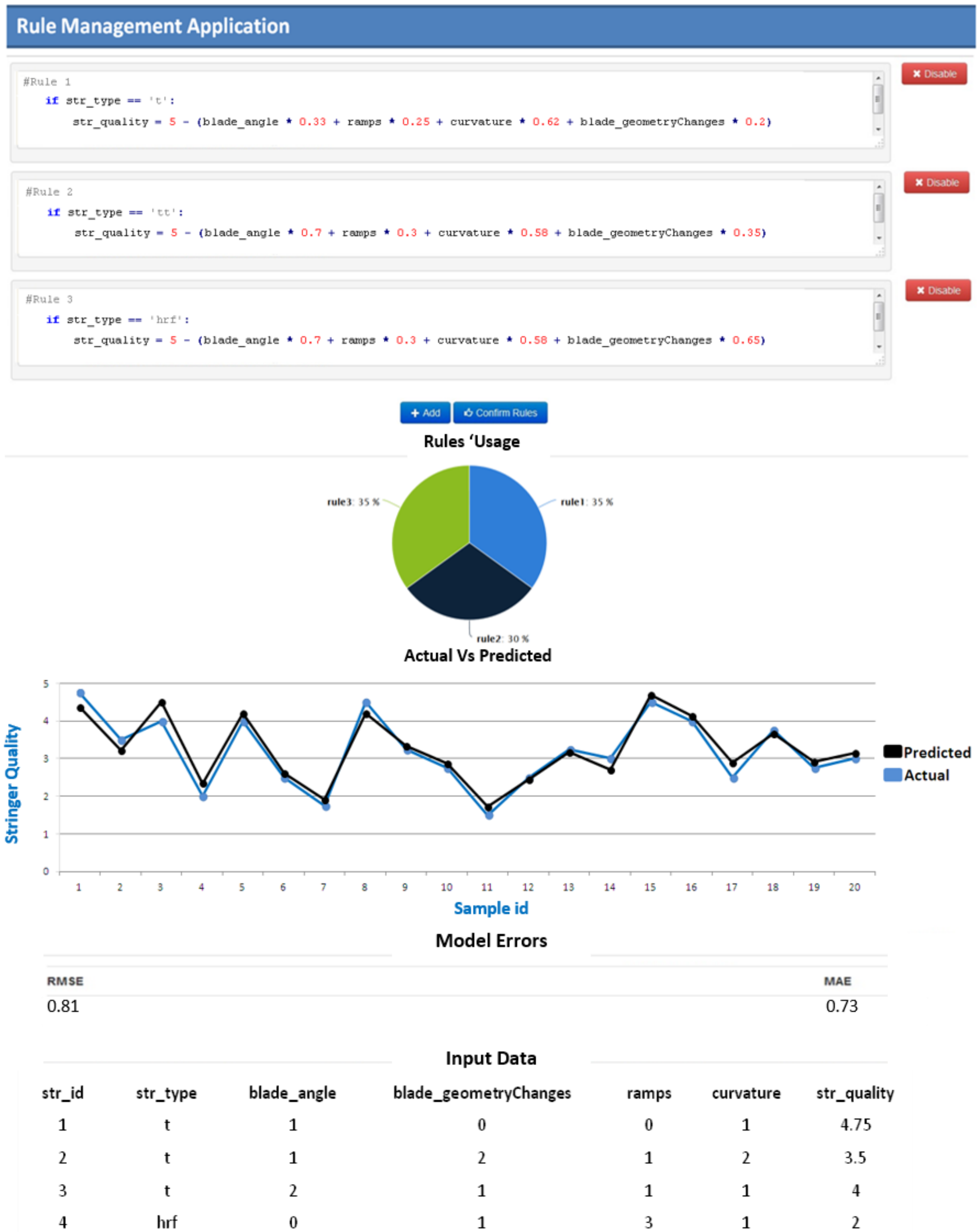


Figure 28. Rule obtained after expert review and pre-validation activities

5.5. Efficient knowledge capture using AI tools

Knowledge capture is considered together with knowledge normalisation as the major bottleneck for KBE development [5]. Current approaches to generate new knowledge in the form of a set of rules modelling the behaviour of a specific problem are highly time consuming tasks for experts. Moreover, experts facing complex problems having poor theory understanding, usually involving immature technologies, often miss correlations in data leading to incomplete models that provide inaccurate estimations. In order to reduce the time in extracting knowledge and deliver solutions for complex problems, this research introduces a new knowledge sourcing framework that integrates experts and AI algorithms as presented in Figure 29.

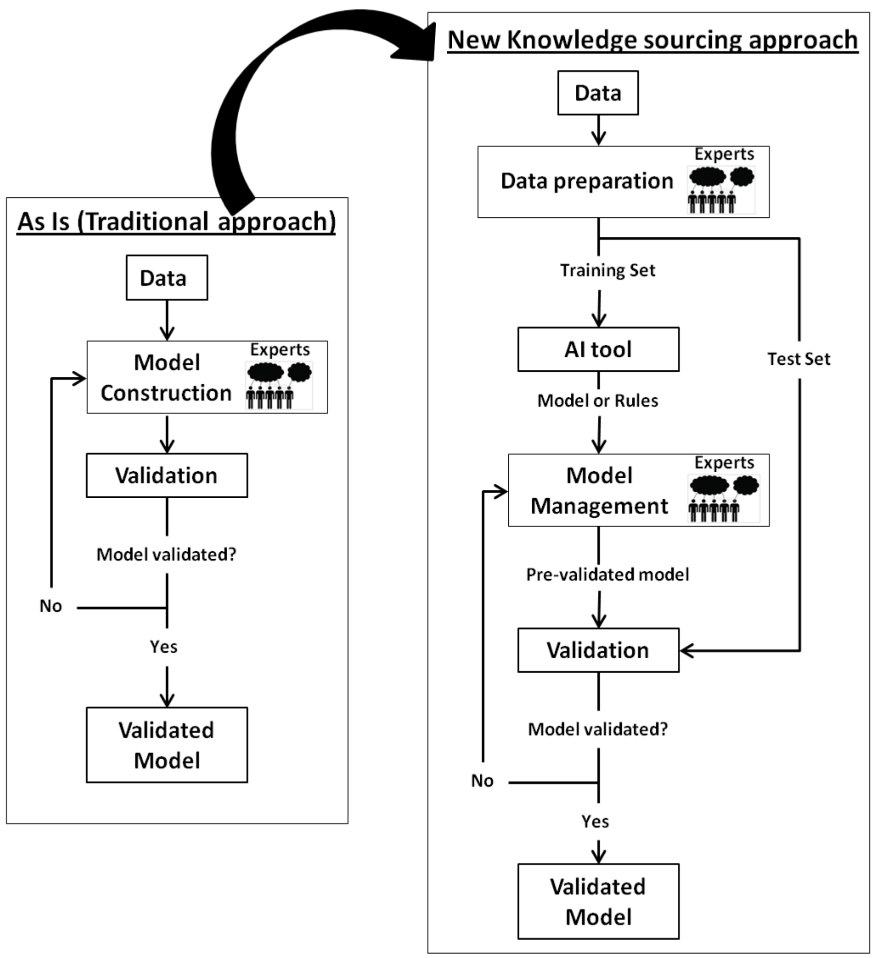


Figure 29. Knowledge capture process: Traditional approach Vs Proposed approach

With the new approach, instead of asking experts to generate a set of rules modelling the stringer quality problem, experts are only required to provide a list of the parameters they believe are driving stringer quality. Figure 30 shows the reduced list of parameters the experts believe have an impact on stringer quality that is obtained from an initial set of parameters related to the stringer manufacturing context.

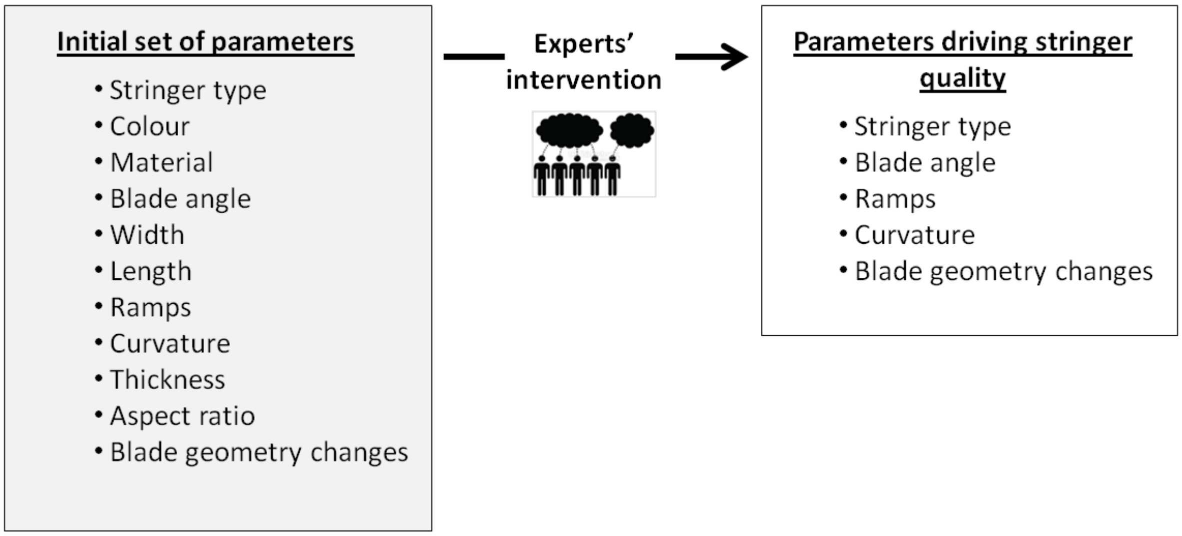


Figure 30. Stringer quality prediction: Data preparation.

The data captured corresponding to the parameters driving stringer quality is divided in two files: “Training Set” and “Test Set”. The “Training Set” is used in the learning process where a machine learning algorithm creates a set of rules modelling the stringer quality prediction problem. In parallel, the “Test Set” is used to certify the model reviewed and pre-validated by experts.

AI algorithms are well suited for process optimisation but these can also be used to source expert knowledge [10],[65]. The appropriate use of AI delivers a knowledge sourcing process which captures expert knowledge more efficiently. To do this, this research integrates two AI engines (WEKA and scikit-learn) within the KSF platform. Additionally, several java and python scripts were created to automate the execution of the WEKA and scikit-learn algorithms respectively (Figure 31).

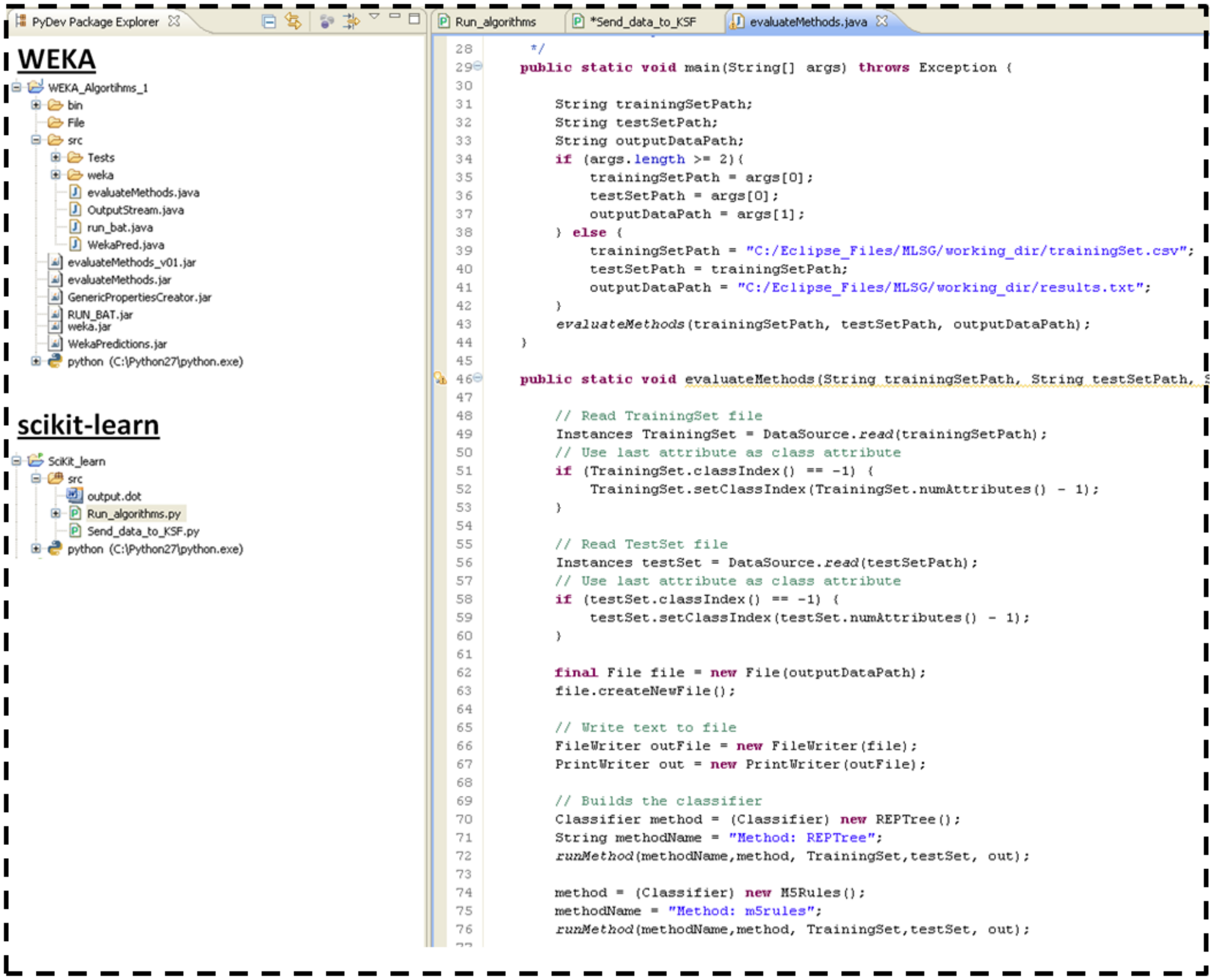


Figure 31. WEKA and scikit-learn scripts

Once each of the algorithms embedded in the KSF generates an explicit model, experts are asked to select the AI method they believe is the most suitable to predict the quality of a stringer. This selection criteria is based on the meaningfulness and understanding of the explicit model and the accuracy of the predictions. After the method is chosen, experts are required to review and validate the rules automatically created with the aim of increasing the reliability of the machine learning rules.

Overall, the collaboration between AI algorithms and experts allows finding correlations in data, delivering a solution to complex problems by generating a set of rules that otherwise:

- Would need to be extracted from experts in high cost knowledge capture sessions.
- Would have missed correlations from large amounts of data that are often missed by experts.

5.6. Concluding remarks

In this chapter, a detail explanation of the achievement of the research objectives related to the knowledge sourcing framework implementation is provided. More precisely, each section of this chapter shows how the research target under study is accomplished. In addition, to provide the reader with a better understanding of the methodology proposed in this research, a simplified use case has been implemented within the knowledge sourcing framework.

Aiming at the validation of the proposed methodology, the realisation of two case studies instantiating the described framework is explained in more details in the following two chapters.

6. Knowledge Sourcing Framework: Case Studies

The aim of this chapter is to prove the foundations of this research through the realisation of two case studies. The first use case intends to carry out the design optimisation of wing covers made of carbon fibre reinforced plastics (CFRP). After this use case was performed a second case study was developed aiming at optimising the design of Lightning Strike Protection of Composites (LSPC).

In order to implement both use cases a Knowledge Sourcing platform was developed. The objective of this platform is the creation, capture and reuse of engineering knowledge by integrating expert knowledge together with knowledge provided by artificial intelligent methods. By achieving this objective, designers will make more informed decisions reducing the time required to select an optimal design.

From this perspective, the case studies show in practice: (i) the main research challenges associated with the process of sourcing engineering knowledge; (ii) how the reliable source of engineering knowledge can be efficiently achieved by enabling the collaboration between experts and automated machine learning algorithms. The domain of this use cases is the aerospace industry. More precisely, the specific context of the first use case is the lay-up process involved in the manufacture of CFRP wing covers whereas the context of the second case study is the lightning strike protection of composites.

This chapter starts with the description of the use cases background in section 6.1 followed by the description of the application of both case studies shown in section 6.2. After that, the results obtained and the procedure followed to validate the case studies is explained in section 6.3. Finally a discussion of the findings and concluding remarks of this chapter are presented in sections 6.5 and 6.6 respectively.

6.1. Background of the use cases

Although the general context of both case studies is the aerospace industry each of them has a specific background as described in the following two points.

6.1.1. Case study 1: Background

The general context of the first use case aiming at predicting the MFG cycle time of composite wing covers is the development of a new aircraft generation (Airbus A30X) that will substitute the current A320 model. Although the manufacturing the A30X model will start 20 years later, there is a need to investigate in advance technologies and capabilities that will be required to achieve the objectives of the A30X program.

More precisely, the Airbus research team involved in the development of this use case is focused on the systematic application of design for manufacturing (DFM) knowledge into working KBE applications predicting the performance of designs from the manufacturing perspective. In this context, engineers are facing two major challenges in order to develop efficient KBE systems:

- **Lack of information at early stages of the engineering design process.** To obtain an optimal design as quick as possible (decreasing the number of design iterations required to achieve it), it is required to make decisions based on relevant information. For instance, descriptions of the aircraft wing design are often given through analytical models. For example, some models describing the structural properties of a wing use Excel-based models rather than detailed CAD data.
- **Dynamic changes to the design requirements and design solutions**. Technologies that will be used to manufacture the A30X aircraft are still under development. Therefore, knowledge used in the KBE applications is constantly changing and even the technologies to be used are often modified. For instance, in the context of manufacture of CFRP parts

Automated Fibre Placement (AFP) is progressively replacing Automated Tape Layup (ATL) since AFP is able to manufacture more complex parts and reduce the amount of material scrap produced. This change of technologies provokes the need for updating, reviewing and validating the knowledge used by the KBE system efficiently.

The objective of the extended KBE system developed in this case study is to predict the manufacturing time of Carbon Fibre Reinforced Plastics (CFRP) wing parts. The CFRP parts employed in this use case are composed of a finite number of plies which are laid one on the top on the other using a roller. In this domain, design practitioners use 3D design data to obtain mass estimation of the parts. In parallel, 3D simulations are also used to make machine layup simulations of the design in order to predict manufacturing time. However, in conceptual design 3D data is generally unavailable. Additionally, existing machine time simulation tools lack automation and need a week of dedicated work from an engineer to estimate the MFG cycle time of a new design. The uncertainties and the need for fast evaluation of designs in conceptual studies make the use of 3D simulations unaffordable in this context. Therefore, the approach for this use case was to produce a fast and accurate solution to evaluate design concepts.

Another approach found in the literature supporting engineers in making more informed decisions –where there is a lack of knowledge– is the use of Artificial Intelligent (AI) methods [187]–[189]. Within the wide range of AI methods, Machine Learning (ML) methods have been widely used in the aerospace industry [190]–[193]. In this direction, [190] successfully used an Artificial Neural Networks (ANNs) approach to developed a decision support system for ultrasound inspection of fiber metal laminates obtaining classification efficiencies of ∼99% where the misclassification of defects was practically null. The criteria to analyse the performance of the method was based on the normalized Efficiencies Product (EP) and the confusion matrix obtaining an $EP = 0.9945 \pm 0.0065$ and misclassification accuracy of ∼0.08% by using Independent Component Analysis technique.

[191] also presented neural network approach in this case used for budget allocation of an aerospace company. The performance analysis was based on the Mean Squared Error (MSE) where the value obtained was practically null for training whereas for testing was of 1.6. Finally, [193] developed a Support Vector Machine (SVM) approach for the analysis of lead times of metallic components in the aerospace industry. The results of were analysed taking into account the Root Mean Squared Error (RMSE) obtaining a value of 2.29 (days) for training and 3.32 for testing.

The major limitation of these methods has been the lack of transparency and reliability. In fact, not being able to understand the rules to work out the predicted values, causes engineers to reject these advanced techniques. A potential solution to tackle the problems of current methods to predict design performance has been proposed in this work. The approach introduced in this research proposes the development of Knowledge Sourcing Framework (KSF) –considered in this study as an extended KBE system– enabling the collaboration between expert knowledge and machine learning algorithms while effectively managing the life cycle of the knowledge generated by experts and AI algorithms. The realisation of the KSF allowed the separation of the knowledge from its application (a machine learning algorithm in this case), storing it in a machine and human readable format. Therefore, the understanding of the knowledge created by machine learning algorithm is facilitated by the KSF.

6.1.1. Case study 2: Background

In current aviation industry, certification standards require evidence on the ability of an aircraft to withstand lightning strikes. Thus the change from metallic to composite materials on aircraft has initiated aerospace organisations to develop large clusters of expertise in the area. These developments have been carried out for programmes, both civil and military, and experts have identified the need to consolidate the learning made in this journey.

The validation of aerospace designs from Lightning Strike Protection of Composites (LSPC) perspective is mostly based on physical testing. A number of these tests can be virtual and based on previous instances. As result the amount of costly physical testing can be reduced. The ability to virtualise LSPC testing is a key input to this cost reduction. From this perspective, the case study developed in this research is to exploit the opportunity to use intelligent data and knowledge mining techniques and knowledge management methods to support the efficient and systematic source of engineering knowledge. In this direction, the main tasks carried out within the scope of this project were:

- Systematically capture and curate data and knowledge coming from both field experts and physical LSPC testing events and transfer it into a virtual knowledge base.
- Support field experts on ways to interpret the data, extract knowledge from it and transform it into a format that can be reused and marketed in future virtual versions of the physical testing.
- Execute virtual testing trials to obtain solutions that would have been obtained at the physical level and assess their accuracy.
- Transfer and exploit the knowledge gained in software tools and online service demonstrators.

In this context, this use case facilitates the understanding of the parameters driving the existing pressure in the nut cap when a sample design containing a Carbon Fibre Reinforced Plastic (CFRP) plate and a set of fastener assemblies is hit by a lightning strike. The nut cap is a piece made of plastic whose function is to isolate the fastener to avoid the spark spreading out.

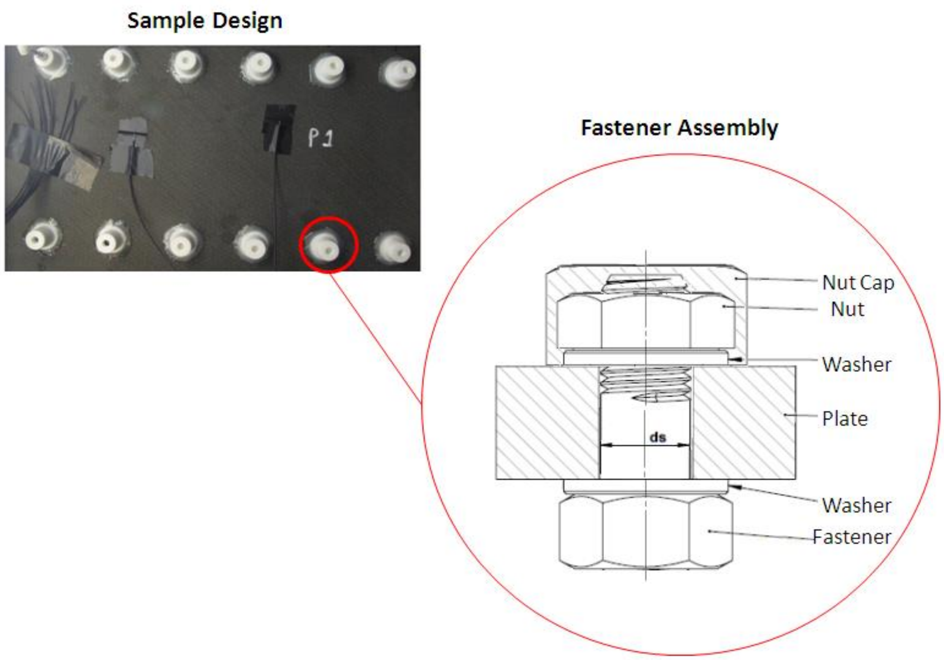


Figure 32.Sample design and fastener assembly.

The use of the proposed methodology will facilitate the analysis of the results obtained in the experiments and will allow the experts to determine what are the variables impacting the pressure in the nut cap and predict the nut cap pressure in new designs. The use of the ML algorithms embedded in the platform to estimate the nut cap pressure of new samples will help engineers to select those more likely to be the optimal design. Following the proposed knowledge sourcing approach, experts will be able to reduce the number of required experiments needed to understand a physical event.

In summary, to achieve an efficient sourcing of engineering knowledge, the solution to be developed must meet the following requirements:

- Simple enough to foster its use.
- Deliver fast and accurate predictions enhancing the optimisation of a wing cover design.
- Allow the analysis of the knowledge and results obtained by the user to increase its transparency and reliability.

6.2. Application of the case studies

To achieve the main goal of each of the use cases, knowledge sourcing methods and KBE applications must be properly integrated. The adequate integration of both elements is obtained using a well-established methodology (KNOMAD) described in section 4.2.1. In this regard, this section presents how the use of KNOMAD allows: (i) the management of the knowledge life cycle; (ii) the methodological collaboration between experts and ML techniques. To accomplish that, a KSF platform permitting the adoption of KNOMAD has been carried out (see section 4.3).

6.2.1. Process description and analysis

The aim of the case studies is to permit the user to efficiently perform the process of predicting the of a target or output class (MFG time in the first use case and nut cap pressure in the second case study) supporting the selection of an optimal design, saving time and costs. The approach proposed is embedded into a web-based platform, facilitating the access to different teams based in multiple locations. The process of predicting the MFG cycle time and nut cap pressure of CFRP parts is modelled in Figure 33 using IDEF0 (Icam DEFinition for Function Modeling, where 'ICAM' is an acronym for Integrated Computer Aided Manufacturing) representation. IDEF0 is a standard methodology for describing manufacturing processes using a functional approach [194]. The illustration below is focused on the overall process presenting the requirements to obtain the required predictions. These requirements are:

- **Input data.** This refers to design data corresponding to a set of design descriptors defined by experts. Moreover, in the first case there was need for the machine time output values provided by a simulation software application whereas in the second use case data generated in lab experiments was required.
- **Resources in the form of expert involvement.** Experts are required to initially define the parameters they believe are driving the output class. They are also needed to evaluate and validate the knowledge generated by the machine learning tool.

- **Set of specific applications**. This refers to the tools developed to enable the automated extraction of input data, machine learning process (including learning and prediction activities) and the advanced visualisation of the results provided by the knowledge sourcing capability.

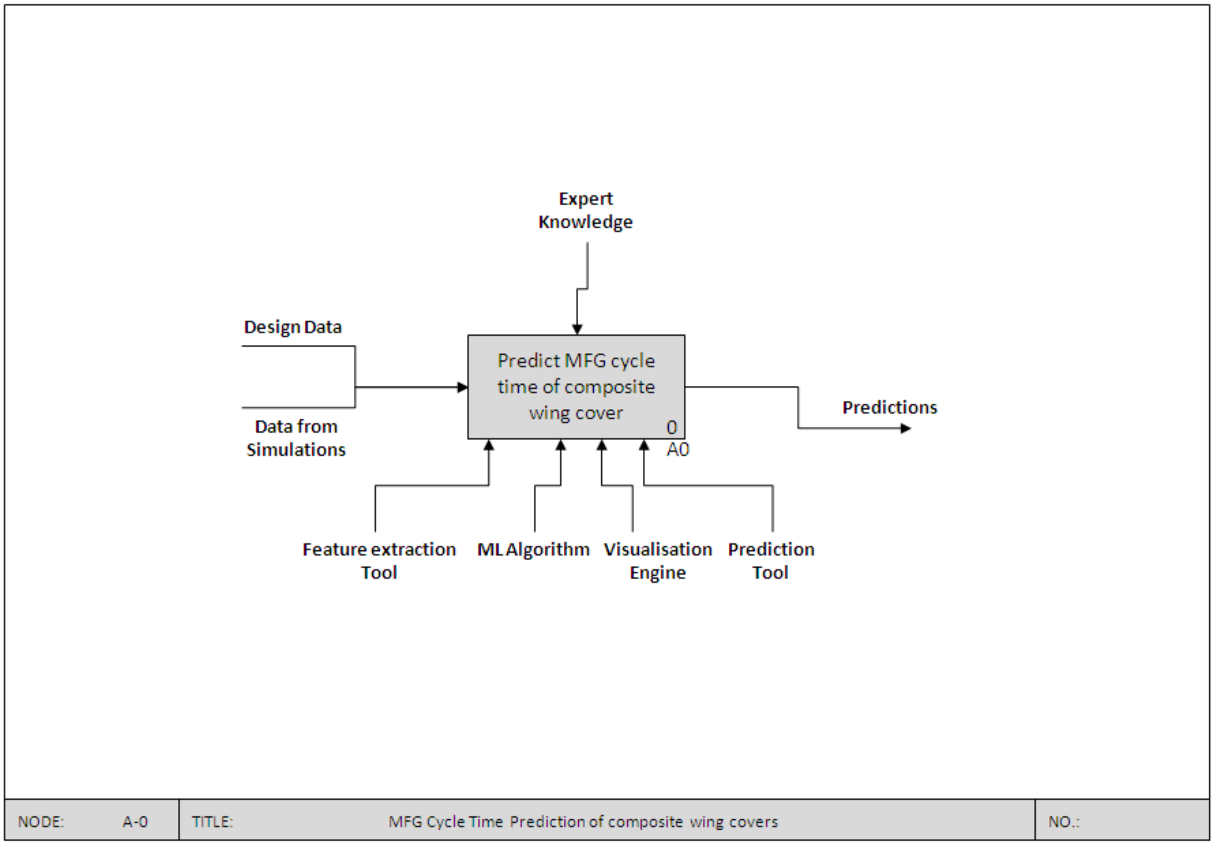


Figure 33.IDEF0: Main process.

The main process is divided in six sub-tasks as observed in Figure 34 and Figure 35. In the first place it is required to identify what are the design descriptors that drive the output class. To do that, experts must intervene using their experience and input data (simulation and design data or experimental data) to understand and identify which parameters –from the wing cover design or fastener assembly– are the ones affecting the output class values. This task was realised following a method described in APPENDIX D. As a result, a list of design descriptors driving the output class were defined and stored within the Content Management System (CMS). After this, a feature extraction tool generates a file named as “Training Set”. This document is obtained as a result of retrieving –from the design input file– the values corresponding to the design descriptors (input values) and output class values –from simulation log files in the first case and from files generated in the lab for the second case.

The “Training Set” file is employed by the machine learning algorithm to create a set of rules describing the system behaviour. These rules are later reviewed and modified by experts. The rule review process is supported by a set of visual analytical tools that permit data filtering and visualisation. Moreover, the format of the generated rules enables their automated execution when output class values (predictions) corresponding to a new design configuration are required.

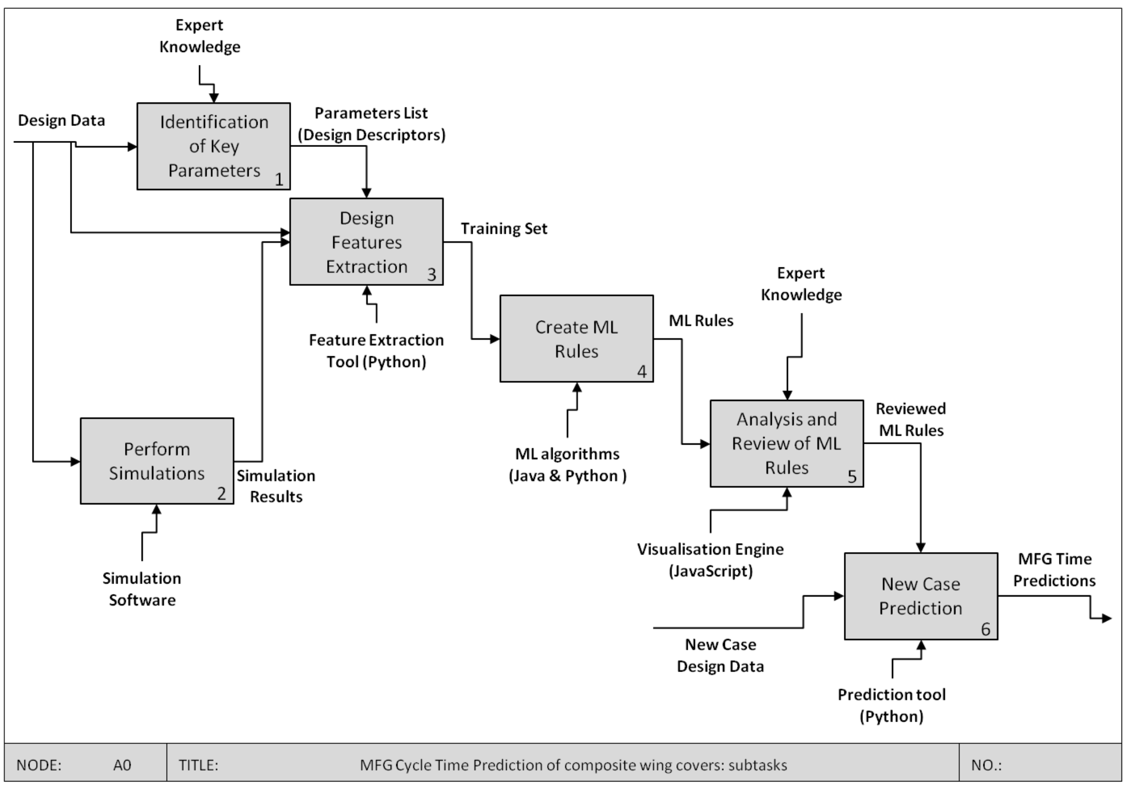


Figure 34.IDEF0: Subtasks for the first case study.

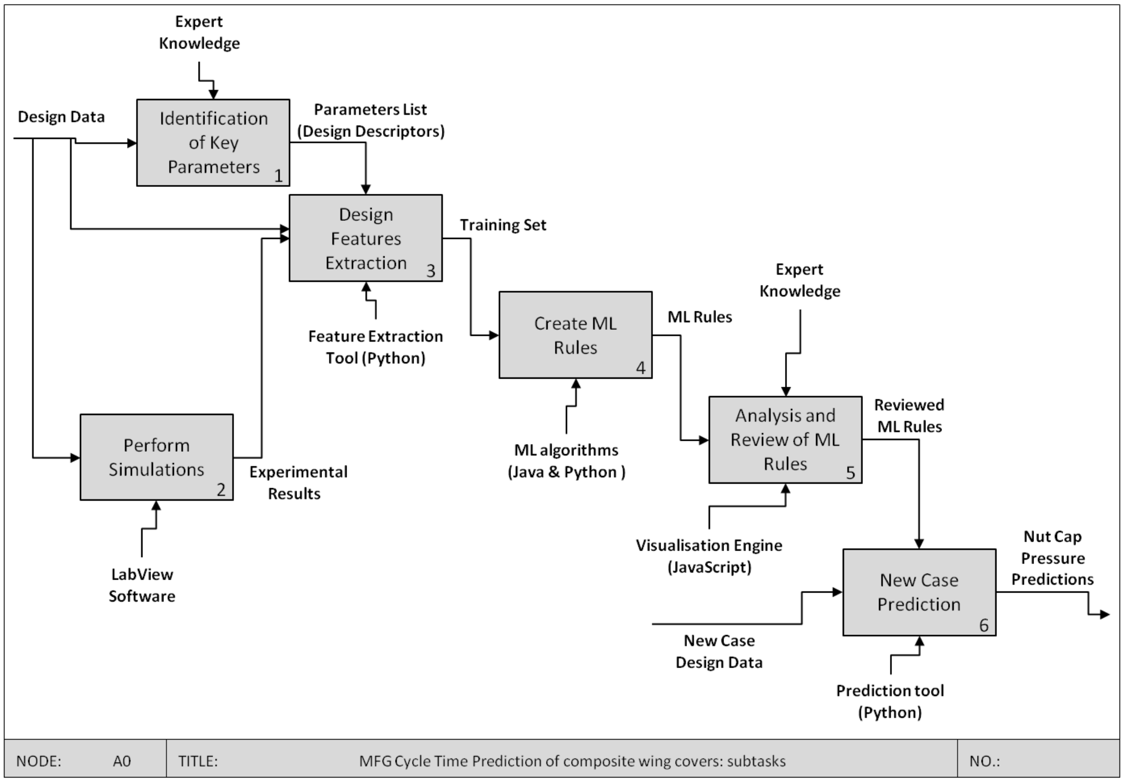


Figure 35.IDEF0: Subtasks for the second case study.

6.2.2. Adoption of KNOMAD methodology for the case studies

The need for a generic methodology in charge of performing the systematic management of engineering knowledge has been established as one of the main challenges of this research. Moreover, due to the disruptive nature of the solution proposed, methodological support is required in order to effectively integrate the developed methods and tools within a common framework.

From this perspective, KNOMAD has been adopted in this study aiming at providing the necessary methodological support. This methodology is based on six main phases: Knowledge Capture (K), Normalisation (N), Organisation (O), Modelling (M) and Implementation, Analysis (A) and Delivery (D) as presented in Figure 36.

In the knowledge capture stage, the scope, objectives and assumptions of the project are identified and knowledge is extracted from explicit and tacit sources. Once the knowledge is captured, the Normalisation tasks are initiated. In this phase, the knowledge is transformed to comply with the organisation's needs. This is followed by the definition of the data structure in the organisation step, which enhances the accessibility of the knowledge stored. At this stage, the use of ontology is advised in order to increase its transparency and applicability.

The next step is modelling where Enterprise Knowledge Resources (EKR's) are utilised to independently store applications and their corresponding knowledge, improving the management of the knowledge life cycle. Knowledge and tools are placed in the framework within three main EKR models: "Knowledge", "Applications" and "Case Reports". Models linked to the Content Management System (CMS) are also required at this stage to enable the interaction between the user and the knowledge sourcing platform. The creation of these CMS models enable web-based activities within the CMS such as automated creation of case reports, creation of tailored visualisations, and files uploading. Once EKR and CMS models are created, the KSF applications are integrated into a semi-automated process.

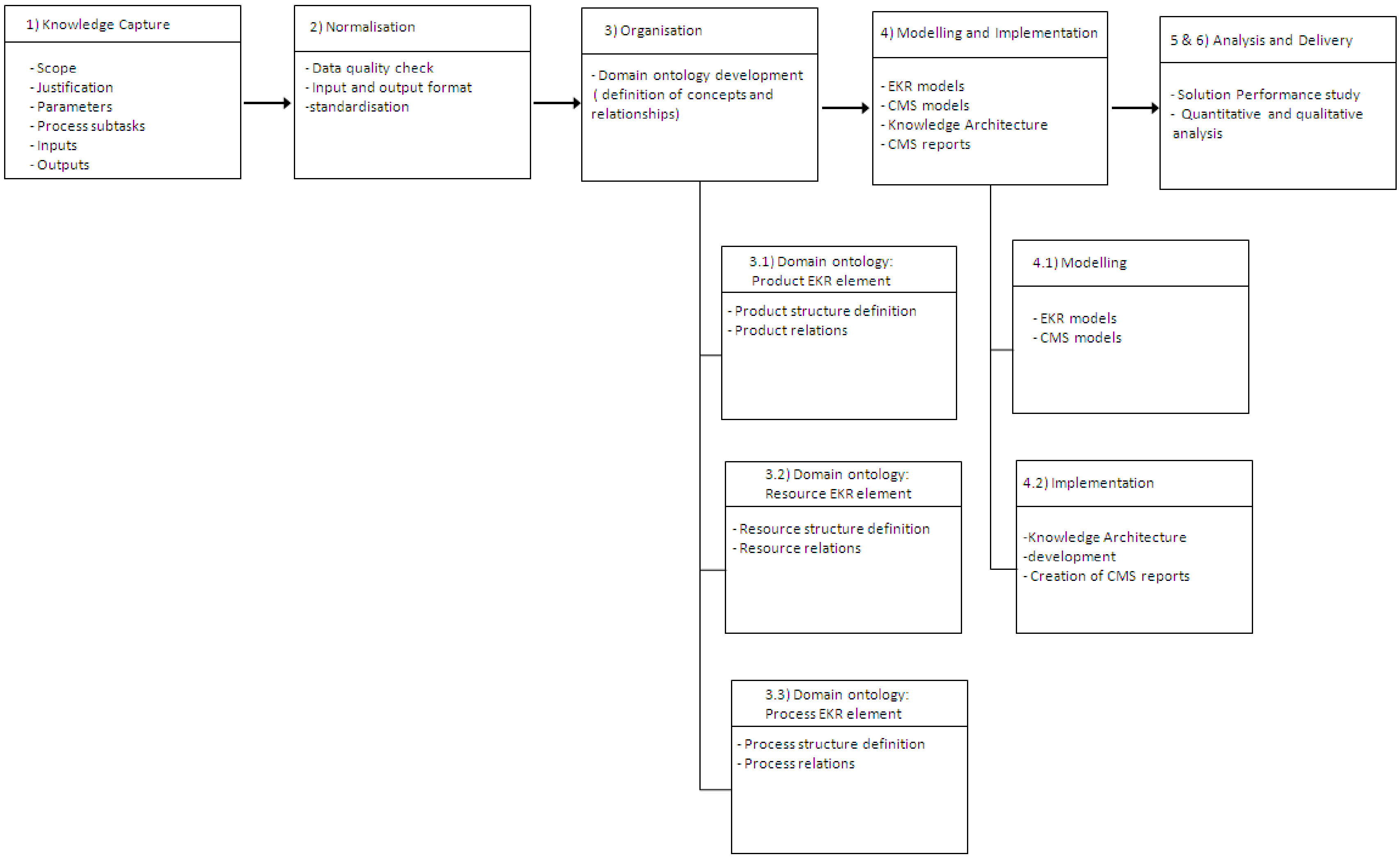


Figure 36.KNOMAD: process flow

The final step entitled as “Analysis and Delivery” is a combination of the two last phases of the KNOMAD methodology. This stage encompasses quantitative and qualitative analyses aiming at the delivery of a robust and reliable capability.

6.2.2.1. Knowledge capture

The knowledge capture or acquisition process is acknowledged by the research community as an essential step to develop an efficient KBE system [26]. This is also supported by the fact that the sourcing of knowledge by machine learning algorithms is highly dependent on the quality of the data captured. Therefore, this process must be adequately realised to ensure the success of the prediction process performed by the ML method.

The knowledge used to obtain the output class predictions is sourced by experts and ML algorithms. Expert knowledge was initially extracted through interviews which enabled the creation of a list of design descriptors. Expert knowledge was also provided in the evaluation of the explicit model generated by the AI algorithm.

The acquisition of expert knowledge started with the interview of domain experts. The interviews aimed at identifying the design descriptors driving the output class. The process followed in the interview activity is explained in more detail in APPENDIX D. Although both use cases followed the same knowledge capture approach there are some details which differ. The details of each use case in terms of the knowledge capture are described in the following two points:

A. Use case 1: Specific details of the knowledge capture process

The interview process was carried by interviewing two experts with more than 15 years of experience on KBE systems where a set of design descriptors driving the MFG cycle times were identified.

To characterise the design of a composite wing cover, each layer of the composite cover was accounted as an independent sample. In doing so, it was possible to obtain a generic set of design features describing a wide variety of design configurations (Figure 37).

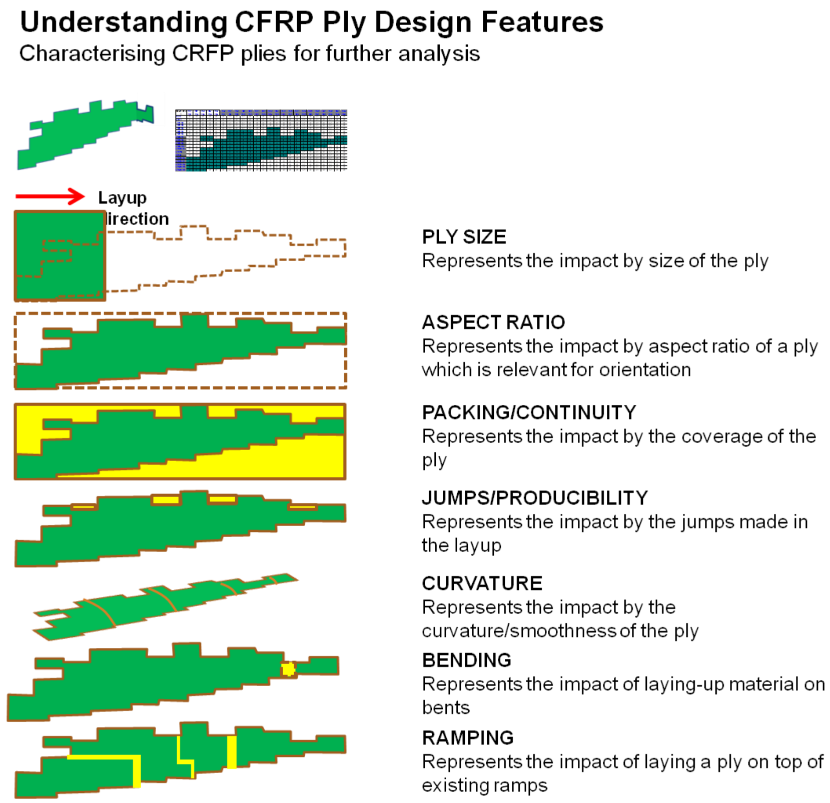


Figure 37. List of Design Descriptors

B. Use case 2: Specific details of the knowledge capture process

The capture of expert knowledge started with the interview of three experts with more than 10 years of experience on the electrostatics domain. The realisation of the interviews allowed the characterisation of the design and the identification of relevant design descriptors.

The designs or samples analysed in this case study are composed of two main parts: CFRP plate and fastener assembly. The latter is composed of a nut cap, two washers, a nut and a bolt. The data is captured by a set of sensors, placed on the elements of the design, when a particular fastener is shot by a lightning strike.

The data captured is divided in two groups: direct parameters and inferred parameters. Direct parameters are extracted from experimental data without the requirement of performing a data pre-process activity. These parameters are commonly used in the LSPC domain such as voltage, current, and pressure. In contrast, inferred parameters are derived from direct parameters by experts in the domain as a result of a detailed analysis of the problem context. Examples of these two groups of data are listed in the table below.

Table 19. Type of parameters captured

Direct parameter examples	**Inferred parameter examples**
Peak voltage on the bolt	**Design complexity:** It refers to the level of complexity of the design configuration. It can be subject to expert opinion.
Resistance between nut and plate	**Derived current:** Parameter that is obtained by applying an ad-hoc equation defined by the expert that captures the current in a fastener due to a shot in a different fastener.

Traditionally, experts are asked to provide the equations modelling a specific problem (e.g. estimation of MFG time of wing covers). By realising the approach proposed in this work, less experts' time is required as it is only necessary to identify the parameters which experts believe are driving the output class.

Once the design descriptors are defined, an application was developed to enable the extraction (from the designs) of the information required to create the "Training Set". The "Training Set" is the input file used by the machine learning algorithm to create a set of rules which emulate the problem behaviour. This input file contains data corresponding to the design descriptors of each sample and their respective output class values obtained from simulations in the case of the first use case or from experimental data in the case of the second use case. The ML algorithm searches for data correlations using the “Training Set” producing as a result a set of rules or explicit model (Figure 38). Finally, these rules are stored within the knowledge repository and go through an expert review and validation process before they can be used to predict the values of the defined output class of new design configurations Figure 39.

M5 Rules Algorithm

“IF THEN” rules exploited to predict MFG cycle time

```
Rule: 1
IF
    Packaging <= 0.441
    Aspect_Ratio <= 2.414
THEN

MachTimeArea =
    16.6668 * Reference_square_alignment=diagonal aligned plies
    + 0.3611 * Reference_square_ply_area
    + 6.7422 * Layup_alignment=across tape direction,diagonal tape direction
    - 0.3938 * Aspect_Ratio
    - 105.9206 * Packaging
    - 20.1005 * Jumps
    - 0.0372 * Orientation
    - 0.0433 * Perimeter
    + 174.6305 [66/9.035%]

Rule: 2
IF
    Packaging <= 0.441
    Jumps > 0.761
THEN

MachTimeArea =
    -5.2342 * Reference_square_alignment=diagonal aligned plies
    - 0.4545 * Reference_square_ply_area
    + 43.6232 * Layup_alignment=across tape direction,diagonal tape direction
    - 7.3034 * Aspect_Ratio
    - 223.4472 * Packaging
    - 30.7608 * Jumps
    - 0.0844 * Perimeter
    + 262.858 [35/27.056%]

Rule: 3
IF
    Jumps > 0.761
THEN

MachTimeArea =
    -3.5922 * Reference_square_alignment=diagonal aligned plies
    - 0.468 * Reference_square_ply_area
    + 17.8775 * Layup_alignment=across tape direction,diagonal tape direction
    + 0.5488 * Aspect_Ratio
    - 71.5921 * Jumps
    + 0.2557 * Perimeter
    + 137.5859 [68/3.977%]

Rule: 4

MachTimeArea =
    -4.2368 * Aspect_Ratio
    - 570.8271 * Packaging
    + 357.8701 [28/54.393%]
```

Figure 38. Machine Learning Rules obtained in the first use case

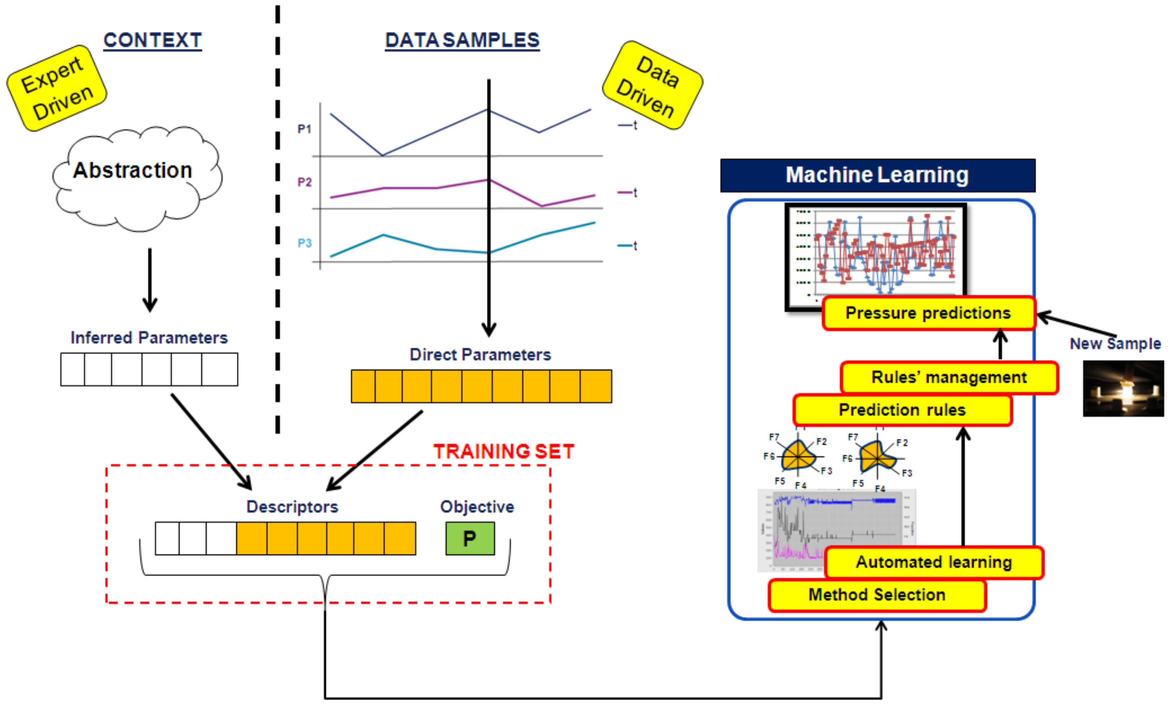
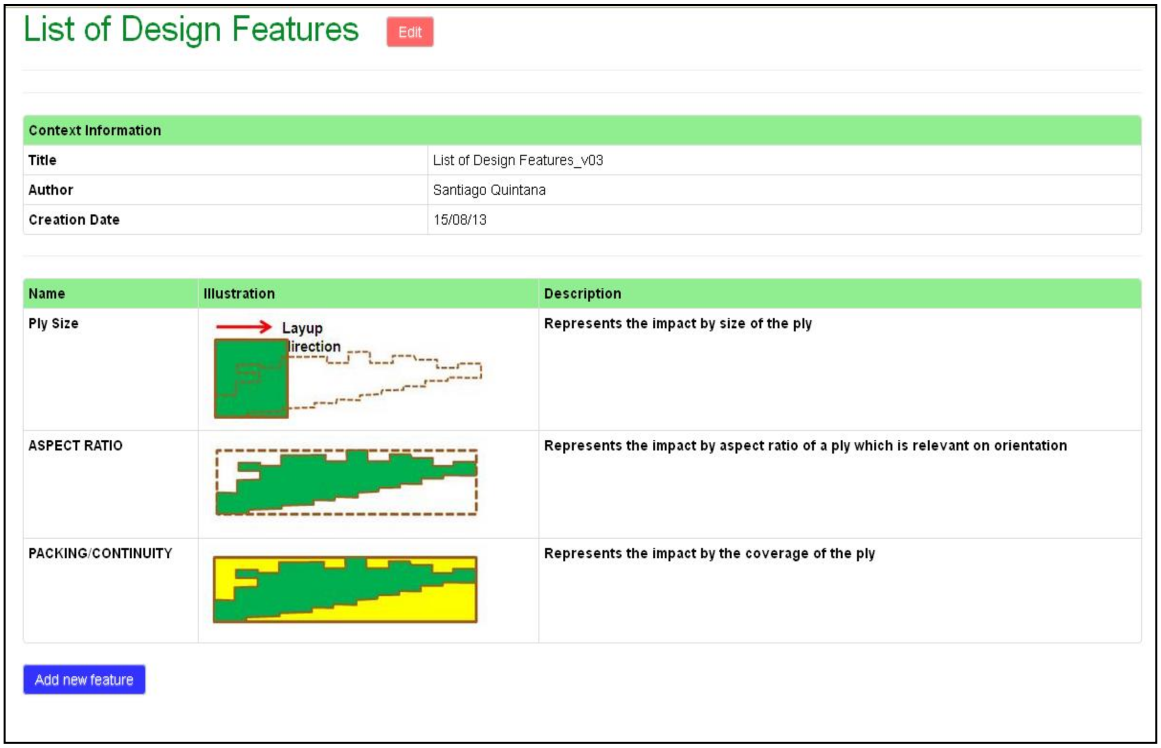


Figure 39. Prediction task: Process flow

6.2.2.2. Normalisation

At this stage, knowledge previously captured is transformed to comply with the standards established by the stakeholders. This step enhances the data retrieval and facilitates the implementation of the ontology described in section 6.2.2.3.

Normalising knowledge enables the storage of the knowledge extracted within formal and informal models, thus improving the knowledge accessibility and maintainability. In this study, informal models contain context data such as article owner, creation date and general information; all of them presented in a human readable language (Figure 40). In parallel, formal models contain machine code in the format of “IF (condition) THEN” rules (Figure 41) which are encoded, thus enabling their automated execution.



List of Design Features

Edit

Context Information	
Title	List of Design Features_v03
Author	Santiago Quintana
Creation Date	15/08/13

Name	Illustration	Description
Ply Size		Represents the impact by size of the ply
ASPECT RATIO		Represents the impact by aspect ratio of a ply which is relevant on orientation
PACKING/CONTINUITY		Represents the impact by the coverage of the ply

Add new feature

Figure 40. Informal Model: List of Design Features for the first use case.

ML rules

Edit

Context Information	
Title	ML Rules_v07
Author	Santiago Quintana
Creation Date	23/09/13

Rule Id	Equation	Action
Rule 1	(see below)	Remove
Rule 2	(see below)	Remove

Rule 1:

```
if Packaging <= 0.441 and AspectRatio <= 2.414:
        MachineTime_v02 = 0.205 * TotalArea - 0.355 * AspectRatio - 84.5633 * Packaging - 19.5469 * Jumps  + 168.3654
        if Layupalignment in ('diagonal tape direction', 'across tape direction'):
                MachineTime_v02 += 6.5101
        if ratio in ('diagonal aligned plies'):
                MachineTime_v02 += 17.1405
```

Rule 2:

```
if Packaging <= 0.441 and Jumps > 0.745 and TotalArea > 6.125:
        MachineTime_v02 = - 0.7059 * TotalArea - 5.6553 * AspectRatio - 97.8651 * Packaging - 111.2078 * Jumps + 275.8134
        if ratio in ('diagonal aligned plies'):
                MachineTime_v02 += -1.3767
        if Layupalignment in ('diagonal tape direction', 'across tape direction'):
                MachineTime_v02 += 36.2812
```

Add new rule

Rules' History

Figure 41. Formal Model: ML Rules for the first use case.

6.2.2.3. Organisation

The main task realised in the organisation phase is the data structure definition. An ontology defines a common vocabulary for engineers. Its use makes knowledge to be understandable by automated search applications, thus enhancing the accessibility and traceability of the knowledge stored while facilitating the knowledge update. KNOMAD fosters the use of an ontology to properly achieve the knowledge architecture. In this regard, a domain specific ontology is constructed in accordance with the main Knowledge Life Cycle (KLC) ontology described in 4.2.2 which specifies the domain concepts and their relationships. The design of a domain specific ontology defines the class hierarchies, relationships and, their attributes and behaviours. Three main classes were used to annotate domain knowledge. These classes are named as “Product”, “Process” and “Resource”.

A. First use case: ontology definition.

- **“Product” class**: its scope is constrained to CFRP wing covers. This class is focused on the layers or plies that constitute a CFRP cover as shown in Figure 42.
- **“Ply” class**: it contains several attributes and behaviours which are related to the descriptors of a ply design. These parameters were previously identified by experts in the knowledge capture process. In this work, behaviours are considered as functions or methods enabling the interaction between instances of a class. For example, the behaviour named as “get_aspect_ratio()” is executed within the class “ply” using as input values the attributes named as “length” and “width”, generating as a result the ply aspect ratio value.

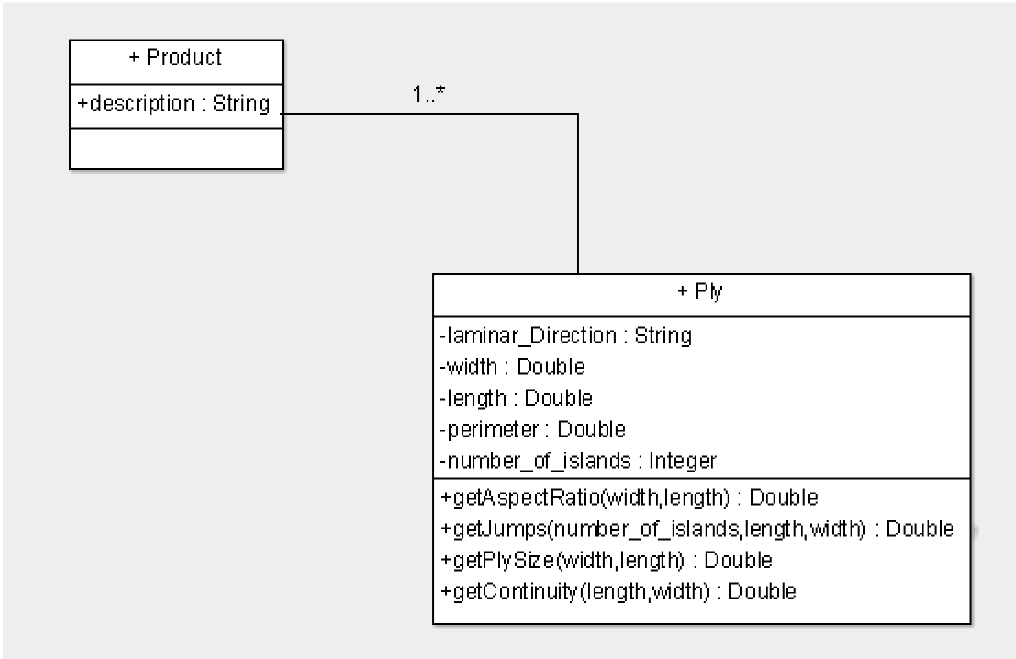


Figure 42. Specific Domain Ontology: Product Class for the first use case

- "**Process" class**: it encompasses a class describing a process objective and a set of sub-classes representing the processes corresponding to the applications integrated within the KSF platform (Figure 43).
- **"Resource" class**: it groups a set of subclasses entitled as "Expert Resource", "Application Resource" and "Document Resource" (Figure 44). The "Expert Resource" class contains knowledge acquired in the knowledge capture process, essential to create the ML rules such as the parameters that drive the MFG cycle time. The "Application" class, includes tools such as the feature recognition application or the implementation in charge of generating ML rules automatically. The "Document Resource" class contains the case reports that are automatically generated when the applications are executed.

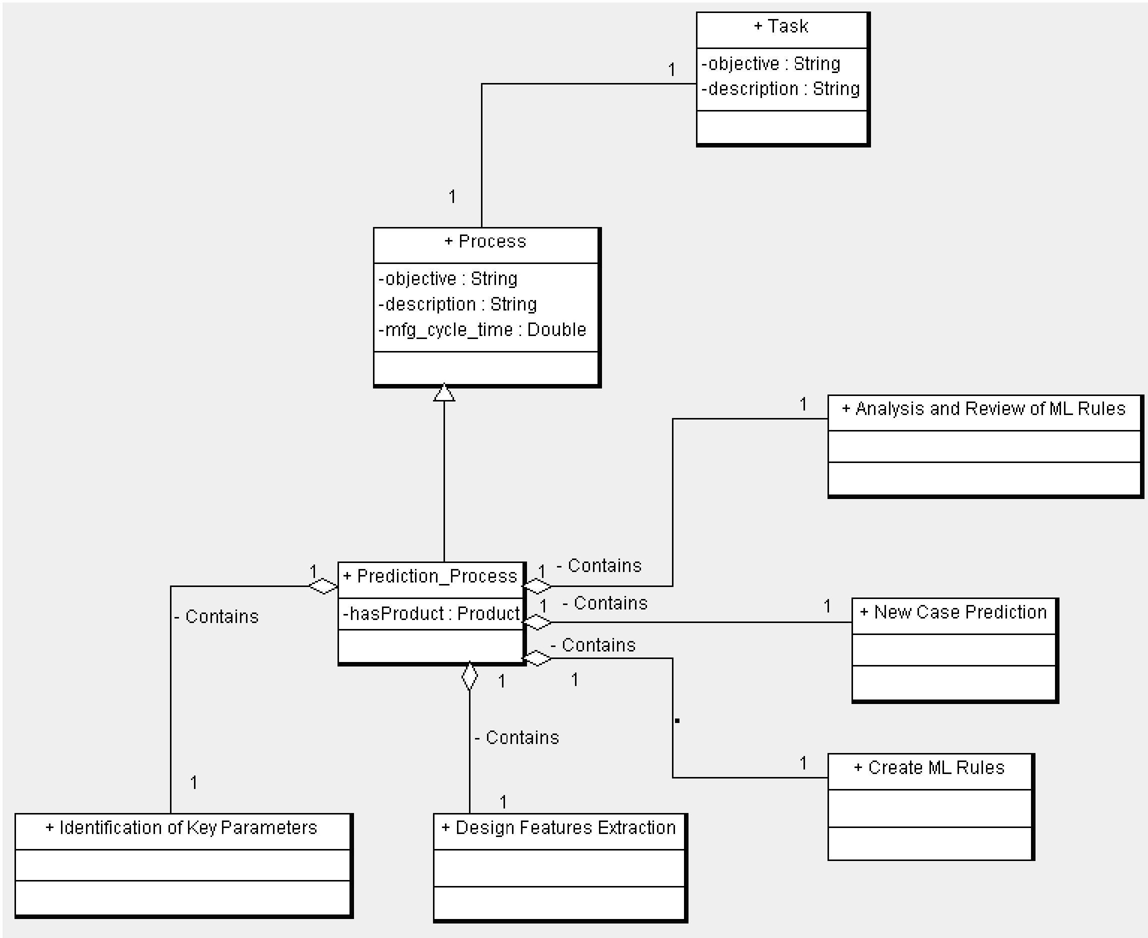


Figure 43. Specific Domain Ontology: Process Class

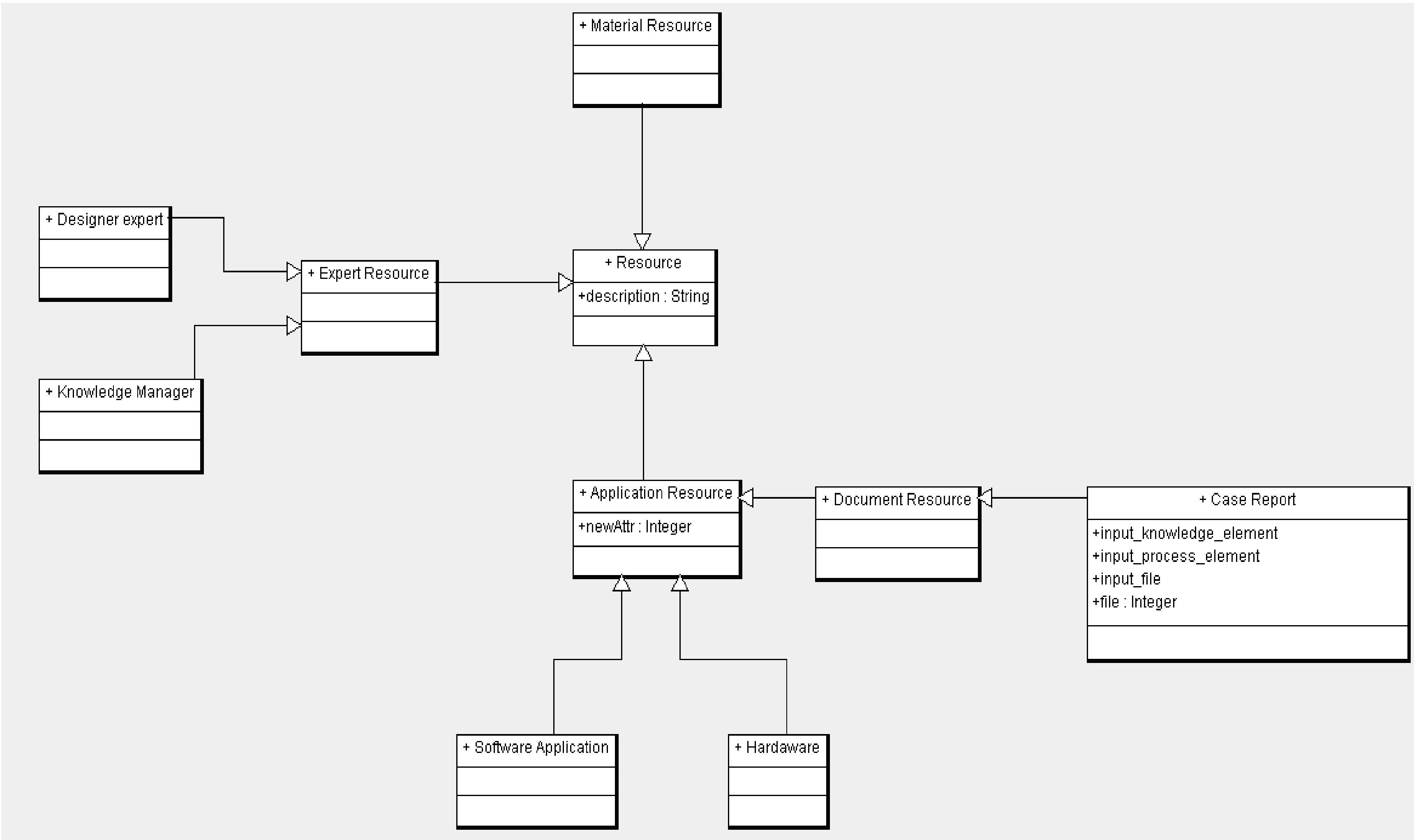


Figure 44. Specific Domain Ontology: Resource Class

B. Second use case: ontology definition.

- **Product class**: it only takes into consideration samples composed of a CFRP panel containing a set of fastener assemblies Figure 45.The “Shot” object is a subclass of the “Product” class and the data contained by this object is captured when a fastener assembly is shot by a lightning strike.

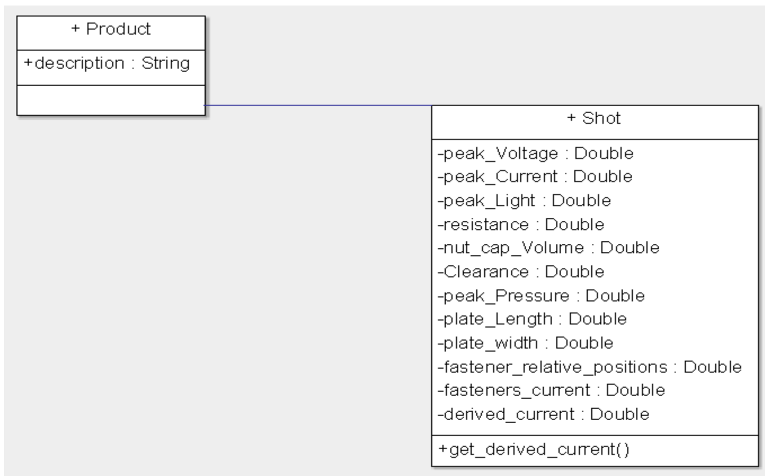


Figure 45. Specific Domain Ontology: Product Class for the second use case

The “Shot” sub-class contains several attributes and behaviours which are related to the main groups of descriptors previously mentioned: contextual and direct parameters. In the context of this work, behaviours are considered as the methods that allow the interaction between instances of a class. For instance, a behaviour named as “get_derived_current()” is executed within the class “Shot”. The output of this process is the data corresponding to the “derived current” parameter. To provide the “derived current” values, the “get_derived_current()” behaviour makes the required calculations using input data coming from the attributes named as “plate_length”, “plate_width” and “fastener_relative_positions” (Equation 3).

Equation 3. Derived Current

Derived Current *= get_derived_current (plate_length, plate_width, fastener_relative_positions)*

- "**Process" class:** it is the second main class within the defined ontology and it contains a set of sub-classes representing the capabilities created with the aim of predicting the pressure in the nut cap (Figure 43).
- **"Resource" class:** it encompasses a set of subclasses named as: "Expert Resource", "Application Resource" and "Document resource" (Figure 44).

6.2.2.4. Modelling and Implementation

This stage of the KNOMAD methodology encompasses the modelling and implementation of the EKR's.

<u>Modelling</u>

The EKR's created under the scope of this study have been annotated and modelled using the specific domain ontology specified in section 6.2.2.3. More precisely, the modelling step is focused on:

- Enabling knowledge elements to be independently stored from its application, thus fostering the use of this knowledge across different engineering problems.
- Providing users with a simple system which facilitates its usability and maintenance.
- Permitting knowledge to be easily updated by allowing the review and validation of ML rules.

The EKR classes modelled are:

- **EKR Knowledge**. This class contains knowledge elements that can be used in multiple processes or problems. In this case study, the knowledge elements are the design descriptors identified by experts in the knowledge capture stage. Within this class, reports or articles utilised to create the list of design descriptors are also included.
- **EKR Process**. In this class, three models are used:

- ➢ Input data model. This allows user to upload data into the content management system.
- ➢ Execution model. This retrieves data from the CMS and executes a script that triggers ML algorithms belonging to WEKA and scikit-learn, thus generating the required predictions.
- ➢ Data Analysis Model. This enables the visualisation and review of the ML rules that were used in the prediction process.

- **EKR Case**. This class contains knowledge elements stored as reports which are automatically created when the process is executed.

<u>Implementation</u>

The implementation process involved the creation of a knowledge-based platform that is built on top of a content management repository. The CMS, containing the knowledge, applications and case reports elements, enables the implementation of the KNOMAD methodology and the consequent development of the KBE system. In fact, the CMS permits the data capture from experts and ML algorithms, the knowledge execution by the KBE applications, and the systematic storage of data. All these activities are possible thanks to the integration of a set of offline and online tasks within a common environment as described in section 4.3. The elements belonging to the "Process" EKR class are integrated within the system architecture as shown in Figure 46.

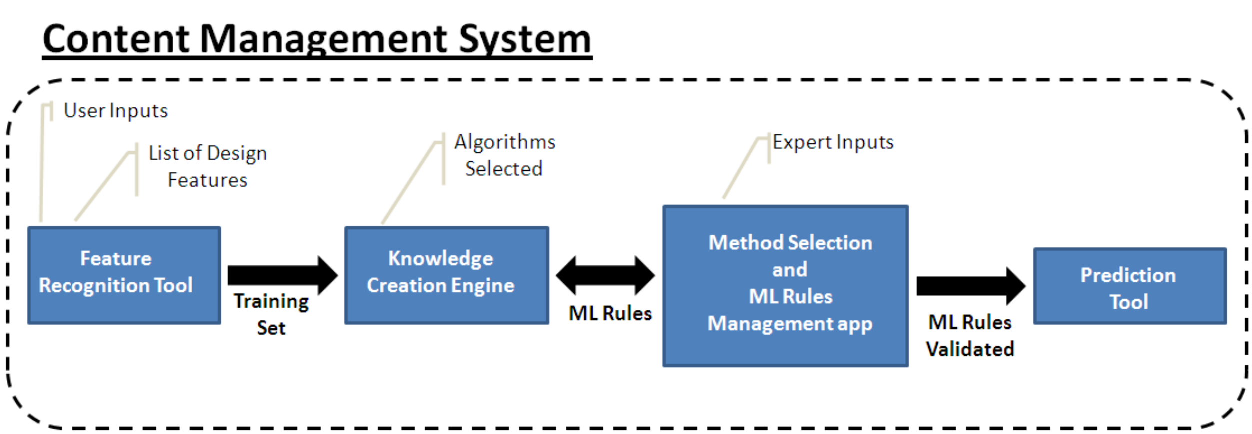


Figure 46. KBE system architecture

To better understand the KBE system implementation, the elements present in the KBE system architecture are described below. The first three boxes of the illustration shown above are applications corresponding to the learning process. These three tools are essential to create a set of ML rules which will be used by the fourth element to accurately predict the output class values of a new design. A detailed description of each application is presented as follows:

A. Feature recognition tool

The aim of this application is to provide the machine learning algorithm with a good quality data set. To do that, once the user uploads a new design within the system, a python script analyses the design input values and extracts the values correspondent to the design descriptors. The script uses a list of design features (previously identified by experts) to extract the values corresponding to the design descriptors. These design descriptor values (input values) are merged with their respective output class values that are extracted from a simulation log file for the first use case and from experimental data for the second use case. After the script is executed, a Comma Separated Value (CSV) file named as “Training Set” containing the values of the design descriptors and machine time for each sample of the design is generated.

The same method is used by the feature recognition tool to generate the “Test Set” file used in the validation process. The specific feature recognition tools built for each use case are described in the following two points in more details.

A1. First use case: Feature recognition tool

In this case study the design data was stored in an excel file containing the 2D shape of each ply of a specific CFRP wing cover. To extract the data correspondent to the defined design descriptors a new algorithm was developed and written in Python. This algorithm is complex and it is in charge of analysing the each of the 2D shapes contained in the design excel file and retrieving information such as the length, width, discontinuity in the material (by identifying the areas within the shape not containing material), packing (considering the accumulation of material or how the material was packed in the direction the material was being laid) and number of ramps (areas where the material was inclined due to the difference on number of layers under a specific area) among others (Figure 37). The calculation of the ramps was considerably difficult due to need for taking into account the data of previous layers to calculate the ramp area of a specific wing cover ply Figure 47.

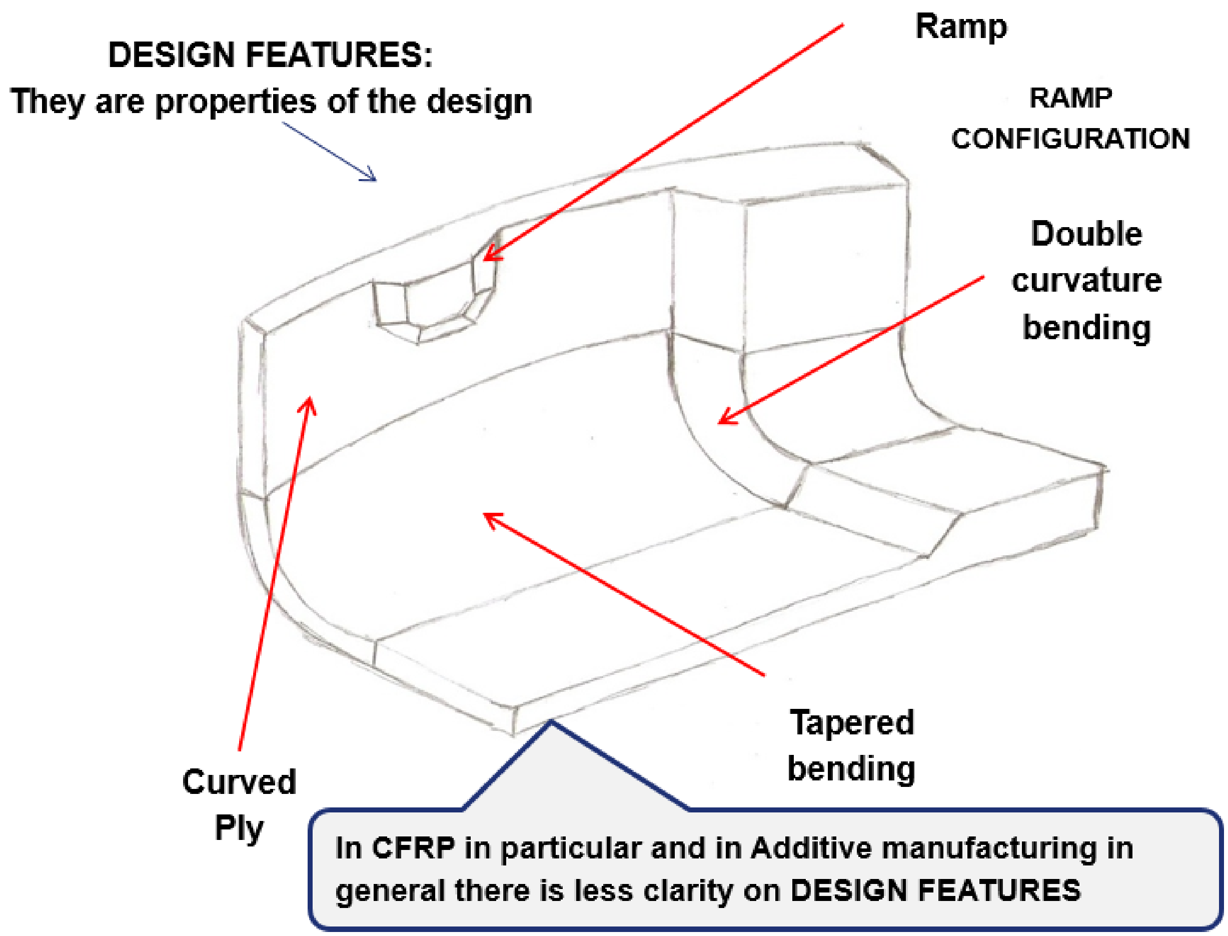


Figure 47. Design features of CFRP wing covers

Once the design data was extracted there was a need of gathering data correspondent to the output class which was stored within simulation log files. In this case, the extraction of the simulated data was a straight forward process. Finally the design and manufacturing times (output class values) were integrated within a single file named as "Training Set" (Figure 48).

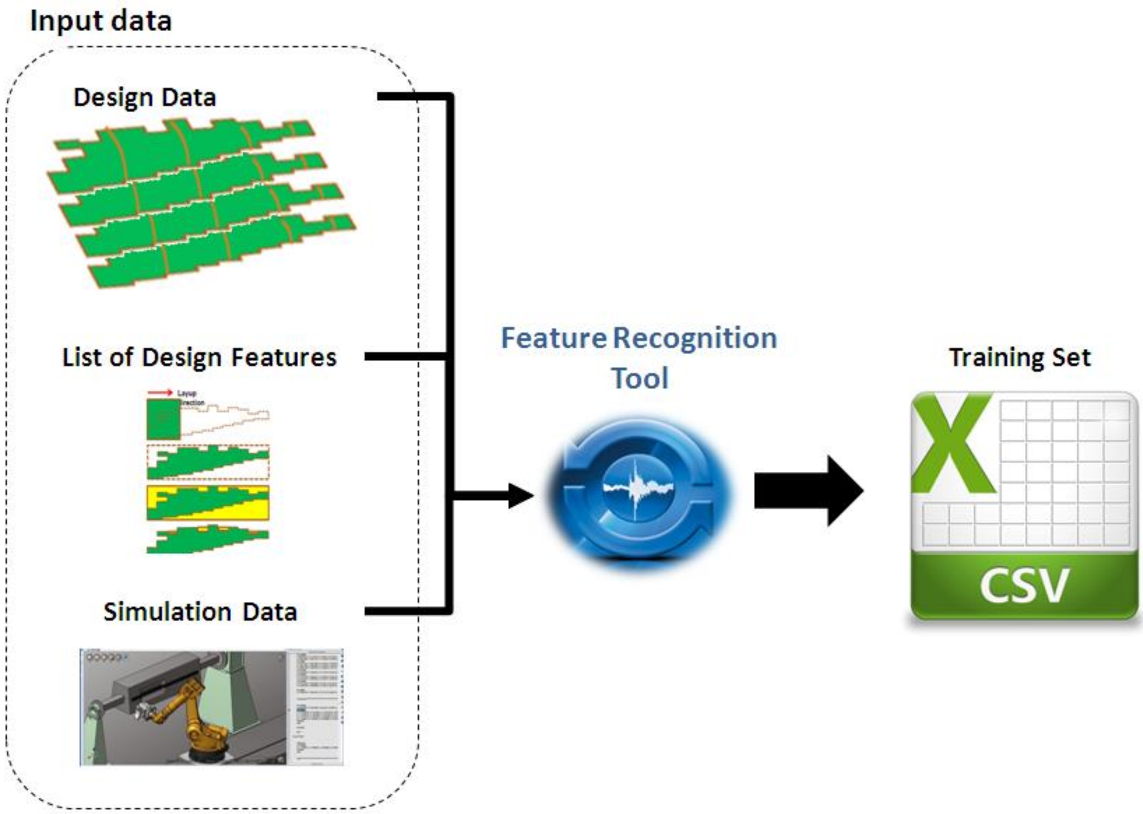


Figure 48. Feature Recognition Tool.

A2. Second use case: Feature recognition tool

In the second use case the feature recognition tool consists of a python script which analyses the design and extracts the values corresponding to the design descriptors. This script uses a list of design features ("Direct" and "Inferred" parameters) previously identified by experts, to create a Comma Separated Value (CSV) file named as "Training Set". This file contains the values of the design descriptors and nut cap pressure for each sample of the design (**Error! Reference source not found.**).

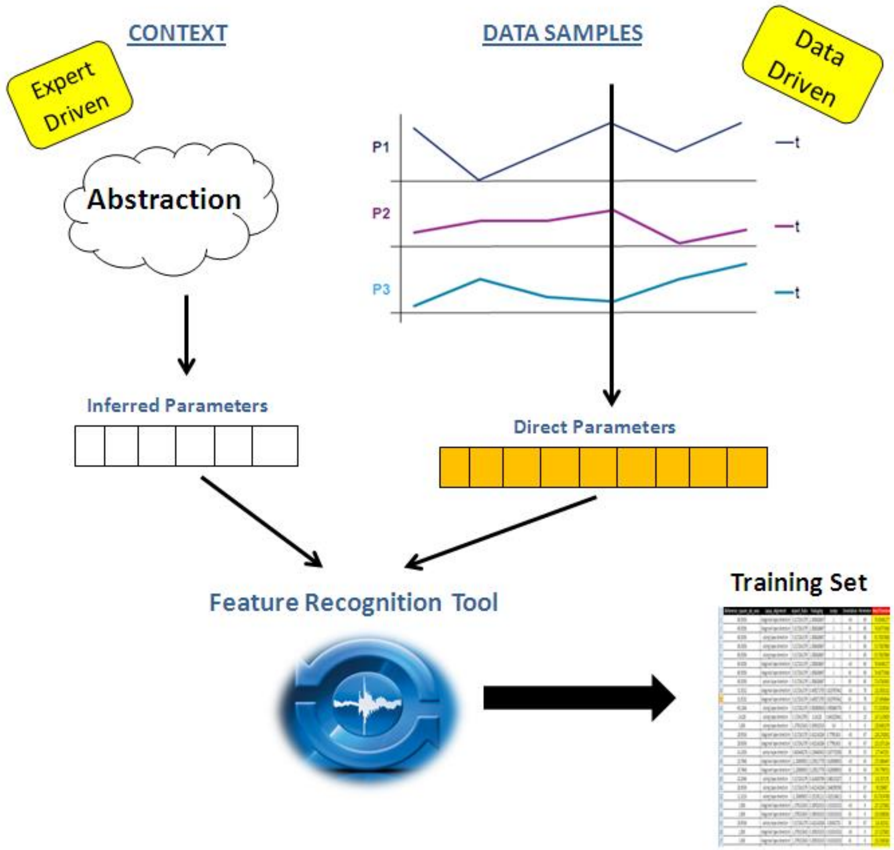


Figure 49. Feature Recognition Tool.

B. Knowledge creation engine.

This tool is in charge of running the learning process where a set of algorithms are executed, creating their corresponding rules modelling the problem. The machine learning algorithm looks for data correlations, creating the minimum number of rules possible (high generality) while keeping a high accuracy rate (Figure 50). This is achieved in many ways (depending on the algorithm applied), for instance when using a decision tree method this is achieved by creating initially a tree with a high number of branches and leaves which are later on pruned until the accuracy does not decrease anymore. This pruning process enables the model to be reduced in size (increasing its generality) while increasing its accuracy.

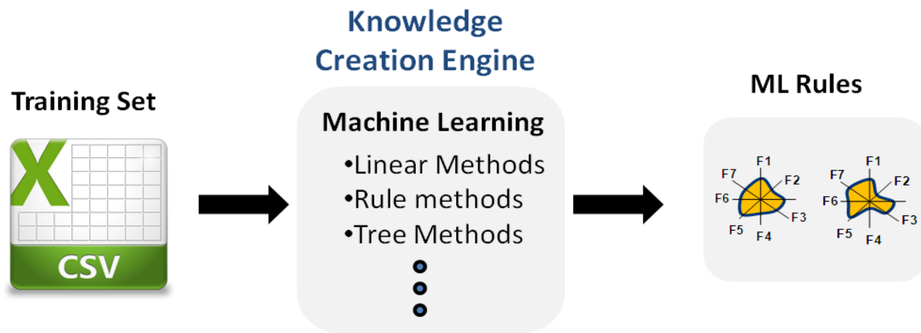


Figure 50. Knowledge Creation Engine.

C. Method selection and ML rules management application.

After creation of the ML rules, experts proceed with the analysis of the rules automatically generated. This evaluation process finalises with the selection of a ML technique which will be further analysed in order to validate the explicit model created. For this purpose, the Rule Management Application (RMA) focuses on the analysis of an explicit model through the utilisation of visual analytical tools. The main objective of RMA is to foster users to review, update and validate the ML model using the CMS. The use of a CMS, that enables the systematic review and validation of the model of the problem, facilitates the knowledge retention and reuse. Moreover, the use of charts and tables permitting the traceability of the results obtain increases the reliability of the ML rules (Figure 51).

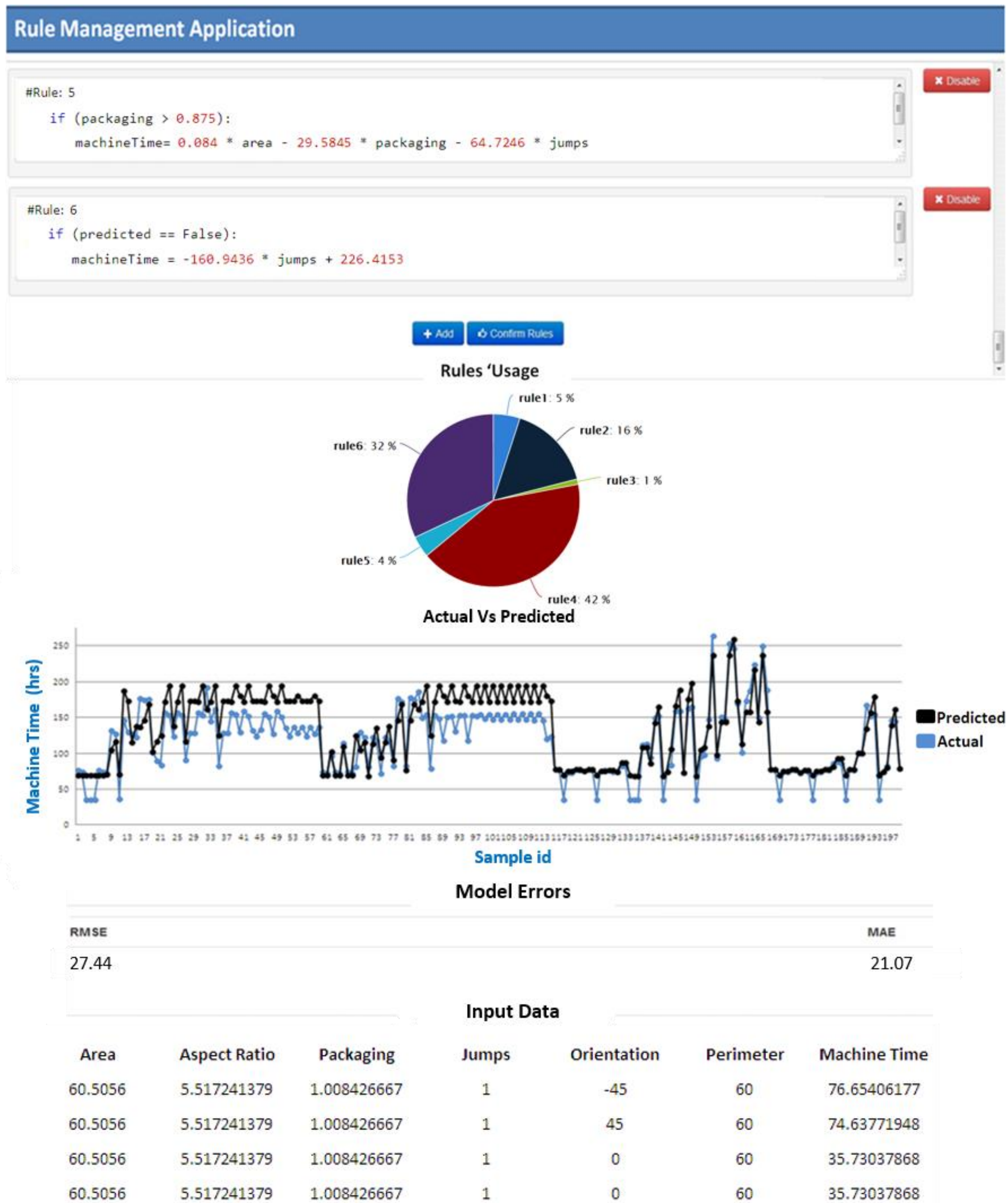


Figure 51. Rules Management application: first use case example.

The display of the model together with its corresponding visualisations facilitates the understanding of the rules. In addition, RMA fosters the improvement of the model by enabling the modification of the rules that is supported by the visual analytical tools. This visual applications allow users to quantify the impact of the changes made in the rules through the analysis of the ML scoring parameters and the charts.

D. Prediction tool

This application encompasses the validation of the model using the “Test set” and the prediction of new cases using the validated rules. After pre-validation of the model using the RMA, the user uploads the “Test set” containing design descriptor and machine time values. Once the file is uploaded, the prediction tool executes a JavaScript function which triggers a python script in charge of retrieving the pre-validated rules stored within the content management system. Then another python script uses the rules retrieved to evaluate the samples contained in the “Test set” and generate the predictions. Finally, this script sends the predictions to the browser where a JavaScript function uses some charting libraries to display the predictions in an intuitive manner.

Based on the accuracy values the experts will or won’t validate the model. After the model is validated, the prediction tool is used to predict the machine times of new design cases. To achieve the predictions of a new design, where the machine time is unknown, a python script retrieves the validated ML rules (stored in the CMS) and executes them using as inputs only the design descriptors of the data belonging to the uploaded design (Figure 52). Every time the prediction process is executed, the script also creates a case report that is stored within the CMS.

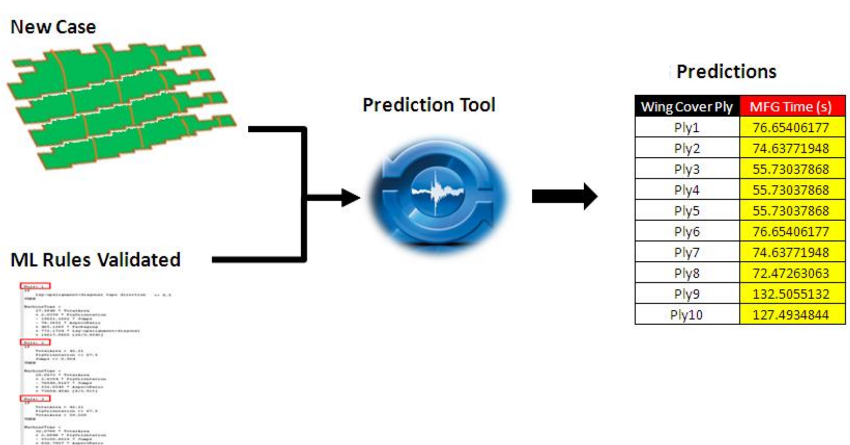


Figure 52. Prediction Tool.

6.2.2.5. Analysis and Delivery

The review of this use case involved qualitative and quantitative analyses used to validate the demonstrator implemented, thus supporting the correctness of the methodology proposed.

The quantification of benefits and costs delivered by the knowledge sourcing framework was not performed. However, a qualitative analysis was realised using a questionnaire completed by the users of the capability developed in the scope of this case study (APPENDIX B). Table 20 lists a summary of the benefits and costs related to the knowledge sourcing platform obtained from the analysis of the questionnaires.

Table 20. Case study 1: Benefits and Costs

Benefits	Knowledge Capture	Exploitation of automated reasoning methods to automatically generate engineering rules that otherwise would have been extracted from experts in high cost knowledge capture sessions, such as interviews or observation.
	Knowledge Retain	Knowledge captured from experts and ML algorithms is systematically stored using specific tools embedded in the knowledge-based platform. This methodological procedure enables the knowledge to be stored within a common data structure, providing users with a capability to easily retrieve and visualise the knowledge captured.
	Knowledge Reuse	Knowledge is stored in a computable format within formal and informal models. This procedure facilitates the exploitation of knowledge across different engineering problems.
	Knowledge Update	The capability enables the knowledge stored within formal and informal models to be updated through the employment of the user interface.
	ML Methods Reliability	The model generated and used by the AI algorithm to generate the predictions is stored in the CMS in a human and machine readable format. This together with the ability of performing actions such as modifying, adding or deleting the automatically created ML rules increases the reliability of the predictions.
Costs	Framework Implementation	The development of a generic KSF for a new use case is a highly time consuming task which requires the creation of a new data processing application and the definition of data structure.
	Data Dependencies	The accuracy of the results provided by this methodology is highly dependent on the quality of the data fed into the system. This data captured is used to carry out the learning process, thus trying to predict new designs which vary considerably from the ones used in the learning process could lead to inaccurate results.

6.3. Results and validation of the first case study

This section shows the outcome achieved as a result of realising the first case study. Moreover, a methodological procedure followed to evaluate and analyse the KSF platform is also described.

This section is organised as follows. In sub-section 6.3.1, a description of the data used in the learning and testing processes is presented. This is followed by a description of the method used to execute the ML algorithms and the selection of a ML technique in sub-sections 6.3.2 and 6.3.3 respectively. Finally, sub-section 6.3.4 presents the steps carried out to validate first the explicit model produced by the selected ML method and then the KSF.

6.3.1. Dataset description

Prior to obtain an appropriate set of rules modelling a problem, it is required to pre-process raw data and define a set of features which drive the problem objective (e.g. area or curvature changes have an impact on MFG time prediction of CFRP wing covers). In fact, this task is essential to create a good quality dataset, containing meaningful features and low noise level, enabling ML algorithms to accurately predict the target values.

The dataset used in this case study ,obtained as a result of the data pre-process carried out by knowledge managers and experts in the domain, consists of 269 samples containing data corresponding to 5 features (design descriptors) and their respective MFG times (output class). The dataset was divided in two different files: “Training Set” and “Test Set”. The “Training Set” is used by the knowledge creation engine to generate an explicit model in a procedure known as “Learning Process”. In this case, the “Training Set” contains around 75% of the existing samples corresponding to 5 wing designs. In parallel, the “Test Set” is the input file utilised in the validation process and it contains 71 samples (2 wing designs). Although the “Test Set” encompasses data corresponding to design descriptors (input values) and MFG cycle times (target values), only design values corresponding to the design descriptors are used to calculate the

MFG times, since the target values are used to evaluate the accuracy of the rules employed.

6.3.2. ML algorithm execution in the learning stage

In the method selection stage and in the learning phase of the machine learning process the approach followed to pre-validate the model was the use of a Cross Validation approach and the intervention of experts to analyse the rules based on their meaningfulness (modifying the rules if required).

Therefore, to allow a reliable comparison of the performance of the pre-selected ML algorithms in the pre-validation phase of the validation process, a technique entitled as Cross Validation (CV) was carried out in the learning process together with the expert evaluation. CV is an iterative process consisting of three phases. Firstly, the data belonging to the “Training Set” was randomly shuffled and divided into 10 cross-validation folders or subsets each containing the same number of samples. Secondly, in each experiment or iteration of the CV, nine out of the ten subsets are used to train; the other subset is used to test the model generated. This activity is realised until every subset has been used to test the model created by a specific iteration (ten times is this case). Finally, a ML model is provided from the averaged performance of the ten models generated in the CV process.

After every algorithm has applied the CV, the models produced were systematically analysed as described in the next section.

6.3.3. ML algorithm selection

The ML algorithms used in the learning process were selected after the realisation of a filtering process where only those ML techniques providing interpretative information where chosen. The selection method starts with the upload of a “Training Set” file into the system. Once the data is uploaded, the ML engine runs a set of algorithms previously selected. As a result, relevant information belonging to these algorithms is displayed on the screen (Figure 53). The algorithms selected in this case study were:

Linear Regression, REPTree (tree method) and M5R (rule based method). The method followed by the experts in order to select an AI technique was to analyse the level of understanding and the prediction accuracy of the AI rules.

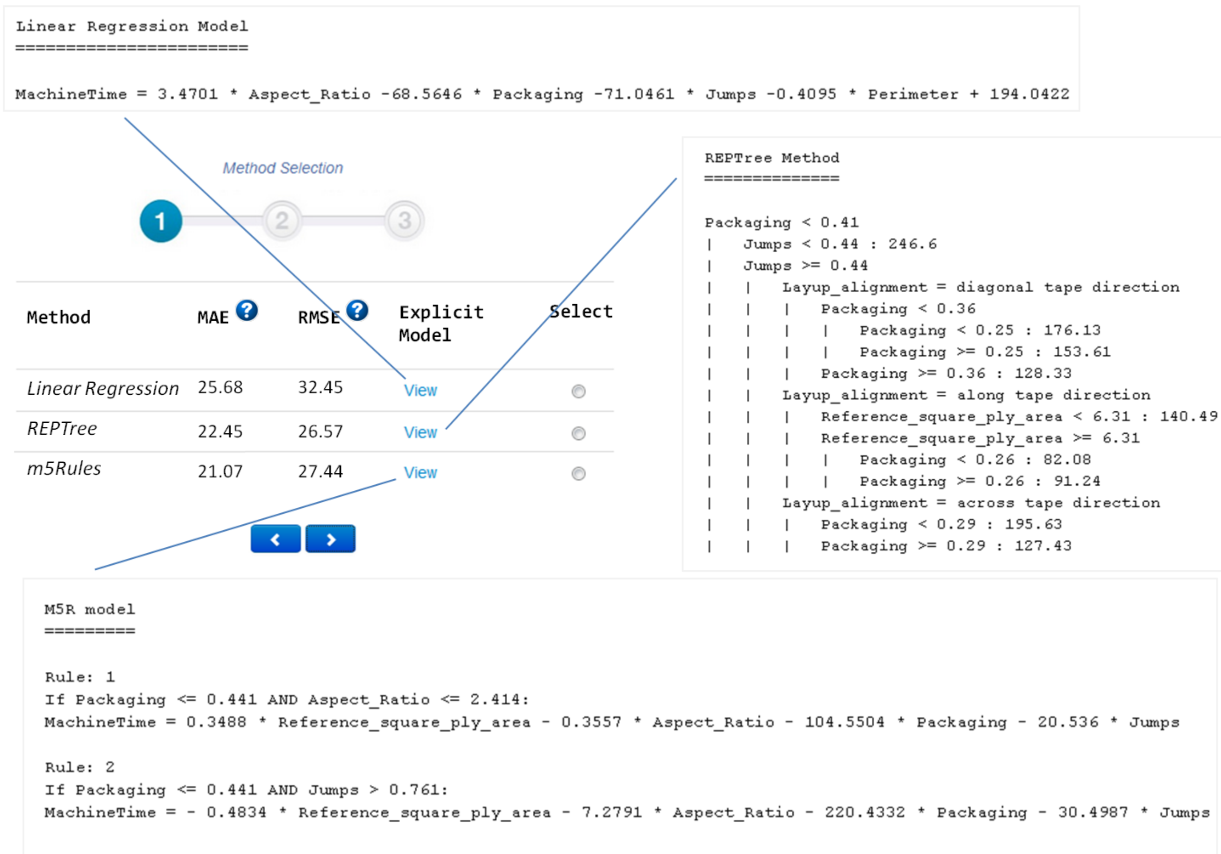


Figure 53. Explicit models displayed by the KSF platform

In this research, the evaluation of the prediction accuracy of the ML algorithms follows a common criterion acknowledged by the research community [195][196]. This criterion is based on using Mean Absolute Error (MAE) and Root Mean Squared Error (RMSE) as scoring values to rate the ML algorithms. Moreover, expert intervention is needed to analyse the performance and the level of understanding of the explicit models. In this case study three experts in problem domain reviewed the ML scoring values (MAE and RMSE) and the level of understanding of each explicit model created. Table 21 shows MAE, RMSE delivered by the ML algorithm utilised. This table also displays an average of the level of understanding of each of the models generated. MAE and RMSE were automatically produced by the ML algorithm whereas the level of understanding was input by the users of the KSF platform.

Table 21. Learning process: results obtained using CV.

ML method	**MAE**	**RMSE**	**Understanding of the model**	
			Level of understanding	Description
Linear Regression	25.68	32.45	Medium	Although the low number of parameters used in the model facilitates its understanding, experts believe that modelling such a complex problem into a single equation decreases the method's reliability.
REPTree	22.45	26.57	Medium	The use of a model containing conditional operators helps the users to better understand which parameters are affecting the manufacturing time. However, the tree format used by this types of algorithms makes it difficult to quantify the impact of each parameter.
M5R	21.07	27.44	High	The use of a familiar method employing "IF THEN..." rules and linear equations to model the problem facilitate the understanding of the rules.

Based on MAE, RMSE and the level of understanding of the models generated, experts selected M5R as the most suitable algorithm to be used in this case study. Although RMSE obtained by REPTree algorithm was lower than the one provided by M5R experts selected M5R due to from their point of view the rules generated by M5R where easier to understand and more meaningful than the ones created by REPTree and the difference in terms of MAE and RMSE was not relevant.

In order to rely on the predictions provided by the algorithm, it is necessary to validate the ML model. In case the results obtained in the validation stage are not satisfactory, the knowledge sourcing application enables experts to select a different algorithm and realise the review and validation phases with the new set of rules.

6.3.4. Rule management and use case validation

Manufacturing time predictions initially provided by the algorithm selected (M5R) had high accuracy rates (measured using MAE and RMSE scoring parameters) but still worse than the values considered as acceptable by the experts in the domain. Therefore, expert review and validation was required in order to allow engineers to rely on the model automatically generated. In this direction, three experts in the area of design for manufacturing reviewed and validated the created ML rules using the Rules Management Application (RMA) encompassed by the KSF platform. RMA enables experts to evaluate and validate the explicit model generated with the help of a visual analytical tool. The use of a visual analytical tool employing charts and tables allowed experts to:

- Identify trends in data.
- Understand the importance and generality of each of the rules generated.
- Trace back the results thanks to the information displayed on the line chart relating to the output values with their corresponding rules.
- Understand the cause of existing inaccuracies by comparing similar samples (e.g. using a table containing the input values used in the learning process).

The ability to iteratively review and modify the rules together with the benefits brought by the visual analytical representations allowed experts to add new knowledge. This activity often increases the accuracy and reliability of the model. Indeed, experts are more confident to base their decisions on the results provided by this capability.

The review and validation activities are part of an iterative process realised by experts consisting of (Figure 54):

- Modification of the ML model if required.
- Use of the visual analytical tool (e.g. charts, tables...) to better understand the model generated. It is also employed to identify trends in data and understand how the changes realised in the model affect the ML scoring values.

- Pre-validation of the ML model if MAE and RMSE values delivered by CV process are acceptable from an engineering point of view. The meaningfulness of the model is also a key factor taking into account for the pre-validation process.
- Use of the pre-validated model to predict the values of new samples contained in the “Test Set” file. If the values of MAE and RMSE delivered using the “Test Set” as input are considered as satisfactory the model is validated. Otherwise, if the model is not validated the user can select a new algorithm and continue with the review and validation activities.

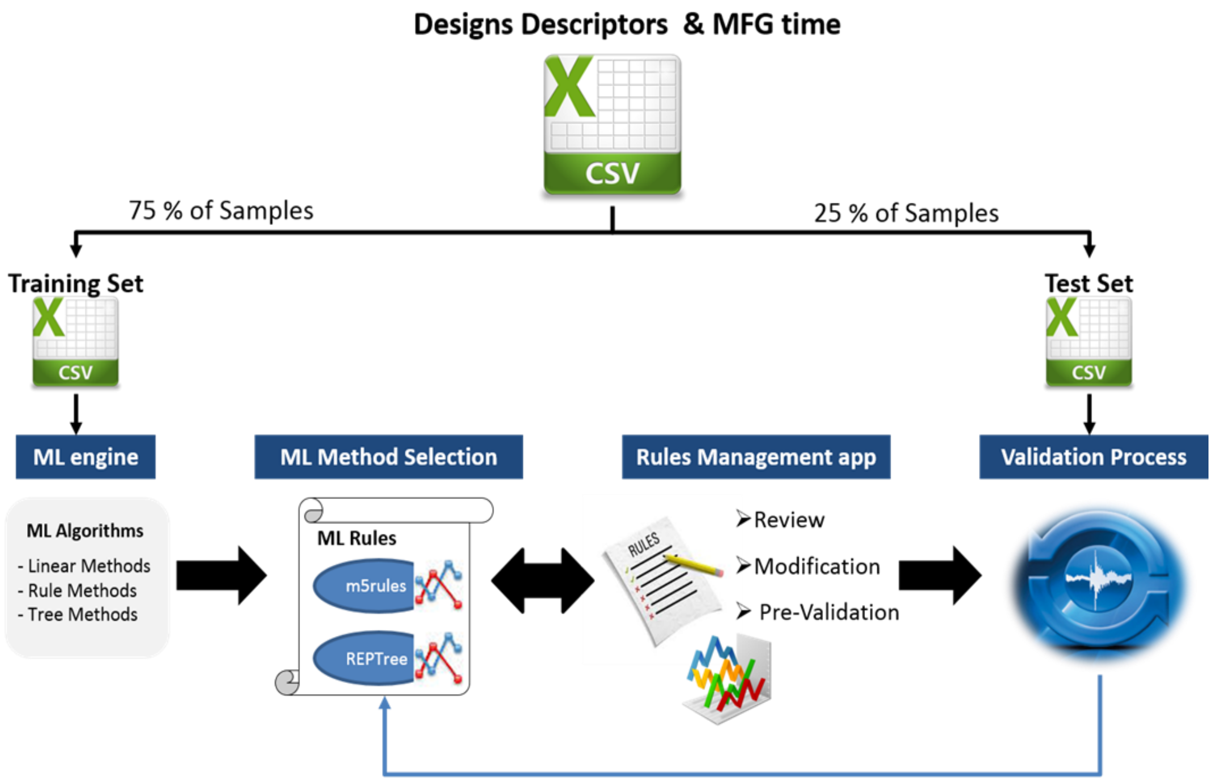


Figure 54. Procedure followed to validate the model.

The results of the learning and validation processes are summarised in Table 22. From this table, a reduction of MAE and RMSE is observed. This was achieved by experts who made the necessary modifications to the ML model in order to reduce the gap between the actual and the predicted values.

The modifications performed on the rules were carried out using the analytical tool provided by the KSF platform. In this regard, Figure 55 and Figure 56 show the results provided by the initial set of rules (delivered by the AI algorithm before expert review) ,and the results delivered by the rules after expert pre-validation respectively. The results displayed on the images below include the rules used to generate the predictions, a pie chart describing how often has each rule being used, the accuracy values used to quantitatively evaluate the performance of the rules, and a table with the input data used in the learning process. In summary, these figures illustrate the improvement of the predictions accuracy when using a set of rules created through the collaboration between experts and machine learning algorithms.

Table 22. Summary of M5R results.

	Learning Process		**Validation Process**	
ML model used	MAE	RMSE	MAE	RMSE
Initial model provided by M5R	21.07	27.44	20.26	27.31
Model reviewed and validated by experts	10.31	15.23	13.91	19.83

The use in the validation process of the pre-validated rules of the M5R algorithm (obtained after expert review and modification of the model) showed a substantial reduction of the MAE and RMSE values (Table 22). The machine learning scoring values obtained in the validation phase were considered by the experts in the domain as acceptable from an engineering point of view.

The pre-validated rules were also considered by the experts as meaningful. In this regard, experts realised that rule 5 in Figure 55 did not account for the impact caused by the direction the material is being laid (represented by the “orientation” parameter). This particular event identified by experts was causing some data inaccuracies which are visually shown in the charts displayed in the RMA user interface. As observed in Figure 56 the modifications realised regarding rule 5 provided better accuracy values (MAE and RMSE) and the deviations represented in the line chart were considerably reduced. Therefore, using the RMA, experts are able to correlate physical events coming from their experience with the results represented in the visualisations.After expert evaluation of the MAE and RMSE scoring values, and the model meaningfulness, the rules were validated.

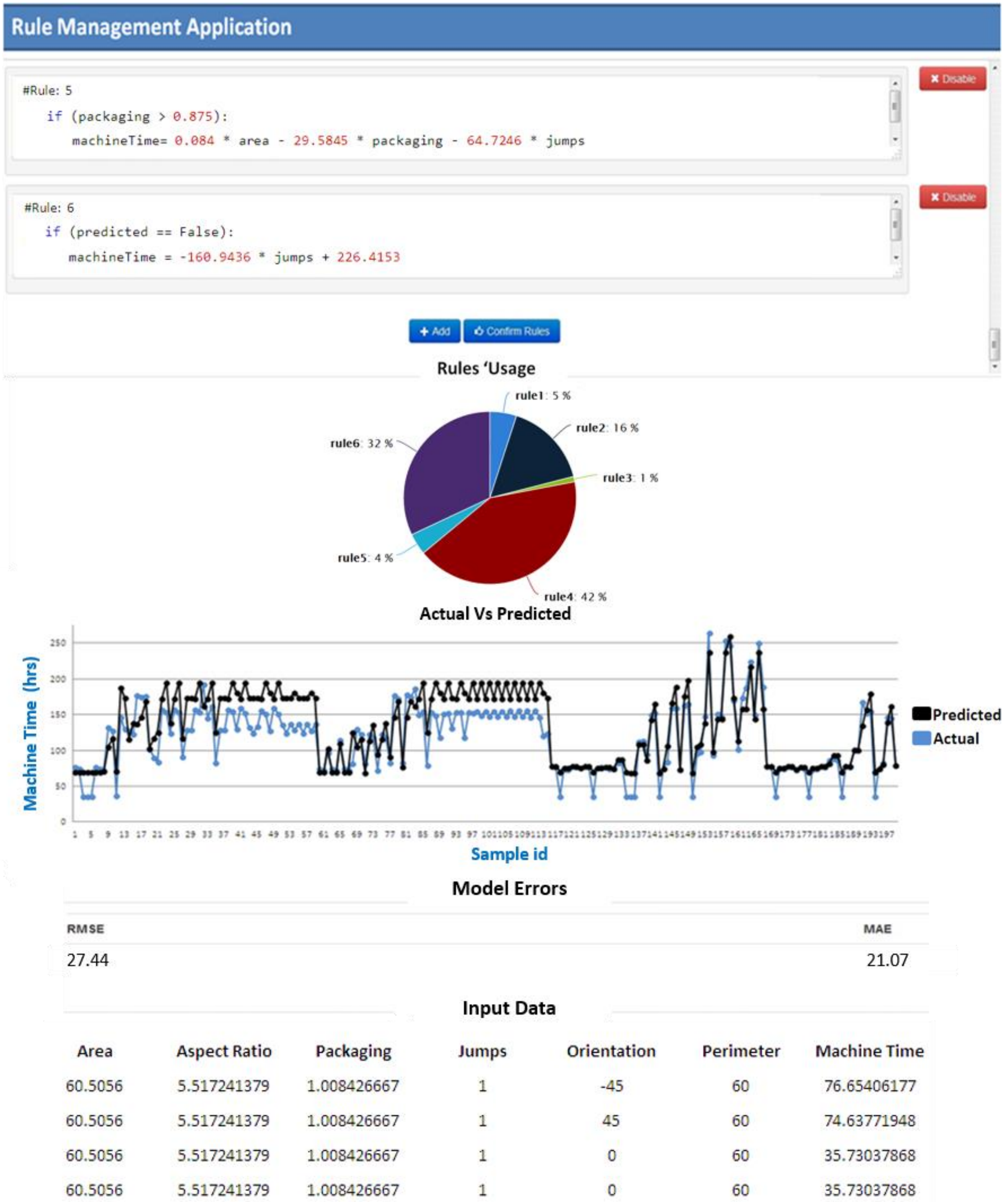


Figure 55. Results generated in the CV process by non-reviewed rules

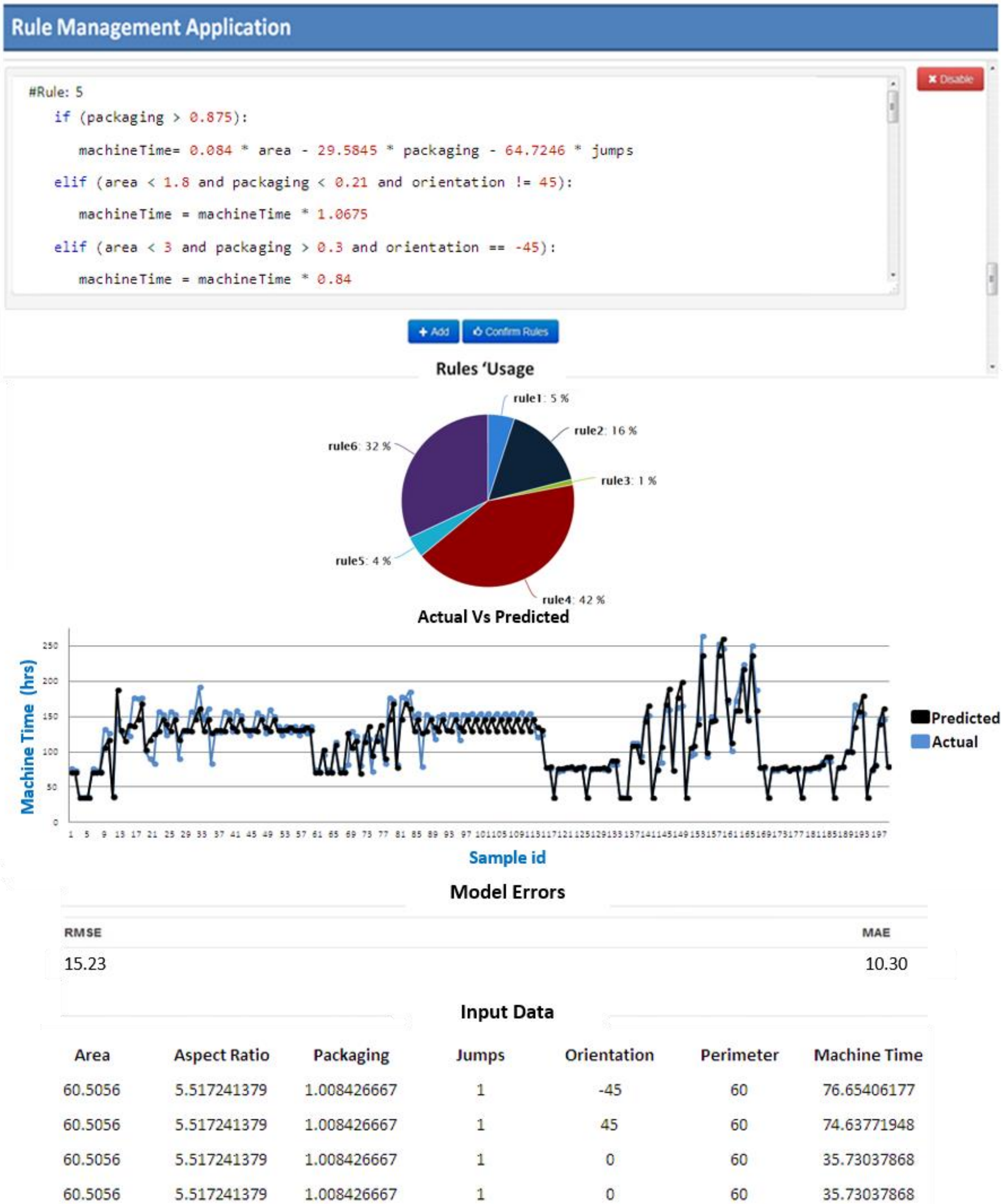


Figure 56. Results generated in the CV process using the rules reviewed and modified by experts

6.4. Results and validation of the second case study

This section shows the outcome achieved as a result of realising the second case study. Moreover, a methodological procedure followed to evaluate and analyse the KSF platform is also described. The process followed in order to obtain the model and validate the results is the same as the one performed in the first use case. Due to data confidentiality, the names of the parameters used as design descriptors are not specified in this chapter. However, the rules and results described in this section are the ones obtained in the real case.

6.4.1. Dataset description

The dataset used in this case study was obtained as a result of the data pre-process carried out by knowledge managers and experts in the domain. It consists of 61 samples containing data corresponding to 5 features (design descriptors) and their respective nut cap pressure values (output class). The dataset was divided in two different files: “Training Set” and “Test Set”. The “Training Set”, containing design descriptors and nut cap pressure values, is used by the knowledge creation engine to generate an explicit model. In this case, the “Training Set” contains around 70% of the existing samples (fastener assemblies). In parallel, the “Test Set” is the input file utilised in the validation process and it contains around 30% of the existing samples (not including the output class values).

6.4.2. ML algorithm selection

The ML algorithms used in the learning process were selected after the realisation of a filtering process where only those ML techniques providing interpretative information where chosen. The selection method starts with the upload of a “Training Set” file into the system. Once the data is uploaded, the ML engine runs a set of algorithms previously selected. As a result, relevant information belonging to the pre-selected algorithms is displayed on the screen (Figure 57). The information displayed on the screen belongs to three machine learning algorithms: Linear Regression, REPTree (tree method) and M5R (rule based method). The method followed by the experts in order to select an AI technique was to analyse the level of understanding and the prediction accuracy of the AI rules (RMSE and MAE scoring values).

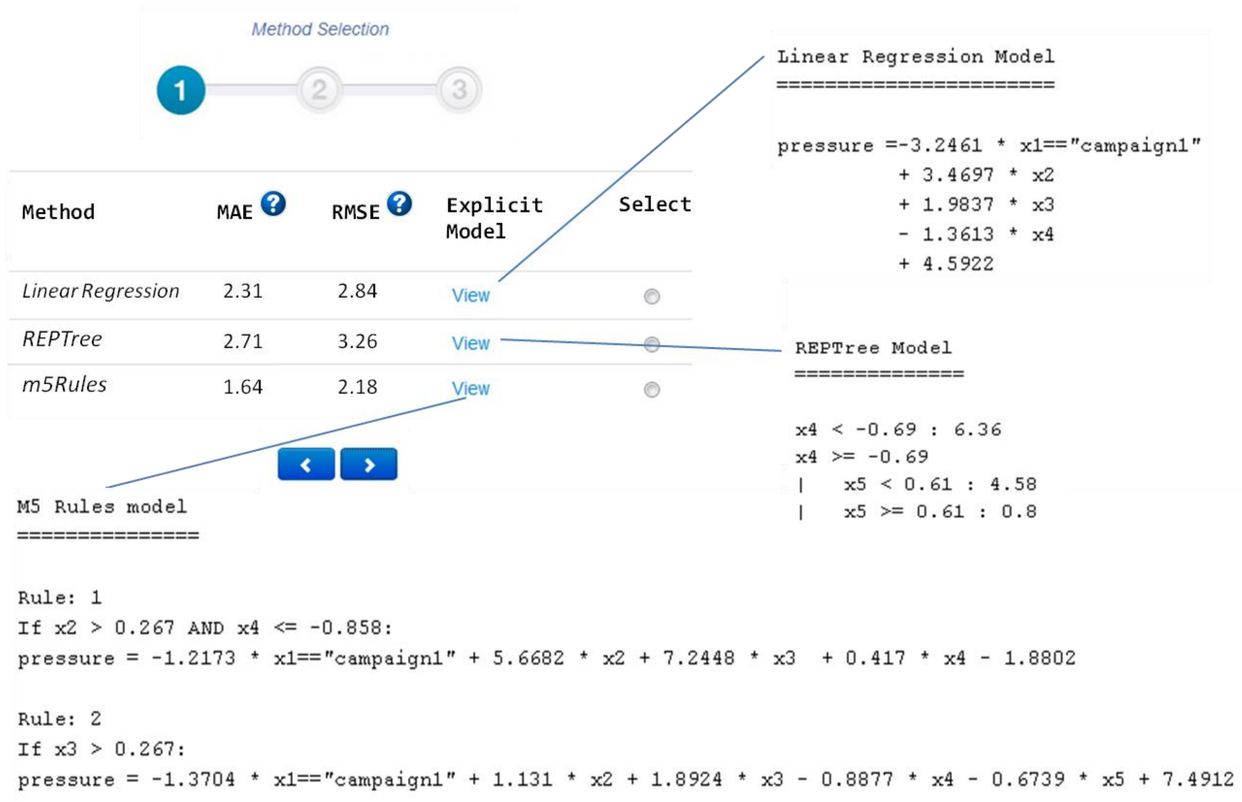


Figure 57. Explicit models displayed by the KSF platform

In this research, the evaluation of the prediction accuracy of the ML algorithms was realised MAE and RMSE as scoring values to rate the ML algorithms. In addition, due to expert intervention is required to validate the ML model generated three experts with more than 10 years of experience in engineering design for lightning protection reviewed the ML scoring values (MAE and RMSE) and the level of understanding of each explicit model created. Table 23 presents MAE, RMSE and an average the level of understanding delivered by each ML algorithm executed within the KSF platform. MAE and RMSE were automatically produced by the ML algorithm whereas the level of understanding was input by the users of the KSF platform.

Table 23. Learning process: results obtained using CV.

ML method	MAE	RMSE	Understanding of the model	
			Level of understanding	Description
Linear Regression	2.31	2.84	High	The low number of parameters used in the model facilitates its understanding.
REPTree	2.71	3.26	Medium	The use of a model containing conditional operators helps the users to better understand which parameters are affecting the manufacturing time. However, the tree format used by this type of algorithms makes it difficult to quantify the impact of each parameter.
M5R	1.64	2.18	Medium/High	The use of a familiar method like "IF THEN..." rules and linear equations to model the problem facilitate the understanding of the rules.

Based on MAE, RMSE and the level of understanding of the models generated, experts selected Linear Regression as most suitable method to be used in this case study. In this case, Linear Regression was initially selected as the most suitable method due to from the domain experts' point of view the model provided by the Linear Regression was the easiest to understand and the difference in terms of MAE and RMSE compared to M5R were not considerably different to decide which method would behave better.

To rely on the predictions generated it is first required to validate the ML model. If the results obtained using the selected model the KSF platform enables the user to go back in the process and select a new ML algorithm and realise the review and validation phases with the new model created.

6.4.3. Rule management and use case validation

Nut cap pressure predictions given by the algorithm selected (Linear Regression) had promising accuracy rates (measured using MAE and RMSE scoring parameters). However, expert review and validation was required due to MAE and RMSE values were slightly higher than the value considered as the maximum error accepted for this specific problem (defined by experts in the domain). In this case study, three experts in the area of lightning protection reviewed and validated the created ML rules using the Rules Management Application (RMA) included in the KSF platform.

Using the RMA experts managed to improve the values of MAE and RMSE initially provided by the AI algorithm. However, the results obtained in the validation process using the "Test Set" were not acceptable from an engineering point of view as the maximum allowance established by domain experts for MAE and RMSE was a value not greater than 3 units (Table 24). This maximum allowance for MAE and RMSE scoring values was decided by experts in the domain based on the range of values of the output class (nut cap pressure) and the engineering design stage corresponding to this problem which in this case was the conceptual design phase. Therefore, the experts did not validate the model and proceed to the selection of a new algorithm.

Table 24. Results summary.

	Linear Regression				M5R			
	Learning Process		Validation Process		Learning Process		Validation Process	
ML model used	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE
Initial model provided by AI method	2.12	2.66	3.87	4.71	1.64	2.18	1.89	2.21
Model reviewed and validated by experts	1.92	2.48	4.67	6.77	1.37	1.88	0.9	1.36

After dismissing linear regression, experts selected M5R algorithm. This decision was made based on the low error values of M5R, and a level of understanding of the rules considered as "medium/high" by experts (Table 23).

Figure 58 and Figure 59 show the results provided by the initial set of rules (generated by the AI algorithm before expert review) and the results delivered by the rules after expert pre-validation respectively (of the M5R algorithm).The figures below show the improvement of the prediction accuracy when using a set of rules created through the collaboration between experts and machine learning algorithms enabled by the .

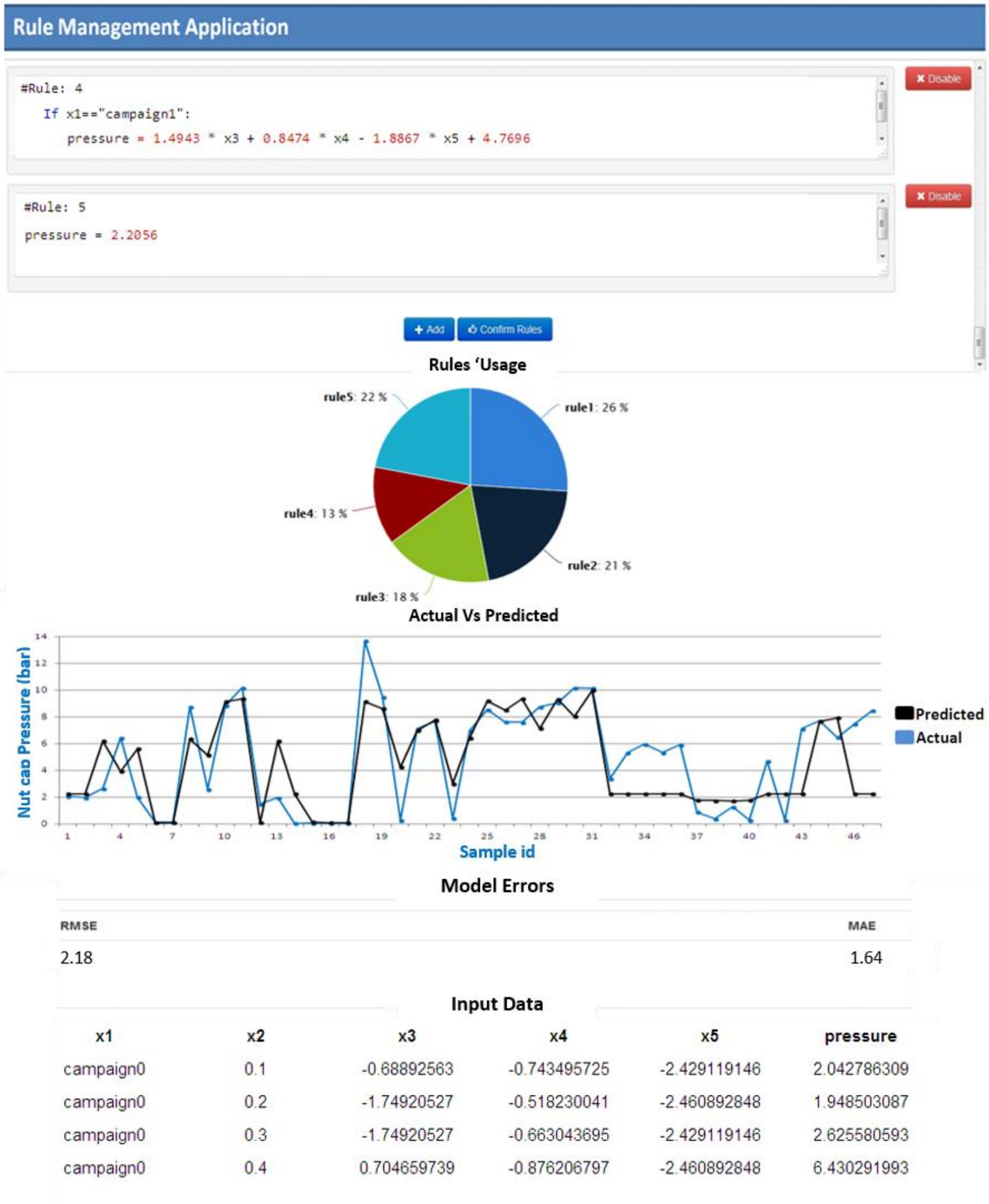


Figure 58. Results generated in the CV process by non-reviewed rules

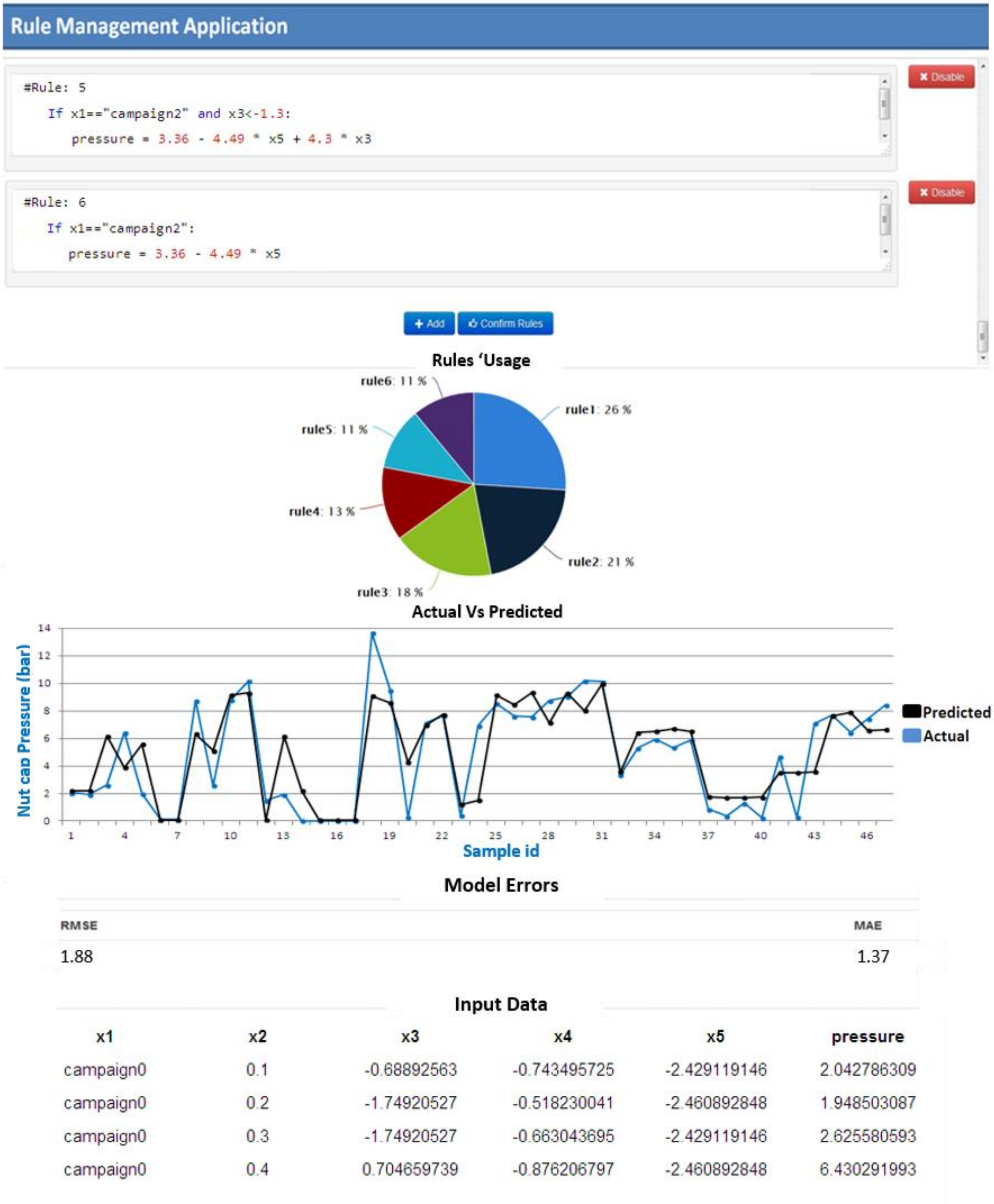


Figure 59. Results generated in the CV process using the rules reviewed and modified by experts

The use of the M5R pre-validated rules in the validation process (obtained after expert review and modification of the model) showed a considerable reduction of the MAE and

RMSE values Table 24. Considering the maximum error allowance established by the experts for this specific problem, MAE and RMSE values of 0.9 and 1.36 were acceptable from an engineering point of view. In addition, experts considered rules to be meaningful. Therefore, after evaluation of the MAE and RMSE scoring values, and the model meaningfulness, experts validated the review rules provided by the M5R algorithm.

6.5. Discussion of the results

The potential of the methodology proposed is evaluated in this case study, showing the benefits of adopting a generic framework for knowledge management, and integrating machine and expert knowledge in a common environment. The use case was developed following the KNOMAD steps aiming at sourcing knowledge in a systematic manner. In this context, special attention was paid to the data preparation activity, carried out as part of the knowledge capture process. Indeed, the data preparation is considered as the most time consuming task in the development of the KSF service [197][198].

To support the assumptions made in this study, the validation of the case study was a must. The use case verification was carried out through the acceptance of the ML model initially provided by the data mining tool and reviewed by experts. To validate the ML model, it was required to use testing data obtained under the same conditions as the training data. In this direction, the feature recognition tool and the simulation software used in the knowledge capture process to generate “Training Set” and “Test Set” did not go through any modification.

The ML methods included in the KSF were executed using data contained in the “Training Set” to generate a set of ML models (one model per algorithm).

A. Discussion of the results obtained in the first use case

The machine time predictions captured varies from 20 to 300. Considering this, experts established a MAE and RMSE lower than 20 as acceptable for the MFG cycle time prediction problem.

The models automatically generated in CV process did not contain low error values as shown in Table 21. However, after the experts selected M5R algorithm and modified its model, the error of the predictions obtained in the CV was considerably reduced from MAE of 21.07 and RMSE of 27.44 to MAE of 10.31 and RMSE of 15.23.

To avoid over-fitting the model to the data, a validation process was performed using the pre-validated model to predict the samples of a new data set named as "Test Set". The "Test Set" did not contain data utilised in the learning process. Therefore, the use of different samples in the validation process increases the reliability of the model. MAE and RMSE values obtained using the "Test Set" were lower than 20 units. Therefore, taking into account the maximum allowance for MAE and RMSE (using a confidence level of 95%), the model was validated by experts.

B. Discussion of the results obtained in the second use case

The use case verification was carried out through the acceptance of the ML model initially provided by the data mining tool and reviewed by experts. Taking into account that pressure values varies from 0 to 12, experts considered an error (MAE and RMSE) lower than 3 units as acceptable for the nut cap pressure prediction problem.

The ML methods selected after the filtering process were executed using data contained in the "Training Set" to generate a set of ML models (one model per algorithm). At the learning stage, the initial model automatically generated using the Linear Regression algorithm had already low error values. However, the poor results obtained in the validation process led the experts to reject the rules provided by the Linear Regression method. As a consequence, experts selected M5R as the algorithm to be further studied. Therefore, review and validation processes were realised again but this time with a different model. After the experts modified the rules provided by the M5R algorithm (using CV to evaluate the model's performance), the error of the predictions was reduced from MAE of 1.64 and RMSE of 2.18 to MAE of 1.37 and RMSE of 1.88.

To avoid over-fitting the model to the data, a validation process was performed using a new data set named as "Test Set". The "Test Set" did not contain data utilised in the learning process. The use of different samples in the validation process increases the reliability of the model. Using the rules modified by experts the error values in the validation phase were considerably reduced from MAE of 1.89 and RMSE of 2.21 to MAE of 0.9 and RMSE of 1.36 (see Table 24). From the evaluation (using a confidence level of 95%) of MAE and RMSE values generated in the validation process and the analysis of the significance of the rules, the pre-validated M5R model was finally validated.

Based on the results, it is proved that the integration of machine and expert knowledge enables the reliable and fast evaluation of design concepts with acceptable accuracy values (obtained through evaluation of MAE and RMSE scoring values).

6.6. Concluding remarks

The creation of two different services providing predictions of manufacturing time for wing covers and the predictions of the pressure in the nut cap of a fastener assembly are presented in this chapter. The results obtained in the first use case supported the foundations of this research whereas the results obtained by the second use case confirmed the assumptions initially made in this report. These assumptions stated that the development of a KBE approach integrating ML algorithms and expert knowledge would deliver an effective and more efficient approach for sourcing engineering knowledge.

Prior the development of this research, knowledge associated to the MFG cycle time or nut cap pressure estimation problems was kept in the minds of the domain experts or in documents locally stored. These practices provoke knowledge leaks in the case of expert leave or retirement, thus making difficult to adequately retain and reuse engineering knowledge. Therefore, it became apparent the need for an effective approach for managing engineering knowledge. In this regard, this research proved to be an effective solution for managing engineering knowledge by providing experts with a methodology fostering the systematic capture, retain and reuse of knowledge. Using the proposed framework, experts are now asked to store the knowledge used by the capability in a human readable format within independent knowledge packages placed in a centralised web-based platform. In doing so, knowledge becomes easily accessible and applicable into different engineering problems across the organisation.

In the case of the first use case, the previous approach used for time estimation lacks a model of the problem due to complexity issues. The complexity of the process is mainly provoked by the use of immature technologies –highly dependent in part complexity– in the manufacture of CFRP wing covers. Experts did not have enough knowledge to create an accurate model capable of predicting the manufacturing cycle time of wing covers. Moreover, any relevant knowledge used for the time estimation was not accessible for the experts as it was hard coded within the simulation software application. Based on these issues regarding the knowledge capture activity, the proposed knowledge sourcing capability proved to be able of efficiently capturing relevant and meaningful knowledge from company data assets and elicit expert knowledge. In this context, the knowledge sourcing platform integrates methods and tools supporting the capture of expert knowledge, and automating the knowledge creation using AI algorithms. Therefore, this research delivers a methodology for fast knowledge capture, thus providing a more efficient knowledge sourcing process. This approach is also a potential solution for problems which are knowledge intensive, complex and have poor theoretical understanding.

The results provided by the knowledge sourcing capability highlight its ability for delivering fast, accurate and reliable evaluations of design concepts. The proposed methodology provides MFG time estimations in just a few seconds whereas the technology commonly used to do the same activity (fidelity simulation software tools) requires more than a week of man work to generate the machine time estimations. Therefore, the knowledge sourcing framework developed is considered more efficient compared to current approaches as it provides faster MFG time estimations with accuracy levels considered as acceptable by experts in the domain.

In parallel, in the context of the second case study before the KSF was implemented the rules modelling the nut cap pressure problem were manually created by experts. To do that, this group of experts used to identify trends in data through the analysis of the data gathered from the design and produced in the lab experiments. To facilitate the data evaluation process, the experts used a range of charts provided by the Excel tool. However, the poor results delivered by the model created led experts to realise about the limitations of using Excel spreadsheets to analyse complex problems which have poor theoretical understanding. More precisely, they considered that using Excel spreadsheets for data analysis makes difficult to identify and measure the impact caused by the parameters' interaction. Therefore, it became apparent the need for more advanced data analysis tools in order to realise the efficient capture of engineering knowledge. In this context, the AI techniques embedded within the KSF provide a solution to complex problems by generating a set of rules modelling the problem from company data assets. The AI algorithms account for data interactions, thus often they provide a more accurate model than using traditional approaches. This is supported in this research by the results provided by the application developed in this case study. Therefore, the new methodology proposed in this work provides a more accurate knowledge capture process, thus delivering a more efficient knowledge sourcing framework.

Summarising, the research objectives 2–6 are accomplished through the development of a knowledge sourcing service following the steps of a well-established and generic methodology for KBE development, and the integration of methods and tools enabling the collaboration of AI techniques and experts. The validation of the case study and the research objectives associated was achieved by:

- Providing expert feedback, generated as a result of the analysis of the ML rules and filling a questionnaire to assess the platform developed (APPENDIX B).

- Quantitative analysis of the results obtained in the review and validation of ML rules.

7. Discussion and Conclusions

This chapter aims at providing the reader with a detailed analysis of the research achievements while describing the degree of quality, generality and applicability of the methodology proposed in section 7.1. This is followed by the description of the key research contributions in section 7.2. In addition, limitations and future work associated to this research are presented in section 7.3 and 7.4 respectively. The chapter is concluded with a summary of the work presented in this research in section 7.5.

7.1. Research achievements

The research reported here presents a new methodology for knowledge sourcing. The aim of the framework is "to develop a new and more efficient Knowledge Sourcing Framework, and create a better link between KBE and knowledge sourcing". By implementing the new approach, engineering knowledge is systematically captured, retained and reused while providing engineers with a more efficient knowledge discovery method.

The definition of research objectives led to the identification of a research challenge regarding the integration of the existing schools of thought aiming at a more appropriate management of engineering knowledge. The achievement of each of the research objectives reported in this research is synthetized in Table 25.

Table 25. Achievement of the research objectives.

Research Objective	Chapter/ Section	Description
1	3	Previous work addressing the efficient sourcing of engineering knowledge is analysed from an industrial perspective. In doing so, limitations representing the key barriers stopping KBE from being a widely used methodology to capture, retain and reuse knowledge were consolidated.
2	5.3	The knowledge sourcing platform reported in this research encompasses a method where the user is guided in the selection of a suitable AI method, and the review and validation of the model automatically generated by the AI technique. Using the information provided by the platform, the user selects a ML algorithm by evaluating the accuracy of the predictions, and the format and

		meaningfulness of the rules.
3	5.5	A new method using AI for faster knowledge discovery is implemented within the knowledge sourcing framework. The AI algorithms deliver a set of explicit rules modelling the problem which provides experts with a better understanding of the problem. In doing so, the faster creation of a definitive model used to generate predictions is facilitated.
4	5.4	This research presents a combined approach integrating AI methods delivering an explicit model and expert knowledge in the form of expert intervention. The involvement of experts at different stages of the knowledge sourcing methodology increases the quality of the data used in the learning process and the reliability of the AI rules (after expert review and validation).
5	5.2	Knowledge and applications associated to the knowledge sourcing capability are stored in independent knowledge packages which are classified within three different engineering resources. This way of managing the elements of the framework facilitates their reuse across different engineering problems and their update.
6	4.2, 6 & 7	The adoption of a well-established methodology supporting the management aspect of KBE is reported in this research. Following a structured approach for knowledge sourcing enables the systematic capture, retain and reuse of engineering knowledge.
7	6 & 7	The knowledge sourcing approach is validated as a result of the successful implementation of two case studies that prove the theoretical foundations of the research.

In addition, the research objective achievement is assessed through the analysis of the knowledge sourcing framework regarding its quality, generality and applicability.

7.1.1. Quality of the research process

To ensure the quality of the key findings obtained in this research, all the knowledge used in the knowledge sourcing framework was systematically captured. In this direction, questions planned to be asked in the meetings were discussed with the interviewees prior to performing the interviews. This step provided the interviewer with feedback regarding the suitability of questions and gave the experts more time to think about the answers, thus enhancing the elicitation of tacit knowledge. Moreover, experts' answers were stored in a content management system, improving knowledge retention. The interviews were subject to expert's availability, however this constraint was minimised as the author of this study was permanently based at the company were the case studies were carried out, thus facilitating the interaction with the experts.

The suitability of the methodology followed to develop the knowledge sourcing framework was ensured through to the use of a well-established and generic methodology (KNOMAD) which supports the development of KBE systems. The selection of this KNOMAD was carried out as a result of evaluating some of the most common methodologies used for KBE development. The evaluation process classified each of the methodologies regarding a set of knowledge management aspects such as the level of support for knowledge reuse and knowledge update. Moreover, these methodologies were also classified within general aspects such as their level of applicability and flexibility.

To validate this research, two case studies were implemented within the knowledge sourcing framework. Both use cases face problems which are knowledge intensive, complex and have poor theoretical understanding. The characteristics of these problems and their context (aerospace industry) led to their selection as case studies and their posterior implementation within the framework developed in this research.

The assessment of the achievement of all the specific research objectives of this study, apart from the first objective which is related to the literature review, was realised through quantitative and qualitative analyses. The aim of the quantitative analyses is the correctness of the results provided by the rules delivered as a result of the interaction between AI applications and expert knowledge. In parallel, two different procedures were carried out regarding the qualitative analysis. The first is focused on assessing the level of understanding of the automated ML rules whereas the second aimed at the collection of evidences about the usability and performance of the developed framework. The procedure followed to validate the outcome of this research is synthesised in Table 26.

Table 26. Validation of the research outcome.

Objective (see section 1.3)	Type of Validation	Description
2. Help engineers to identify a suitable AI tool to solve a specific problem.	Qualitative and quantitative	The selection of a suitable ML algorithm performed by a group of experts was realised using a procedure which combines qualitative and quantitative features. The evaluation of ML numerical scoring values (RMSE and MAE) and the analysis of the ML rules based on their level of understanding

		constitute respectively the quantitative and qualitative characteristics of the validation process.
4. Increase the reliability of AI knowledge-based implementations, allowing experts to review and validate the ML rules generated	Quantitative	After the selection of a specific ML technique, the user (expert in the problem domain) evaluates the explicit model automatically generated and modifies the rules employing the RMA. In the iterative process followed when using the RMA, the pre-validation of the rules is achieved once the expert accepts RMSE and MAE as acceptable from an engineering point of view. The possibility of reviewing and validating the explicit model (responsible for generating predictions) increases the reliability of using machine learning algorithms to support engineers in solving complex problems.
3. Enable the efficient sourcing of engineering knowledge using AI techniques.	Quantitative and qualitative	The verification of this objective encompasses the two objectives listed above and the validation of the explicit model. The validation of the explicit model is realised evaluating RMSE and MAE values obtained when using the pre-validated rules to predict the output class values corresponding to new samples contained in the "Test Set" input file. The model is validated as long as the experts consider this RMSE and MAE values as acceptable from an engineering point of view.
2 – 7	Qualitative	The applicability and performance of the presented framework were qualitatively analysed by asking the experts involved in this research to fill a questionnaire (APPENDIX B). This questionnaire enabled the assessment (from a user point of view) of the appropriateness of the framework, thus providing the author with useful information about the importance of the features delivered by the presented methodology.

7.1.2. Generality of the research methodology

Regarding the scope of the methodology presented in this study, the author believes the framework is generic enough to be applied in different business contexts. This is supported by the fact that similar research studies have been successfully implemented in other type of industries [100],[199],[200]. However, no safe claim can be realised as the two case studies implemented in this work have been only carried out in the aerospace context.

The scope of this research is constrained to the aerospace industry. In this context, the characteristics of the aerospace industry defining the scope of this work are:

- Lack of knowledge at early design stages, intrinsically related with the complexity of the products

- Complexity issues faced in the definition and application of tools in early design stages. This raises some challenges associated with the need to increase the efficiency of the knowledge sourcing practices.
- High impact of the design choices made at early stages. There is significant interest from aerospace companies to increase the knowledge sourcing for early knowledge exploitation.

In short, this research proved to be an effective and more efficient approach for knowledge sourcing in the context of the aerospace industry. Further research is required in order to prove the success of the proposed framework for a wider scope.

7.1.3. Applicability of knowledge sourcing framework

The analysis of the potential business impact of the implementation of the knowledge sourcing framework (KSF) is explained in this section. The application of the KSF enables engineers to make informed decisions when facing complex problems which have poor theoretical understanding; for instance, the design and manufacturing of CFRP components employing immature technologies which are highly dependent on part complexity. Additionally, designers working at early stages of the engineering design process could use the KSF to generate new knowledge based on past experiences. More precisely, the KSF would work as a design scoring tool helping designers to choose an optimal design.

Apart from the advantages mentioned above, a major argument to implement the proposed framework is the need of every company to retain expert knowledge, thus minimising knowledge leaks caused by employees either leaving or retiring.

To adequately implement a new service in the KSF developed in this research, there is a need to adapt the methodology to the client requirements. First, the definition of the data structure and its domain specific ontology is required, enabling systematic knowledge capture and storage. Once the structure of the knowledge to be captured is defined and, the required EKR models and the inference tool in charge of extracting data from input files are created, minor changes must be made to the KSF in order to implement a new service. Only the requirement of new algorithms apart from the ones already embedded in the existing KSF would create the need for a programmer to integrate the desired method within the KSF. This could be a simple task but it will depend on the library or software package containing the algorithm requested. Furthermore, the KSF has been designed to be widely used and therefore its use does not need any special training. Indeed, the KSF has been considered by the users as an intuitive platform simple to use –as registered in the questionnaire about the KSF shown in Appendix B.

In summary, the impact of implementing a new service in the KSF is considered to be from medium to low depending on the complexity of the process or processes to be modelled. The business impact will mainly depend on the level of expert involvement required at the knowledge capture stage. The efficient creation and reuse of engineering knowledge are the major features motivating the implementation of the KSF presented.

7.2. Key Research Contributions

In the last two decades, the increasing interest and research on AI and KM areas has led to new ways of realising engineering design activities. In the engineering design context, researches having into account both areas are mainly focus on providing a solution to knowledge intensive problems (see section 2.4.2). In this context, capturing and retaining expert knowledge in a transferable format enabling its future reuse is of high interest for organisations. An example of this is the EPSRC grand challenge project on information and knowledge management [201]. Despite of the efforts realised by the research community to deliver a more effective and efficient knowledge sourcing approach, there are still some key barriers in the way expert knowledge is capture and the generality of the approach used for managing engineering knowledge.

Existing KBE approaches usually capture expert knowledge through traditional elicitation techniques which are highly time consuming for experts. Moreover, KBE frameworks are rarely supported by a generic methodology (such as MOKA or

CommonKADS) providing the required methodological support to manage the complete knowledge life cycle. No KBE research was found presenting a capability managing engineering knowledge systematically (using a generic approach) while using advanced techniques for fast knowledge capture. This represents the key research finding identified in this work.

The research reported here, presents a methodology delivering a solution to the knowledge capture and management aspects of KBE. This approach is also considered by the author as an integration of personalisation and codification EKM views for improved knowledge sourcing. Personalisation tools aim at improving the knowledge sourcing process by providing a framework for better collaboration between engineers whereas codification tools seeks the enhancement of knowledge sourcing through the automation of tasks [202].

The methodological support provided by the adoption of a well-established methodology for KBE development is a personalisation feature whereas the use of AI techniques for knowledge discovery is a codification feature. The specific contributions obtained as a result of combining these two features are listed as follows:

- Knowledge capture in engineering practices.
- Knowledge life cycle management in engineering design.

7.2.1. Knowledge capture in engineering practices.

Nowadays KBE is considered by the research community as a key feature to capitalise company efforts on past activities. The knowledge capture aspect of KBE has been identified as the major bottleneck to hamper its use in real industrial applications. Therefore, the use of an efficient knowledge capture technique to retain and enable the transfer of lessons learn is crucial for organisations.

Engineering knowledge is usually extracted using traditional techniques such as expert's interview or observation. In this context, the use of AI for the efficient capture of engineering knowledge is considered a strong candidate. However, little research is found where AI is used within a KBE framework for knowledge capture as AI is mostly used for “active” design optimisation whereas sourcing of information and knowledge is perceived as a secondary and therefore “passive” need. Identified researches integrating a knowledge based system and AI tools proved to be a solution capable of reducing the time required from experts [66],[70],[90],[91]. Nevertheless, none of these researches use a generic methodology capable of systematically managing the complete knowledge life cycle.

Reliability is the main limitation of AI being applied in an engineering context. Most of the existing algorithms behave as a “black box” application not providing engineers with an explicit set of rules modelling the problem, thus they are not able to trace back from the results obtained. Only a few studies employ algorithms delivering an explicit model (see section 0). Therefore, it becomes apparent the need for a using AI algorithms delivering interpretative information enabling the explicit link between the predictions and the rules employed to generate them. Moreover, some researches using AI for knowledge discovery highlight the need for a closer link between AI algorithms and experts in the form of expert collaboration [67]–[69],[89],[91]. From these studies using AI and expert intervention, only [62] employs a knowledge based system for knowledge retain. However it doesn’t provide support in other stages of knowledge life cycle such as knowledge modelling and knowledge update.

The approach reported here was developed following the steps of a generic methodology supporting the management aspect of KBE. This approach includes a structured method combining AI algorithms and experts for faster and more reliable knowledge capture compared to existing researches. The AI techniques used in this approach were selected after realising an assessment process. This evaluation is based on the ability of the AI technique to provide an explicit set of rules modelling a problem which allows the expert to realise the corresponding review and validation activities. The knowledge sourcing approach proposed encompasses an application which facilitates the expert intervention at different stages of the knowledge capture process. More precisely, the method defined requires expert collaboration in the data preparation, model selection and, model review and validation tasks (Figure 60).

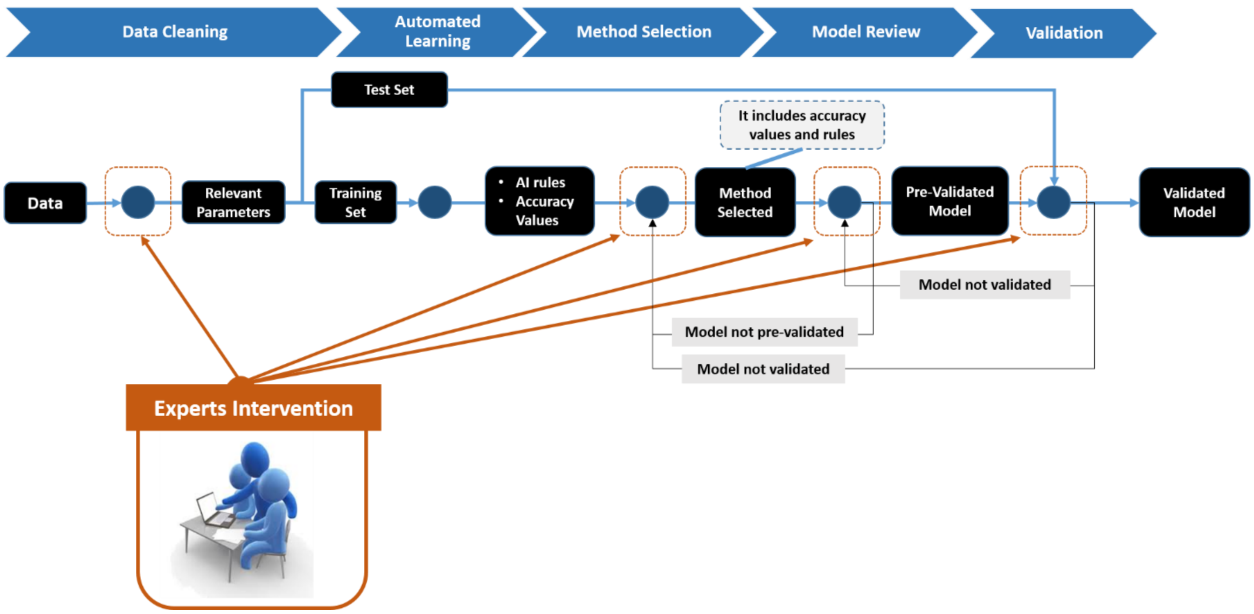


Figure 60. Knowledge capture process proposed

7.2.2. Knowledge life cycle management in engineering design

The advances in ICT and web-based technologies achieved in the last 20 years have led to the wider use knowledge repositories in the engineering design process. Data used in engineering design activities is often stored within knowledge models, thus facilitating the automatic execution of engineering tasks by KBE technology. Apart from the ability to automate tasks, some authors argue that KBE is capable of enabling an effective capture, retain and reuse of engineering knowledge [5]. Therefore, KBE is acknowledged by the research community as a potential solution to manage the complete knowledge life cycle of the engineering employed and produced in the engineering design process. However, KBE used in industrial applications is limited by [5] [6]:

- High development costs.
- Lack of KBE development methodology.
- Knowledge capture is a highly time consuming process for experts.
- Lack of an approach to measure the benefits provided by the implementation of KBE.

In the last decade, research efforts have been focus on providing a KBE development methodology. In this direction, MOKA and CommonKADS are the two prominent frameworks supporting KBE development. These methodologies have proved their ability to provide the methodological support required by KBE. However, there are still some limitations associated to them as presented in Table 27 and further described in section 4.2.

Table 27.Main limitations associated to MOKA and CommonKADS.

Methodology limitations	
MOKA	**CommonKADS**
Knowledge engineer oriented rather than end user	Not enough number of guidelines supporting the completion of activities.
Limited support for KBE maintenance and reuse	Limited support for KBE maintenance and reuse
Knowledge modelling support is limited	

Based on the limitations presented by the methodologies supporting the management aspect of KBE, [183] presents KNOMAD methodology aiming at delivering the benefits provided by MOKA and CommonKADS while addressing knowledge maintain and reuse.

The research reported here adopts KNOMAD methodology to support the development, maintain and update of an extended KBE development process aiming at the effective management of engineering knowledge. The methodological support required by KBE to effectively manage engineering knowledge is successfully provided by adopting KNOMAD. The effectiveness of the methodology adopted was validated through the completion of a questionnaire fulfilled by the experts and users involved in the development and use of the KSF. Therefore, the approach presented here tackles the identified KBE limitation associated to the lack of a generic methodology for KBE development.

Additionally, the knowledge sourcing methodology presented also tackles the KBE barrier linked to the need of a more efficient knowledge capture process through the integration of experts and AI within a knowledge-based framework as described in the previous section.

7.3. Key research limitations

In this section, limitations related to the realised research are described. Two types of limitations have been identified. The first group correspond to those limitations known prior to the realisation of the research. In this study, only one limitation has been classified within this group. This constraint exists as a result of selecting the qualitative approach as a research strategy. The research strategy followed makes it difficult to replicate the KSF methodology presented. In parallel, the limitations belonging to the second group were identified while performing the research. These limitations are related to the:

- **Validation of the knowledge sourcing framework.** As previously mentioned, no quantification of the performance of the KSF has been realised. This causes difficulties in proving the validity of the methodology. However, in order to minimise this, a qualitative validation process was performed, asking experts in the problem domain and users of the KSF to fill a questionnaire. There is a lack of a statistical analysis of the answers of the expert questionnaires due to the reduced number of engineers involved in the development and use of the knowledge sourcing capabilities developed in each use case.
- **Complexity of the KSF.** The modular approach followed in this research enables the KSF to grow quickly increasing its complexity, thus making more difficult the replication of the KSF presented in this work.
- **Industrial context of the case studies developed**. The two use cases described in this study have been implemented in different areas (design for manufacturing and lightning strike effects), nevertheless both cases belong to the aerospace industry. The application of a third use case within a different industrial context would have helped to improve the generality and applicability of the methodology proposed. However, it has not been realised due to experts and researcher time constraints.
- **Complexity of the model provided by the AI capability.** This limitation refers to the problems faced by experts when trying to understand the explicit model or

set of rules automatically created by the ML algorithms. More precisely, experts required a considerable amount of time to understand the explicit rules. This highlighted the need for more advanced visual analytical tools to speed up the process of understanding ML rules.

7.4. Future work

This section lists a number of research areas identified by the author where further work can be undertaken in order to enhance the source of engineering knowledge.

- **Knowledge representation and modelling.** A more intuitive way of modelling knowledge could be achieved by using knowledge representation formalisms such as semantic networks and conceptual graphs. By representing knowledge using these formalisms, various type of knowledge could be expressed apart from inference rules, thus providing better expressiveness and a deployable inference process.
- **ML Rules interpretation.** Further work is advised in order to facilitate the understanding of the knowledge automatically generated by ML methods. In this direction, more advanced visual analytical tools would be beneficial, providing the user with more intuitive formats of visualising the information related to the ML rules.
- **Rule validation process**. The framework developed includes an engineering workflow enabling the effective validation of the rules generated. However, the possibility of enhancing the defined workflow including features (such as rules' weighting and characterisation) would enhance the reliability of the KSF capability.
- **Data management**. The use of unstructured databases would allow the KSF to deal with larger amounts of data, with both structured and non-structured information such as data from manuals and emails respectively. This will provide experts with more relevant information to solve a specific problem.

7.5. Conclusions

In this research, a framework aiming at an extended and generic KBE methodology enabling the effective management of the complete knowledge life cycle and the efficient sourcing of knowledge was developed. The KSF development responds to a set of industrial needs identified by the author. The aspects of this PhD thesis related to the industrial motivations are listed as follows:

- **Reduce the time dedicated to Knowledge capture activities**. Literature highlights the knowledge elicitation task as the major bottle neck to boost the use of KBE systems seeking sourcing knowledge adequately. In this regard, a key feature included in this work is the use of AI applications in collaboration with experts to reduce the time required to model a problem. The research reported here facilitates the task of finding correlations in data in which knowledge is hidden, delivering a solution to those complex problems in which previous knowledge is limited.
- **Capitalise on efforts in the development of systems to source engineering knowledge**. Often, no profits are obtained from the large number of hours spent in developing systems focused on the elicitation and reuse of expert knowledge. In most of the cases, this is due to the way the knowledge is stored and structured. This makes knowledge gathered from previous experiences difficult to understand and transfer it within the organisation. In order to tackle this issue, the methodology proposed here stores knowledge in a human and computable format within independent knowledge packages, facilitating the knowledge transfer and its applicability across different engineering problems. For instance in the first case study, formal models contain knowledge in the format of IF-THEN rules. In doing so knowledge becomes more comprehensive while its automated execution using external applications becomes easier to overcome.
- **Increase reliability of AI methods in the engineering design context.** A majority of the existing machines learning algorithms fail to provide an explicit model (considered as "black box" techniques) that permits the users to trace

back the results. This constraint makes difficult for engineers to correlate the results provided by the AI algorithm to physical events. The use in this research of only those methods providing engineers with a set of rules describing the problem, not only allows users to trace back the results but also gain a better understanding of the problem. The reliability of AI techniques for knowledge capture is also enhanced by the intervention of experts in the problem domain at different stages of the knowledge sourcing procedure proposed.

To ensure the delivery of a KSF providing an effective solution to the challenges listed above, the following key activities were undertaken:

- Adoption of a state of the art methodology providing the knowledge sourcing capability with the required methodological support that enables the systematic capture, retain and reuse of expert knowledge.
- Use of AI techniques for fast knowledge capture, thus delivering a more efficient knowledge sourcing capability. In addition, this research integrates experts and AI methods within the knowledge sourcing process to generate more accurate and reliable predictions.
- Quantitative evaluation of the predictions provided in the learning and validation stages. This analysis was realised following an acknowledge criterion that facilitates the validation of the rules obtained as a result of the interaction between machine learning algorithms and experts.
- Qualitative assessment of the knowledge sourcing framework performed by a group of experts involved in the implementation of the case studies of this research. A questionnaire was created and experts were asked to complete it in order to facilitate the confirmation and validation of the assumptions made regarding the management aspect of the KSF.

The validation of the KSF was achieved through the successful implementation of two case studies aiming at optimising engineering design activities. The former case study – focus on the MFG cycle time prediction of CFRP wing covers– reduced the time required to obtain accurate time estimations from one week to a few minutes. The latter

–focus on the prediction of the pressure in the nut cap of a fastener assembly after it is hit by a lightning strike– provided a more comprehensive and realistic model of the problem in a similar amount of time when compared to the existing approach. Based on this, the KSF has proved to be a more efficient approach for knowledge sourcing. In parallel, KSF demonstrated to be effective methodology for the management of the complete knowledge life cycle. This became apparent through the implementation of both case studies following the steps of KNOMAD and the qualitative expert assessment of the framework.

References

[1] H. Zhu et al., “A Web-based Product Service System for aerospace maintenance, repair and overhaul services,” *Comput. Ind.*, vol. 63, pp. 338–348, May 2012.

[2] R. Stark et al., “Advanced Technologies in Life Cycle Engineering,” *Procedia CIRP*, vol. 22, pp. 3–14, 2014.

[3] Airbus, “Power8 prepares way for ‘New Airbus,’” 2007. [Online]. Available: http://www.airbus.com/presscentre/pressreleases/press-release-detail/detail/power8-prepares-way-for-new-airbus/. [Accessed: 05-May-2013].

[4] J. Wallin et al., “Developing PSS Concepts from Traditional Product Sales Situation: The Use of Business Model Canvas,” in *5th CIRP International Conference on Industrial Product-Service Systems*, 2013, pp. 263–274.

[5] G. La Rocca, “Knowledge based engineering: Between AI and CAD. Review of a language based technology to support engineering design,” *Adv. Eng. Informatics*, vol. 26, no. 2, pp. 159–179, Apr. 2012.

[6] W. J. C. Verhagen et al., “A critical review of Knowledge-Based Engineering: An identification of research challenges,” *Adv. Eng. Informatics*, vol. 26, no. 1, pp. 5–15, Jan. 2012.

[7] A. Lowe et al., “Characterising the requirements of engineering information systems,” *Int. J. Inf. Manage.*, vol. 24, no. 5, pp. 401–422, Oct. 2004.

[8] P. Bermell-Garcia, “A metamodel to annotate knowledge based engineering codes as enterprise knowledge resources,” 2007.

[9] P. Bermell-Garcia et al., “A framework for management of Knowledge-Based Engineering applications as software services: Enabling personalization and codification,” *Adv. Eng. Informatics*, vol. 26, no. 2, pp. 219–230, 2012.

[10] J. Liebowitz, “Knowledge management and its link to artificial intelligence,” *Expert Syst. Appl.*, vol. 20, no. 1, pp. 1–6, 2001.

[11] I. S. Fan et al., “A rule level knowledge management system for knowledge based engineering applications,” in *Proceedings of DETC’02. . ASME International 22nd Computers and Information in Engineering*, 2002.

[12] MOKA, “A Framework for Structuring and Representing Engineering Knowledge,” 2000. [Online]. Available: http://web1.eng.coventry.ac.uk/moka/kbe.htm. [Accessed: 24-Mar-2013].

[13] C. Robson, *Real world research*. Oxford: Oxford Blackwell, 2002.

[14] E. Gummesson, *Qualitative Methods in Management Research*. USA: Sage Publications, 1991.

[15] G. Pahl and W. Beitz, *Engineering Design. A Systematic Approach.* Springer-Verlag, 1996.

[16] S. Pugh, *Total design: Integrated methods for successful product engineering.* Addison Wesley Publishing Company., 1991.

[17] V. Hubka and W. E. Eder, *Theory of technical systems: a total concept theory for engineering design.* Springer-Verlag, 1988.

[18] E. Dieter, *Engineering Design*. New York: McGraw-Hill, 2000.

[19] C. Hales, *Managing Engineering Design*. Longman Scientific & Technical, 1993.

[20] G. Pahl and W. Beitz, “Engineering Design,” *Des. Counc.*, 1984.

[21] L. A. Mendes et al., “Designing automated test systems: An adapted methodology inspired on Pahl and Beitz ’ s systematic approach,” *Robot. Comput. Integr. Manuf.*, vol. 25, no. 6, pp. 945–950, 2009.

[22] M. A. Robinson, “How design engineers spend their time: Job content and task satisfaction,” *Des. Stud.*, vol. 33, no. 4, pp. 391–425, 2012.

[23] C. Hales, *Analysis of the engineering design process in an industrial context*, Unpublishe. University of Cambridge, UK, 1986.

[24] Y. Chen et al., “On technology for functional requirement-based reference design retrieval in engineering knowledge management,” vol. 44, pp. 798–816, 2008.

[25] R. Marsh, *The capture and utilisation of experience in engineering design*, Unpublishe. University of Cambridge, UK, 1997.

[26] M. Hoegl et al., “How to Support Knowledge Creation in New Product Development: An Investigation of Knowledge Management Methods,” *Eur. Manag. J.*, vol. 23, no. 3, pp. 263–273, 2005.

[27] J.-J. Huang, “The evolutionary perspective of knowledge creation – A mathematical representation,” *Knowledge-Based Syst.*, vol. 22, no. 6, pp. 430–438, Aug. 2009.

[28] C. Chapman and M. Pinfold, “Design engineering—a need to rethink the solution using knowledge based engineering,” *Knowledge-Based Syst.*, vol. 12, no. 5–6, pp. 257–267, Oct. 1999.

[29] J. S. Gero, *The situated function-behaviour - structure framework Design studies*. Vol 25 No 4 pp 373 - 391, 2004.

[30] W. Browne et al., “Knowledge-elicitation and data-mining: Fusing human and industrial plant information,” vol. 19, pp. 345–359, 2006.

[31] H. Soltan, *Knowledge management lecture notes*. Cranfield University, UK, 2006.

[32] G. Schreiber et al., *Knowledge Engineering and management: the CommonKADS methodology*. USA: MIT Press, 2000.

[33] M. Curran et al., “Attitudes to expert systems: A card sort study,” *foot J.*, vol. 15, no. 4, pp. 190–197, 2005.

[34] R. Mammen et al., “New approaches to qualitative interviewing: Development of a card sort technique to understand subjective patterns of symptoms and responses,” *Int. J. Nurs. Stud.*, 2016.

[35] R. Hoffman, “The problem of extracting the knowledge of experts from the perspective of experimental psychology,” *AI Magazine*, pp. 53–67, 1987.

[36] F. Todhunter, “Using Concurrent Think-Aloud and Protocol Analysis to Explore Student Nurses’ Social Learning Information Communication Technology Knowledge and Skill Development,” *Nurse Educ. Today*, vol. 35, no. 6, pp. 815–822, 2015.

[37] D. R. Deeter-Schmelz et al., “Understanding sales manager effectiveness — Linking attributes to sales force values.,” *Ind. Mark. Manag.*, vol. 31, no. 7, pp. 617–626, 2002.

[38] T. Gruber et al., “The desired qualities of customer contact employees in complaint handling encounters,” *J. Mark. Manag.*, vol. 22, no. 5–6, pp. 619–642, 2006.

[39] K. Grunert and T. Bech-Larsen, “Explaining choice option attractiveness by beliefs elicited by the laddering method,” *J. Econ. Psychol.*, vol. 26, no. 2, pp. 223–241, 2005.

[40] C. Chen et al., “A strategy for acquiring customer requirement patterns using laddering technique and ART2 neural network,” *Adv. Eng. Informatics*, vol. 16, no. 3, pp. 229–240, 2002.

[41] M. L. Markus, “Toward a theory of knowledge reuse: types of knowledge reuse situations and factors in reuse success,” *J. Manag. Inf. Syst.*, vol. 18, no. 1, pp. 57–93, 2001.

[42] R. H. Weise, *Representing the corporation: strategies for legal counsel,*. USA: Aspen Law & Business Publishers, 1996.

[43] C. K. Kwong et al., "A GA based knowledge discovery system for the design of fluid dispensing processes for electronic packaging," *Expert Syst. Appl.*, vol. 36, pp. 3829–3838, 2009.

[44] P. Wriggers et al., "Intelligent support of engineering analysis using ontology and case-based reasoning," *Eng. Appl. Artif. Intell.*, vol. 20, no. 5, pp. 709–720, 2007.

[45] T. Kocsis et al., "Case-Based Reasoning system for mathematical modelling options and resolution methods for production scheduling problems: Case representation, acquisition and retrieval," *Comput. Ind. Eng.*, vol. 77, pp. 46–64, 2014.

[46] Y. Guo et al., "Research on CBR system based on data mining," *Appl. Soft Comput. J.*, vol. 11, no. 8, pp. 5006–5014, 2011.

[47] T. M. Connollya et al., "A systematic literature review of empirical evidence on computer games and serious games," *Comput. Educ.*, vol. 59, no. 2, pp. 661–686, 2012.

[48] T. Lainema and O. P. Hilmola, "Learn more, better and faster: computer-based simulation gaming of production and operations," *Int. J. Bus. Perform. Manag.*, vol. 7, no. 1, pp. 34–59, 2005.

[49] M. A. Lewis and H. R. Maylor, "Game playing and operations management education," *Int. J. Prod. Econ.*, vol. 105, pp. 134–149, 2007.

[50] D. Battini et al., "Logistic Game: learning by doing and knowledge-sharing," *Prod. Plan. Control*, vol. 20, no. 8, pp. 724–736, 2009.

[51] H. H. Adelsberger et al., "A Simulation Game Approach for Efficient Education in Enterprise Resource Planning Systems," in *Proceedings of the 13th European Simulation Multiconference*, 1999, pp. 454–460.

[52] T. Carron and S. Talbot, "Industrial Knowledge Assessment for Capitalization Purpose via Learning Techniques," *World Appl. Sci. J.*, vol. 31, pp. 28–35, 2014.

[53] F. S. C. Tseng and C.-L. Chen, "Extending the UML concepts to transform natural language queries with fuzzy semantics into SQL," *Inf. Softw. Technol.*, vol. 48, no. 9, pp. 901–914, Sep. 2006.

[54] D. Baxter et al., "An Engineering Design Knowledge Reuse Methodology Using Process Modelling 1 Introduction," vol. 18, no. 1, pp. 37–48, 2007.

[55] C. A. Costa et al., “Assessment of a Product Range Model concept to support design reuse using rule based systems and case based reasoning,” *Adv. Eng. Informatics*, vol. 26, no. 2, pp. 292–305, Apr. 2012.

[56] C. H. Y. Lam et al., “A hybrid case-GA-based decision support model for warehouse operation in fulfilling cross-border orders,” *Expert Syst. Appl.*, vol. 39, no. 8, pp. 7015–7028, Jun. 2012.

[57] J. Owen et al., “Knowledge reuse and transfer in a project management environment,” *J. Inf. Technol. Cases Appl. Res.*, 2004.

[58] D. Baxter et al., “A framework to integrate design knowledge reuse and requirements management in engineering design,” vol. 24, pp. 585–593, 2008.

[59] G. F. Luger, *Artificial Intelligence: Structures and Strategies for Complex Problem Solving*. Addison-Wesley Publishing Company, 1999.

[60] N. Deshpande, *Artificial Intelligence*. Technical publications, 2009.

[61] R. E. Bellman, *An introduction to artificial intelligence: Can computers think?* USA: Boyd & Fraser Pub. Co, 1978.

[62] R. Kurzweil, *The age of intelligent machines*. MIT Press, 1990.

[63] R. J. Schalkoff, *Artificial intelligence: An engineering approach*. New York: McGraw-Hill, 1990.

[64] M. Negnevitsky, *Artificial Intelligence: A Guide To Intelligent Systems.* Addison Wesley Longman, 2001.

[65] E. Tsui et al., “The role of artificial intelligence in knowledge management,” vol. 13, pp. 235–239, 2000.

[66] C. F. Cheung et al., “A multi-faceted and automatic knowledge elicitation system (MAKES) for managing unstructured information,” *Expert Syst. Appl.*, vol. 38, no. 5, pp. 5245–5258, May 2011.

[67] H. C. W. Lau et al., “Development of a process mining system for supporting knowledge discovery in a supply chain network,” *Int. J. Prod. Econ.*, vol. 122, no. 1, pp. 176–187, Nov. 2009.

[68] C. Palmer et al., “Generating rules from data mining for collaboration moderator services,” *J. Intell. Manuf.*, vol. 24, no. 2, pp. 313–330, Sep. 2011.

[69] P. Potes Ruiz et al., “Generating knowledge in maintenance from Experience Feedback,” *Knowledge-Based Syst.*, vol. 68, pp. 4–20, 2014.

[70] J.-B. Yang, "A rule induction-based knowledge system for retaining wall selection," *Expert Syst. Appl.*, vol. 23, no. 3, pp. 273–279, Oct. 2002.

[71] R. D. Sriram, *Distributed and integrated collaborative engineering design*. Sarvan Publishers, 2002.

[72] L. Wang et al., "Collaborative conceptual design — state of art and future trends," *Comput. Des.*, vol. 34, no. 13, pp. 981–996, 2002.

[73] M. Aurisicchio et al., "EVALUATION OF HOW DRED DESIGN RATIONALE IS INTERPRETED," no. August, pp. 1–12, 2007.

[74] R. L. Nagel et al., "SIGNAL FLOW GRAMMAR FROM THE FUNCTIONAL BASIS," no. August, pp. 1–10, 2007.

[75] S. K. Chandrasegaran et al., "Computer-Aided Design The evolution , challenges , and future of knowledge representation in product design systems," *Comput. Des.*, vol. 45, no. 2, pp. 204–228, 2013.

[76] J. J. Shah et al., "A discourse on geometric feature recognition from cad models," *J. Comput. Inf. Sci. Eng.*, vol. 1, no. 1, pp. 41–51, 2001.

[77] G. Turner and D. Anderson, "An object-oriented approach to interactive, feature-based design for quick turnaround manufacturing," in *Proceedings of the 1988 ASME international computers in engineering conference*, 1988, pp. 551–6.

[78] D. Anderson and T. Chang, "Geometric reasoning in feature based design and process planning," *Comput. Graph. An Int. J.*, vol. 14, no. 2, pp. 225–235, 1990.

[79] K. H. Ko et al., "An embedded system for knowledgebased cost evaluation of molded parts," *Knowledge-Based Syst.*, vol. 20, pp. 291–299, 2007.

[80] J. W. Choi et al., "Knowledge-based engineering system to estimate manufacturing cost for composite structures," *J. Aircr.*, vol. 42, pp. 1396–1402, 2005.

[81] M. Pinfold and C. Chapman, "The application of KBE techniques to the FE model creation of an automotive body structure," *Comput. Ind.*, vol. 44, no. 1, pp. 1–10, Jan. 2001.

[82] W. Skarka, "Application of MOKA methodology in generative model creation using CATIA," *Eng. Appl. Artif. Intell.*, vol. 20, no. 5, pp. 677–690, Aug. 2007.

[83] D. J. Allsopp et al., "A database architecture for reusable CommonKADS agent specification components," *Knowledge-Based Syst.*, vol. 15, no. 5–6, pp. 275–283, 2002.

[84] R. Curran et al., “A multidisciplinary implementation methodology for knowledge based engineering: KNOMAD,” *Expert Syst. Appl.*, vol. 37, no. 11, pp. 7336–7350, Nov. 2010.

[85] N. Shukla et al., “Algorithm portfolios for logistics optimization considering stochastic demands and mobility allowance,” *Int. J. Prod. Econ.*, vol. 141, no. 1, pp. 146–166, 2013.

[86] S. R. Yadav et al., “An algorithm portfolio based solution methodology to solve a supply chain optimization problem,” *Expert Syst. Appl.*, vol. 36, no. 4, pp. 8407–8420, 2009.

[87] T. Segaran, *Programming Collective Intelligence: Building Smart Web 2.0 Applications*, 1st editio. O’Reilly Media, 2007.

[88] I. Witten et al., *Data Mining: Practical Machine Learning Tools and Techniques*, 3rd Editio. Morgan Kaufmann, 2011.

[89] B. Dolšak and M. Novak, “Intelligent decision support for structural design analysis,” *Adv. Eng. Informatics*, vol. 25, no. 2, pp. 330–340, Apr. 2011.

[90] H. Dettmar et al., “Knowledge-based data generation,” *Knowledge-Based Syst.*, vol. 11, no. 3–4, pp. 167–177, Nov. 1998.

[91] T. W. Liao et al., “An integrated database and expert system for failure mechanism identification: Part I, automated knowledge acquisition,” vol. 6, 1999.

[92] F. Alonso et al., “Cooperation between expert knowledge and data mining discovered knowledge: Lessons learned,” *Expert Syst. Appl.*, vol. 39, no. 8, pp. 7524–7535, Jun. 2012.

[93] D. Steenhuizen and M. van Tooren, “The implementation of a knowledge-based framework for the aerodynamic optimization of a morphing wing device,” *Adv. Eng. Informatics*, vol. 26, no. 2, pp. 207–218, Apr. 2012.

[94] H. Z. Yang et al., “Implementation of knowledge-based engineering methodology in ship structural design,” *Comput. Des.*, vol. 44, no. 3, pp. 196–202, Mar. 2012.

[95] A. Kochan, “Jaguar uses knowledge-based tools to reduce model development times,” *J. Assem. Autom.*, 1999.

[96] K. Amadori et al., “Flexible and robust CAD models for design automation,” *Adv. Eng. Informatics*, vol. 26, no. 2, pp. 180–195, Apr. 2012.

[97] Z. Lou et al., “Development of an integrated knowledge-based system for mold-base design,” *J. Mater. Process. Technol.*, vol. 150, no. 1–2, pp. 194–199, Jul. 2004.

[98] V. Naranje and S. Kumar, “A knowledge based system for automated design of deep drawing die for axisymmetric parts,” *Expert Syst. Appl.*, vol. 41, no. 4, pp. 1419–1431, Mar. 2014.

[99] D. Xie et al., “Object-oriented knowledge framework for modelling human mastication of foods,” *Expert Syst. Appl.*, vol. 36, no. 3, pp. 4810–4821, 2009.

[100] Y.-H. Wu and H.-J. Shaw, “Document based knowledge base engineering method for ship basic design,” *Ocean Eng.*, vol. 38, no. 13, pp. 1508–1521, Sep. 2011.

[101] J. R. Corney et al., “Putting the crowd to work in a knowledge-based factory,” *Adv. Eng. Informatics*, vol. 24, no. 3, pp. 243–250, Aug. 2010.

[102] S. Kumar and R. Singh, “An intelligent system for selection of die-set of metal stamping press tool,” *J. Mater. Process. Technol.*, vol. 164–165, pp. 1395–1401, May 2005.

[103] S. Kumar and R. Singh, “An automated design system for progressive die,” *Expert Syst. Appl.*, vol. 38, no. 4, pp. 4482–4489, Apr. 2011.

[104] G. Colombo et al., “Towards the design of intelligent CAD systems: An ontological approach,” *Adv. Eng. Informatics*, vol. 21, no. 2, pp. 153–168, Apr. 2007.

[105] C. B. Chapman and M. Pinfold, “The application of a knowledge based engineering approach to the rapid design and analysis of an automotive structure,” *Adv. Eng. Softw.*, vol. 32, no. 12, pp. 903–912, Dec. 2001.

[106] J.-W. Choi, “Architecture of a knowledge based engineering system for weight and cost estimation for a composite airplane structures,” *Expert Syst. Appl.*, vol. 36, no. 8, pp. 10828–10836, Oct. 2009.

[107] W.-C. Chen et al., “A systematic optimization approach for assembly sequence planning using Taguchi method, DOE, and BPNN,” *Expert Syst. Appl.*, vol. 37, no. 1, pp. 716–726, Jan. 2010.

[108] G. La Rocca, “Advanced Engineering Informatics Knowledge based engineering : Between AI and CAD . Review of a language based technology to support engineering design,” *Adv. Eng. Informatics*, vol. 26, no. 2, pp. 159–179, 2012.

[109] C. van der Velden et al., “Adaptable methodology for automation application development,” *Adv. Eng. Informatics*, vol. 26, no. 2, pp. 231–250, Apr. 2012.

[110] J. Nacsa et al., “Knowledge management support for machine tool designers using expert enablers,” *Int. J. Comput. Integr. Manuf.*, vol. 18, no. 7, pp. 561–571, Oct. 2005.

[111] N. Gardan and Y. Gardan, “An application of knowledge based modelling using scripts,” *Expert Syst. Appl.*, vol. 25, no. 4, pp. 555–568, Nov. 2003.

[112] R. Utpal, “An Intelligent interface between symbolic and numeric analysis tools required for the development of an integrated CAD system,” *Comput. Ind. Eng.*, vol. 30, no. 1, pp. 13–26, 1996.

[113] L. Sheremetov et al., “Knowledge-based collaborative engineering of pipe networks in the upstream and downstream petroleum industry,” *Comput. Ind.*, vol. 59, no. 9, pp. 936–948, Dec. 2008.

[114] R. M. C. Ratnayake, “KBE development for criticality classification of mechanical equipment: A fuzzy expert system,” *Int. J. Disaster Risk Reduct.*, vol. 9, pp. 84–98, Sep. 2014.

[115] D. Monticolo et al., “An agent-based system to build project memories during engineering projects,” *Knowledge-Based Syst.*, vol. 33, Jan. 2014.

[116] I. Jagielska et al., “An investigation into the application of neural networks, fuzzy logic, genetic algorithms, and rough sets to automated knowledge acquisition for classification problems,” *Neurocomputing*, vol. 24, no. 1–3, pp. 37–54, Feb. 1999.

[117] M. Sawhney et al., “Collaborating to create: The Internet as a platform for customer engagement in product innovation,” *J. Interact. Mark.*, vol. 19, no. 4, pp. 4–17, Jan. 2005.

[118] D. Lee, “Knowledge-based system for safety control of damaged ship,” *Knowledge-Based Syst.*, vol. 19, no. 3, pp. 187–191, Jul. 2006.

[119] A. Corallo et al., “Enhancing product development through knowledge-based engineering (KBE): A case study in the aerospace industry,” *J. Manuf. Technol. Manag.*, vol. 20, no. 8, pp. 1070–1083, 2009.

[120] S. Kumar, “A knowledge based reliability engineering approach to manage product safety and recalls,” *Expert Syst. Appl.*, vol. 41, no. 11, pp. 5323–5339, Sep. 2014.

[121] R. Hunter et al., “A functional approach for the formalization of the fixture design process,” *Int. J. Mach. Tools Manuf.*, vol. 46, no. 6, pp. 683–697, May 2006.

[122] J.-W. Choi et al., “A knowledge-based engineering tool to estimate cost and weight of composite aerospace structures at the conceptual stage of the design process,” *Aircr. Eng. Aerosp. Technol.*, vol. 79, no. 5, pp. 459–468, 2007.

[123] I. Mendikoa et al., “Distributed Product Design and Manufacturing Based on KBE,” in *Computer Supported Cooperative Work in Design II*, vol. 3865, Springer, 2006, pp. 404–413.

[124] M. Ruschitzka et al., “Ontology-Based Approach in Hybrid Engineering Knowledge Representation for Stamping Die Design,” in *New World Situation: New Directions in*

Concurrent Engineering, J. Pokojski, S. Fukuda, and J. Salwiński, Eds. London: Springer London, 2010, pp. 205–212.

[125] E. Shehab and H. Abdalla, “An Intelligent Knowledge-Based System for Product Cost Modelling,” *Int. J. Adv. Manuf. Technol.*, vol. 19, no. 1, pp. 49–65, Jan. 2002.

[126] Y.-E. Nahm and H. Ishikawa, “A new 3D-CAD system for set-based parametric design,” *Int. J. Adv. Manuf. Technol.*, vol. 29, pp. 137–150, Apr. 2006.

[127] S. Kumar and R. Singh, “A short note on an intelligent system for selection of materials for progressive die components,” *J. Mater. Process. Technol.*, vol. 182, no. 1–3, pp. 456–461, Feb. 2007.

[128] M. İpek et al., “An expert system based material selection approach to manufacturing,” *Mater. Des.*, vol. 47, pp. 331–340, May 2013.

[129] F. Elgh and M. Cederfeldt, “Documentation and Management of Product Knowledge in a System for Automated Variant Design – A Case Study,” in *New World Situation: New Directions in Concurrent Engineering, Proceedings of the 17th Conference on Concurrent Engineering*, 2010, pp. 237–245.

[130] F. Elgh, “Supporting Management and Analysis of Quotations in a Design Automation Approach to Customization,” in *New World Situation: New Directions in Concurrent Engineering, Proceedings of the 17th Conference on Concurrent Engineering*, 2010, pp. 401–409.

[131] P. Modeling et al., “Parametric Modelling of movables for structural analysis,” *J. Aircr.*, vol. 42, no. April, pp. 1606–1614, 2004.

[132] J. Kulon et al., “Applying knowledge-based engineering to traditional manufacturing design,” *Int. J. Adv. Manuf. Technol.*, vol. 30, no. 9–10, pp. 945–951, Dec. 2005.

[133] J. Kulon et al., “A knowledge-based engineering design tool for metal forging,” *J. Mater. Process. Technol.*, vol. 177, no. 1–3, pp. 331–335, Jul. 2006.

[134] E. M. Shehab and H. S. Abdalla, “A Cost-Effective Knowledge-Based Reasoning System for Design for Automation,” *Proc. Inst. Mech. Eng. Part B J. Eng. Manuf.*, vol. 220, no. 5, pp. 729–743, Jan. 2006.

[135] S. Van Der Elst and M. Van Tooren, “Application of a Knowledge Engineering Process to Support Engineering Design Application Development,” in *Collaborative Product and Service Life Cycle Management for a Sustainable World, Proceedings of the 15th Conference on Advanced Concurrent Engineering*, 2008.

[136] J. Cui and D. Wang, “Application of knowledge-based engineering in ship structural design and optimization,” *Ocean Eng.*, vol. 72, pp. 124–139, Nov. 2013.

[137] W.-R. Jong et al., “Application of knowledge-based engineering for automated slide design,” *Int. J. Adv. Manuf. Technol.*, Jun. 2014.

[138] G. N. Koini et al., “A software tool for parametric design of turbomachinery blades,” *Adv. Eng. Softw.*, vol. 40, no. 1, pp. 41–51, Jan. 2009.

[139] V. Naranje and S. Kumar, “A Knowledge Based System for Selection of Components of Deep Drawing Die,” *Am. J. Intell. Syst.*, vol. 2, no. 2, pp. 1–11, 2012.

[140] K. H. Ko et al., “An embedded system for knowledge-based cost evaluation of molded parts,” *Knowledge-Based Syst.*, vol. 20, no. 3, pp. 291–299, Apr. 2007.

[141] J. Choi, “Architecture of a knowledge based engineering system for weight and cost estimation for a composite airplane structures,” *Expert Syst. Appl.*, vol. 36, no. 8, pp. 10828–10836, 2009.

[142] C. L. Emberey and N. R. Milton, “Application of Knowledge Engineering Methodologies to Support Engineering Design Application Development in Aerospace,” in *7th AIAA Aviation Technology, Integration and Operations Conference (ATIO)*, 2007, pp. 18–20.

[143] S. Kumar and R. Singh, “A knowledge-based system to automate the selection of progressive die components,” *Int. J. Comput. Mater. Sci. Surf. Eng.*, vol. 1, pp. 85–96, 2007.

[144] S. Kumar and R. Singh, “A low cost knowledge base system framework for progressive die design,” *J. Mater. Process. Technol.*, vol. 153–154, pp. 958–964, 2004.

[145] L. Pomponio and M. Le Goc, “Reducing the Gap between Experts’ Knowledge and Data: The TOM4D Methodology,” *Data Knowl. Eng.*, vol. 94, pp. 1–37, Jul. 2014.

[146] V. Allee, “The Future of Knowledge: Increasing Prosperity through Value Networks,” *Oxford: Butterworth-Heinemann*, 2002.

[147] L. C. M. Tang et al., “Codification vs personalisation: A study of the information evaluation practice between aerospace and construction industries,” *Int. J. Inf. Manage.*, vol. 30, no. 4, pp. 315–325, Aug. 2010.

[148] R. R. WJC Verhagen, P Bermell-Garcia, P Mariot, JP Cotton, D Ruiz, “Knowledge-based cost modelling of composite wing structures,” *Int. J. Comput. Integr. Manuf.*, vol. 25, no. 4–5, pp. 368–383, 2012.

[149] G. Gonzalez Lozano, “Knowledge extraction to support wing conceptual design evaluation,” Msc Thesis, Cranfield University, 2012.

[150] Y. Abu-Mostafa and M. Magdon-Ismail, *Learning From Data*. AMLBook, 2012.

[151] M. Bishop, *Pattern Recognition and Machine Learning*. Springer, 2007.

[152] K. Murphy, *Machine Learning - A Probabilistic Perspective*. MIT Press, 2012.

[153] K. Kourou et al., "Machine learning applications in cancer prognosis and prediction," *Comput. Struct. Biotechnol. J.*, vol. 13, pp. 8–17, 2015.

[154] Q. R. Hashim and C. Roy, "A systematic extreme learning machine approach to analyze visitors′ thermal comfort at a public urban space," *Renew. Sustain. Energy Rev.*, vol. 58, pp. 751–760, 2016.

[155] M. Hall et al., "The WEKA Data Mining Software: An Update," *SIGKDD Explor.*, 2009.

[156] Pedregosa et al., "Scikit-learn: Machine Learning in Python," *JMLR 12*, 2011.

[157] G. K. F. Tso and K. K. W. Yau, "Predicting electricity energy consumption: A comparison of regression analysis, decision tree and neural networks," *Energy*, vol. 32, no. 9, pp. 1761–1768, Sep. 2007.

[158] Z. Elouedi et al., "Belief decision trees: theoretical foundations," *Int. J. Approx. Reason.*, vol. 28, pp. 91–124, 2001.

[159] S. Portnoy and R. Koenker, "The Gaussian hare and the Laplacian tortoise: computability of squared-error versus absolute-error estimators," *Stat. Sci.*, vol. 12, no. 4, pp. 279 – 300, 1997.

[160] X. Chen and H. Ishwaran, "Random forests for genomic data analysis," *Genomics*, vol. 99, no. 6, pp. 323–9, Jun. 2012.

[161] H. Kang et al., "Senti-lexicon and improved Naïve Bayes algorithms for sentiment analysis of restaurant reviews," *Expert Syst. Appl.*, vol. 39, no. 5, pp. 6000–6010, Apr. 2012.

[162] A. Pérez et al., "Bayesian classifiers based on kernel density estimation: Flexible classifiers," *Int. J. Approx. Reason.*, vol. 50, no. 2, pp. 341–362, Feb. 2009.

[163] A. Gusnanto et al., "Partial least squares and logistic regression random-effects estimates for gene selection in supervised classification of gene expression data.," *J. Biomed. Inform.*, vol. 46, no. 4, pp. 697–709, Aug. 2013.

[164] S. Dreiseitl and L. Ohno-Machado, "Logistic regression and artificial neural network classification models: a methodology review," *J. Biomed. Inform.*, vol. 35, no. 5–6, pp. 352–359, Oct. 2002.

[165] G. Mountrakis et al., "Support vector machines in remote sensing: A review," *ISPRS J. Photogramm. Remote Sens.*, vol. 66, no. 3, pp. 247–259, May 2011.

[166] N. Barakat and A. P. Bradley, “Neurocomputing Rule extraction from support vector machines: A review,” *Neurocomputing*, vol. 74, no. 1–3, pp. 178–190, 2010.

[167] M. Koç and A. Barkana, “Application of Linear Regression Classification to low-dimensional datasets,” *Neurocomputing*, vol. 131, pp. 331–335, May 2014.

[168] D. Mackay, “Introduction to gaussian processes,” in *Neural Networks and Machine Learning*, C. M. Bishop, Ed. Berlin: Springer, 1998, pp. 133–165.

[169] G. Holmes et al., “Generating Rule Sets from Model Trees,” in *Advanced Topics in Artificial Intelligence*, vol. 1747, 1999, pp. 1 – 12.

[170] R. Quinlan, “Learning with continuous classes,” in *Proceedings of the 5th Australian Joint Conference On Artificial Intelligence*, 1992, pp. 343 – 348.

[171] Y. Wang and I.Witten, “Induction of model trees for predicting continuous classes,” in *European Conference on Machine Learning Poster Papers*, 1997, pp. 128 – 137.

[172] M. Paliwal and U. a. Kumar, “Neural networks and statistical techniques: A review of applications,” *Expert Syst. Appl.*, vol. 36, no. 1, pp. 2–17, Jan. 2009.

[173] J. Kacprzyk and W. Pedrycz, *Springer Handbook of Computational Intelligence*. Springer-Verlag Berlin Heidelberg, 2015.

[174] K. erano, T., Sugeno, M., Mukaido, M., Shigemasu, *Fuzzy Engineering toward Human Friendly Systems*. 1992.

[175] D. Freedman, “Statistical Models: Theory and Practice,” 2009.

[176] Y. Xin, *Linear Regression Analysis: Theory and Computing*. World Scientific, 2009.

[177] L. Zadeh, “Fuzzy Sets,” *Inf. Control*, vol. 8, no. 3, pp. 338–353, 1965.

[178] Z. Chi et al., “Fuzzy Algorithms with Applications to Image Processing and Pattern Recognition,” *World Scientific*, 1996.

[179] H. Ishibuchi et al., *Classification and Modeling with Linguistic Information Granules: Advanced Approaches to Linguistic Data Mining*. Springer-Verlag, 2005.

[180] R. Babuska, “Fuzzy Modeling for Control,” *Kluwer Academic Press*, 1998.

[181] B. Kompare et al., “Prediction of rainfall runoff from catchment by intelligent data analysis with machine learning tools within the artificial intelligence tools,” *Acta hydrotechnica*, p. 16/16, 1997.

[182] D. P. Solomatine and K. N. Dulal, “Model tree as an alternative to neural network in rainfall-runoff modelling,” *Hydrol. Sci.*, vol. 48, no. 3, pp. 399–411, 2003.

[183] C. Verhagen, “An Ontology-Based Approach for Knowledge Lifecycle Management within Aircraft Lifecycle Phases,” 2013.

[184] M. Stokes, *Managing engineering knowledge - MOKA: Methodology for Knowledge Based Engineering Applications*. Professional Engineering Publishing Limited, 2001.

[185] R. Curran et al., “Expert Systems with Applications A multidisciplinary implementation methodology for knowledge based engineering: KNOMAD,” *Expert Syst. Appl.*, vol. 37, no. 11, pp. 7336–7350, 2010.

[186] V. Challagulla et al., “Empirical assessment of machine learning based software defect prediction techniques,” in *10th IEEE international workshop on object-oriented real-time dependable systems, Sedona*, 2005, pp. 263–70.

[187] P. Guan et al., “Adaptive fuzzy sliding mode control for flexible satellite,” *Eng. Appl. Artif. Intell.*, 2005.

[188] M. VijayaKumar et al., “A direct adaptive neural command controller design for an unstable helicopter,” *Appl. Artif. Intell.*, 2009.

[189] F. Cantu-Ortiz, “Advancing artificial intelligence research and dissemination through conference series: Benchmark, scientific impact and the MICAI experience,” *J. Expert Syst. with Appl.*, 2014.

[190] E. F. Simas Filho et al., “Decision support system for ultrasound inspection of fiber metal laminates using statistical signal processing and neural networks.,” *Ultrasonics*, vol. 53, no. 6, pp. 1104–11, Aug. 2013.

[191] Y.-C. Tang, “An approach to budget allocation for an aerospace company—Fuzzy analytic hierarchy process and artificial neural network,” *Neurocomputing*, vol. 72, no. 16–18, pp. 3477–3489, Oct. 2009.

[192] M. a. Zaidan et al., “Bayesian Hierarchical Models for aerospace gas turbine engine prognostics,” *Expert Syst. Appl.*, vol. 42, no. 1, pp. 539–553, Jan. 2015.

[193] F. J. de Cos Juez et al., “Analysis of lead times of metallic components in the aerospace industry through a supported vector machine model,” *Math. Comput. Model.*, vol. 52, no. 7–8, pp. 1177–1184, Oct. 2010.

[194] C. Feldmann, *The Practical Guide to Business Process Reengineering Using Idef0*. Dorset House Publishing Co, 1998.

[195] P. Tüfekci, “Prediction of full load electrical power output of a base load operated combined cycle power plant using machine learning methods,” *Int. J. Electr. Power Energy Syst.*, 2014.

[196] K. Kavaklioglu, “Modeling and prediction of Turkey’s electricity consumption using Support Vector Regression,” *Appl. Energy*, 2011.

[197] R. Mlynarski et al., “A new method of data preparation for cardiological decision support.,” in *Computers in Cardiology*, 2006, pp. 273–276.

[198] J. Alcalá-Fdez et al., “KEEL: A software tool to assess evolutionary algorithms for data mining problems.,” *Soft Comput. Fusion Found. Methodol. Appl.*, vol. 13, no. 3, pp. 307–318, 2009.

[199] a. M. M. S. Ullah and K. H. Harib, “A human-assisted knowledge extraction method for machining operations,” *Adv. Eng. Informatics*, vol. 20, no. 4, pp. 335–350, Oct. 2006.

[200] U. Roy, “An Intelligent interface between symbolic and numeric analysis tools required for the development of an integrated CAD system,” *Comput. Ind.*, vol. 30, no. 1, pp. 13–26, 1996.

[201] KIM, “Knowledge and Information Management Through Life,” 2007. [Online]. Available: http://www-edc.eng.cam.ac.uk/kim/. [Accessed: 15-Jul-2014].

[202] P. Taylor et al., “Knowledge management in engineering design: personalization and codification,” no. February 2012, pp. 37–41, 2007.

[203] G. L. S. Cooper, I.S. Fan, “Achieving Competitive Advantage Through Knowledge Based Engineering – A Best Practice Guide, Cranfield University, Bedford, UK,” 2001.

[204] M. Stickdorn and J. Schneider, *This is service design thinking*. Uitgeverij, 1 edition, 2012.

[205] S. Wang and R. M. Summers, “Machine learning and radiology,” *Med. Image Anal.*, vol. 16, no. 5, pp. 933–951, 2012.

[206] A. Khodayari-rostamabad et al., “Clinical Neurophysiology A machine learning approach using EEG data to predict response to SSRI treatment for major depressive disorder,” *Clin. Neurophysiol.*, vol. 124, no. 10, pp. 1975–1985, 2013.

[207] C. Catal, “Expert Systems with Applications Software fault prediction: A literature review and current trends,” *Expert Syst. Appl.*, vol. 38, no. 4, pp. 4626–4636, 2011.

[208] K. O. Elish and M. O. Elish, “Predicting defect-prone software modules using support vector machines,” *J. Syst. Softw.*, 2008.

[209] G. Meyfroidt, “Machine learning techniques to examine large patient databases,” *J. Clin. Neurophysiol.*, 2013.

[210] S. Ganzert et al., “Analysis of respiratory pressure-volume curves in intensive care medicine using inductive machine learning,” *Artif. Intell. Med.*, vol. 26, no. 1–2, pp. 69–86, 2002.

[211] D. Martens et al., “Comprehensible credit scoring models using rule extraction from support vector machines,” *Eur. J. Oper. Res.*, vol. 183, no. 3, pp. 1466–1476, 2007.

[212] M. A. H. Farquad et al., “Expert Systems with Applications Support vector regression based hybrid rule extraction methods for forecasting,” *Expert Syst. Appl.*, vol. 37, no. 8, pp. 5577–5589, 2010.

[213] P. Zhu and Q. Hu, “Knowledge-Based Systems Rule extraction from support vector machines based on consistent region covering reduction,” *Knowledge-Based Syst.*, vol. 42, pp. 1–8, 2013.

[214] B. Heisele et al., “Hierarchical classiÿcation and feature reduction for fast face detection with support vector machines,” *Pattern Recognit.*, vol. 36, no. 2003, pp. 2007–2017, 2007.

[215] J. Yang et al., “Neurocomputing Effective search for genetic-based machine learning systems via estimation of distribution algorithms and embedded feature reduction techniques,” *Neurocomputing*, vol. 113, pp. 105–121, 2013.

[216] F. G. Zöllner et al., “SVM-based glioma grading: Optimization by feature reduction analysis,” *Z. Med. Phys.*, vol. 22, no. 3, pp. 205–214, 2012.

APPENDIX A. Terminology

This appendix aims at providing the reader with a better understanding of the concepts used in this study. More precisely, the concepts studied in this section are EKM, KBE and knowledge sourcing due to existence of wide variety of researchers' with a different view on KBE [6] and its relationship with the other two terms.

In this context, based in [5][203][54] the author considers KBE as an software application focused in the automation of repetitive tasks and support to the engineering design optimisation aiming at delivering a more efficient product development process. KBE falls into EKM discipline which is presented in [28] as *"a key for the organisations attempting to capitalise their expertise and know-how".* Other authors point out in their EKM definitions to the systematic process in which knowledge is organized in an efficient manner by making it available and facilitating its exploitation and reuse [5] [26][27]. Therefore, as presented in [5] and illustrated in Figure 61 KBE emphasis is on the technical area of developing KBE tools while EKM is focused in making a more effective and efficient use of the company data assets.

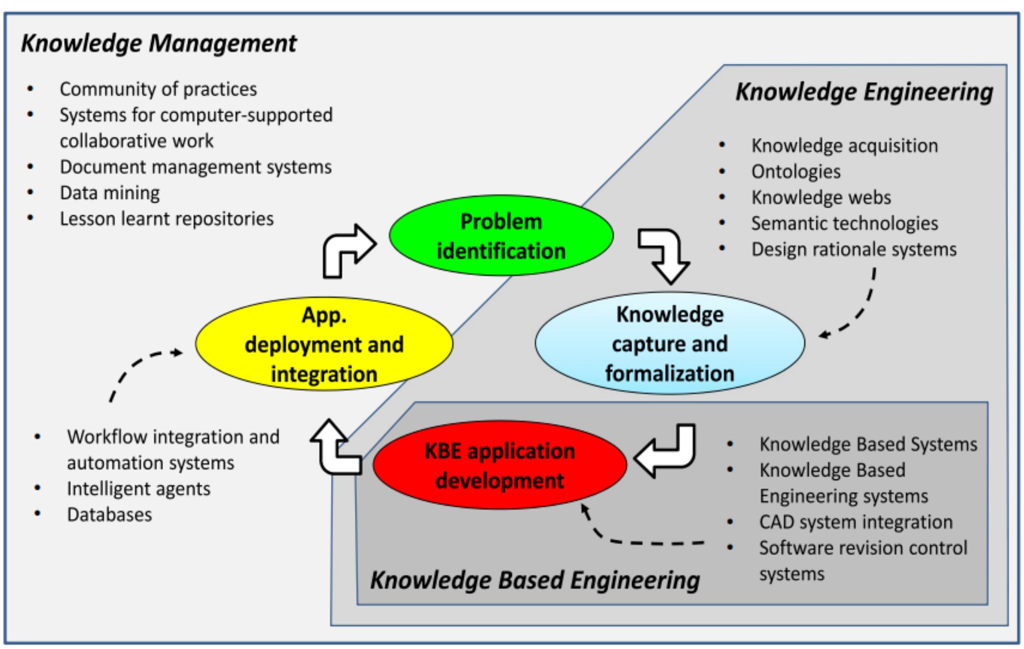


Figure 61. Relative positioning of KBE respect to knowledge management and knowledge engineering [5].

Finally, in this research knowledge sourcing emphasis is on the following EKM aspects:

5. **Creating new knowledge:** Enable knowledge sourcing not only from experts but as well from company data assets in which often relevant knowledge is hidden.
6. **Retaining the knowledge:** Enable the sharing and the modification of the knowledge.
7. **Making the knowledge ready to be exploited into KBE applications:** Capturing knowledge from experts and representing this knowledge in a human- and computer-readable language. It allows the exploitation of knowledge across different engineering problems.

The notion of “knowledge sourcing” here refers to the capacity not only to access and manage the change of knowledge in KBE applications. In addition it refers to the ability to efficiently extract knowledge in engineering scenarios characterised by:

- The complex networks of engineers that do not work anymore in the same office but in distributed teams across organisations with high level of internal and external mobility. Empirical research on engineering design indicates that that designers spend around 30% of their time looking for information that is already available [7]. The threat to engineering firms stands at the chance of engineers being able only to still use the same 30%. But on a more complex knowledge sourcing exercise motivated by factors like the lack of experts and their delocalisation.
- The explosion of data generated by digital engineering. Digitalisation of engineering work has brought significant efficiencies to the design process. However, it has also multiplied the availability of data coming from digital tools and the complexity to manage its value. Rather than a threat, data availability brings the opportunity to extract usable knowledge from information sources and thus increasing the return of investment on digitalisation.

APPENDIX B. KSF Framework Questionnaire

This appendix provides the details of the questionnaires realised in order to qualitatively validate the knowledge sourcing framework developed in this study. The aim of the questionnaires is to prove the suitability of the platform to enable the efficient and systematic source of engineering knowledge. This framework has been tested in two different domains where experts in the field were asked to fulfill the questionnaire presented below.

B.1 Questionnaires of case study 1

Questionnaire 1.1

Interviewed id: EXPERT_1

Information about the questionnaire

What is the aim of the questionnaire?

The objective of this questionnaire is to qualitatively validate the knowledge sourcing framework developed in this research to predict the MFG time required to manufacture a wing cover.

What is the purpose of this questionnaire?

The aim of this questionnaire is to validate and consolidate the assumption that a effective and more efficient way of source engineering knowledge can be realised by integrating artificial intelligence knowledge-based tools within a knowledge Sourcing framework.

How the proposed methodology is validated?

The validation of the methodology was planned to be achieved by the development of two use cases. Therefore, the validation of these two case studies will consequently validate the framework proposed.

The validation of the use case encompasses three main tasks as shown in Figure 62. The first one is linked to the process of reviewing and modification the ML rules obtained as a result of training the ML algorithm with 198 samples. In this case study M5R was selected due to the meaningfulness of its rules (assessed by experts) and the accuracy of the results provided.

Once the explicit model created was validated by experts, the last version of the ML rules (pre-validated model) were used to predict the MFG times of 71 samples corresponding to the "Test Set" file. This file contains data of two new design configurations which are not used in the learning process. The values of MAE and RMSE obtained in the validation process are 13.91 and 19.83 respectively which were considered by the experts as acceptable from and engineering point of view.

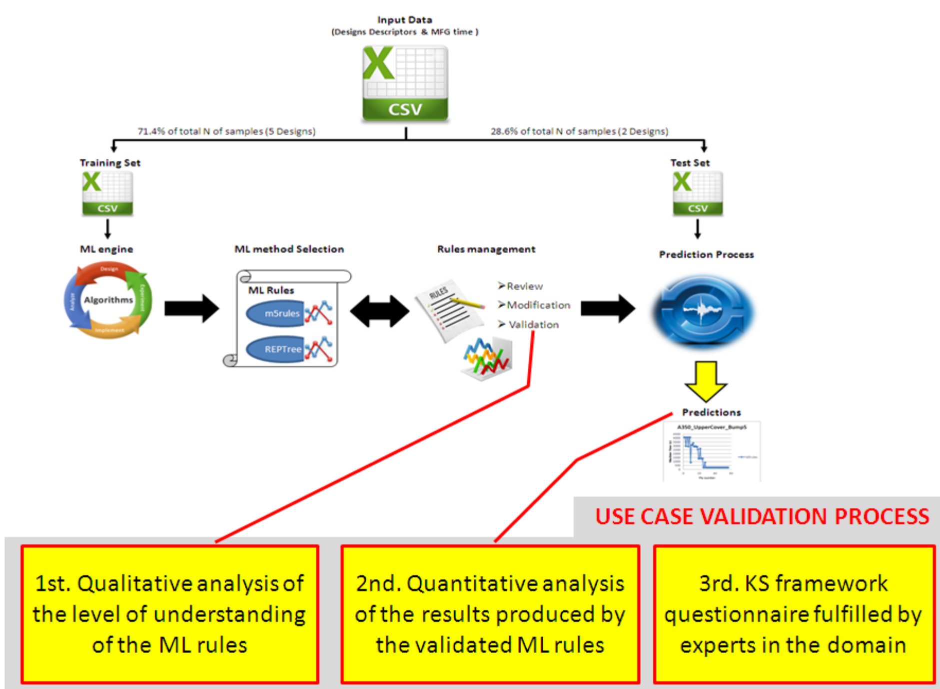


Figure 62. Use case validation process.

Finally, the last step to validate the use case was achieved by performing a qualitative analysis of the KS framework. This was realised by asking the experts to fulfil this questionnaire.

Instructions to complete the questionnaire.

- Questions should be answered ticking the correspondent box on the white area of the tables.
- Comments should be answered on the grey area of the tables.
- Questions only have one possible answer.

Technical questions

This questionnaire is intended to analyse the knowledge sourcing (KS) framework developed in this PhD study.

The KS platform developed in this study is a user friendly system fostering its wider use within the company.			
Strongly Agree ☐	Agree ☒	Disagree ☐	Strongly Disagree ☐
Comments:			

The KS framework is simple enough to be used by a non-programmer and its knowledge easy to keep up to date.			
Strongly Agree [X]	Agree [ ]	Disagree [ ]	Strongly Disagree [ ]
Comments:			

The system enables the user to easily review and validate the automated rules created.			
Strongly Agree [ ]	Agree [X]	Disagree [ ]	Strongly Disagree [ ]
Comments: It could be useful the use of an automated workflow to validate the rules more efficiently.			

The system allows the knowledge to be stored separately from its application fostering the reuse of knowledge.			
Strongly Agree ☒	Agree ☐	Disagree ☐	Strongly Disagree ☐
Comments: Knowledge is usually hard coded within the application. Therefore, this is a key feature for as due to expert knowledge is often loss when and expert leaves or gets retire.			

An expert or group of experts could not have been able to derive the same rules obtained by the KSF.			
Strongly Agree ☒	Agree ☐	Disagree ☐	Strongly Disagree ☐
Comments: I do agree but with enough time a group of experts could eventually obtain a set of meaningful and accurate rules.			

The ML algorithm could not have been able to obtain the same validated rules (obtained as a result of the interaction between ML algorithms and experts) without expert intervention.

Strongly Agree	Agree	Disagree	Strongly Disagree
☐	☒ X	☐	☐

Comments:

The knowledge sourcing methodology enables to easily trace back the results generated.

Strongly Agree	Agree	Disagree	Strongly Disagree
☐	☒ X	☐	☐

Comments: The use of a set of particular algorithms providing an explicit model enables tracing back from the predictions. This increases the accuracy of this framework compared to other systems using “black box” applications for prediction.

The knowledge sourcing capability developed is a robust system which doesn't fail easily and recovers quickly from no ordinary circumstances.				
Strongly Agree	Agree	Disagree	Strongly Disagree	N/A
☐	X	☐	☐	☐
Comments:				

The use of a well-established methodology (KNOMAD) to develop the knowledge sourcing framework delivers a generic capability which enables the systematic capture, retaining and reuse of engineering knowledge.				
Strongly Agree	Agree	Disagree	Strongly Disagree	N/A
☐	X	☐	☐	☐
Comments: KNOMAD delivers the required methodological support required by KBE for effective knowledge sourcing.				

Questionnaire 1.2

Interviewed id: EXPERT_2

Information about the questionnaire

What is the aim of the questionnaire?

The objective of this questionnaire is to qualitatively validate the knowledge sourcing framework developed in this research to predict the MFG time required to manufacture a wing cover.

What is the purpose of this questionnaire?

The aim of this questionnaire is to validate and consolidate the assumption that a effective and more efficient way of source engineering knowledge can be realised by integrating artificial intelligence knowledge-based tools within a knowledge Sourcing framework.

How the proposed methodology is validated?

The validation of the methodology was planned to be achieved by the development of two use cases. Therefore, the validation of these two case studies will consequently validate the framework proposed.

The validation of the use case encompasses three main tasks as shown in Figure 62. The first one is linked to the process of reviewing and modification the ML rules obtained as a result of training the ML algorithm with 198 samples. In this case study M5R was selected due to the meaningfulness of its rules (assessed by experts) and the accuracy of the results provided.

Once the explicit model created was validated by experts, the last version of the ML rules (pre-validated model) were used to predict the MFG times of 71 samples corresponding to the “Test Set” file. This file contains data of two new design configurations which are not used in the learning process. The values of MAE and

RMSE obtained in the validation process are 13.91 and 19.83 respectively which were considered by the experts as acceptable from and engineering point of view.

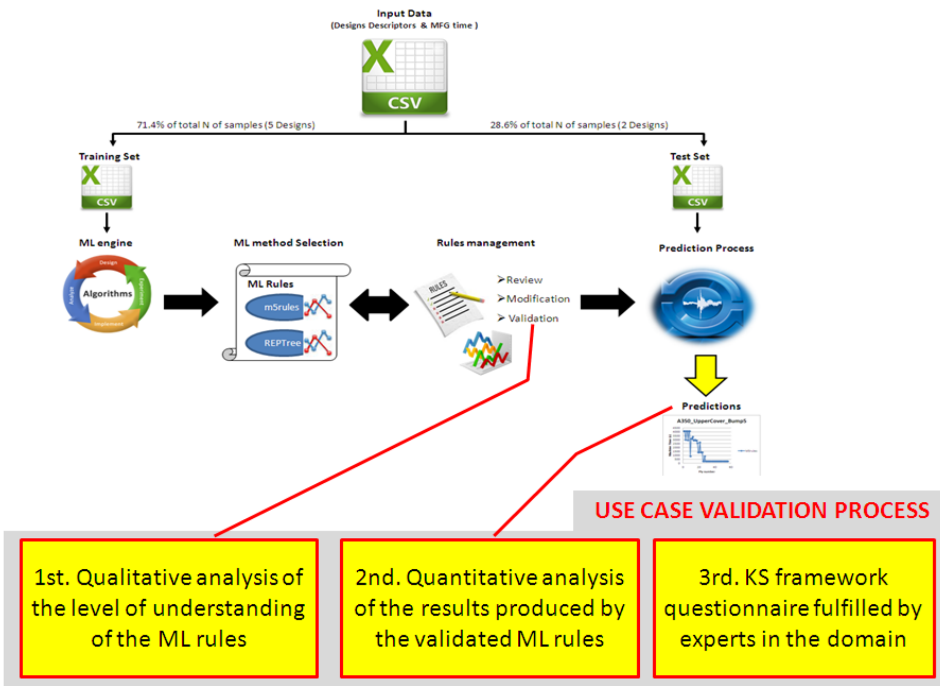


Figure 63. Use case validation process.

Finally, the last step to validate the use case was achieved by performing a qualitative analysis of the KS framework. This was realised by asking the experts to fulfil this questionnaire.

Instructions to complete the questionnaire.

- Questions should be answered ticking the correspondent box on the white area of the tables.
- Comments should be answered on the grey area of the tables.
- Questions only have one possible answer.

Technical questions

This questionnaire is intended to analyse the knowledge sourcing (KS) framework developed in this PhD study.

The KS platform developed in this study is a user friendly system fostering its wider use within the company.			
Strongly Agree ☐	Agree ☒ X	Disagree ☐	Strongly Disagree ☐
Comments: The process of familiarising with the tool was fast and little help was required from the knowledge manager in order to review and validate the rules automatically generated.			

The KS framework is simple enough to be used by a non-programmer and its knowledge easy to keep up to date.			
Strongly Agree ☐	Agree ☒ X	Disagree ☐	Strongly Disagree ☐
Comments:			

The system enables the user to easily review and validate the automated rules created.			
Strongly Agree ☐	Agree ☒ X	Disagree ☐	Strongly Disagree ☐
Comments:			

The system allows the knowledge to be stored separately from its application fostering the reuse of knowledge.			
Strongly Agree ☒ X	Agree ☐	Disagree ☐	Strongly Disagree ☐
Comments: Usually the knowledge used by CAD tools is hard-coded making difficult to reuse that knowledge in other problems. Fostering the storage of the knowledge out of the application is highly convenient for enhanced knowledge re-use.			

An expert or group of experts could not have been able to derive the same rules obtained by the KSF.			
Strongly Agree ☐	Agree ☒ X	Disagree ☐	Strongly Disagree ☐
Comments: Not in the same time frame but a group of experts without the intervention of the AI algorithms could lead a set of rules with more meaningful rules. However, some correlations could be not taken into account without the use of advanced analytical tools.			

The ML algorithm could not have been able to obtain the same validated rules (obtained as a result of the interaction between ML algorithms and experts) without expert intervention.			
Strongly Agree ☐	Agree ☒ X	Disagree ☐	Strongly Disagree ☐
Comments:			

The knowledge sourcing methodology enables to easily trace back the results generated.			
Strongly Agree X	Agree ☐	Disagree ☐	Strongly Disagree ☐
Comments:			

The use of a well-established methodology (KNOMAD) to develop the knowledge sourcing framework delivers a generic capability which enables the systematic capture, retaining and reuse of engineering knowledge.				
Strongly Agree ☐	Agree X	Disagree ☐	Strongly Disagree ☐	N/A ☐
Comments: The existence of predefined guidelines for the development of the knowledge sourcing framework increases the applicability of the framework.				

Questionnaire 1.3

<u>Interviewed id:</u> EXPERT_3

<u>Information about the questionnaire</u>

What is the aim of the questionnaire?

The objective of this questionnaire is to qualitatively validate the knowledge sourcing framework developed in this research to predict the MFG time required to manufacture a wing cover.

What is the purpose of this questionnaire?

The aim of this questionnaire is to validate and consolidate the assumption that a effective and more efficient way of source engineering knowledge can be realised by integrating artificial intelligence knowledge-based tools within a knowledge Sourcing framework.

How the proposed methodology is validated?

The validation of the methodology was planned to be achieved by the development of two use cases. Therefore, the validation of these two case studies will consequently validate the framework proposed.

The validation of the use case encompasses three main tasks as shown in Figure 62. The first one is linked to the process of reviewing and modification the ML rules obtained as a result of training the ML algorithm with 198 samples. In this case study M5R was selected due to the meaningfulness of its rules (assessed by experts) and the accuracy of the results provided.

Once the explicit model created was validated by experts, the last version of the ML rules (pre-validated model) were used to predict the MFG times of 71 samples corresponding to the “Test Set” file. This file contains data of two new design configurations which are not used in the learning process. The values of MAE and

RMSE obtained in the validation process are 13.91 and 19.83 respectively which were considered by the experts as acceptable from and engineering point of view.

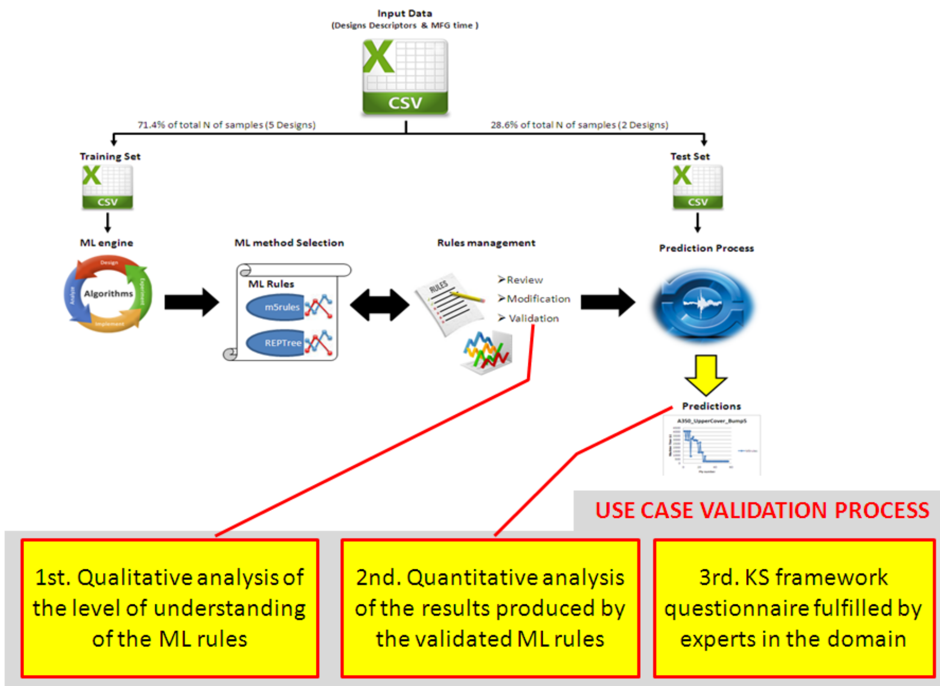


Figure 64. Use case validation process.

Finally, the last step to validate the use case was achieved by performing a qualitative analysis of the KS framework. This was realised by asking the experts to fulfil this questionnaire.

Instructions to complete the questionnaire.

- Questions should be answered ticking the correspondent box on the white area of the tables.
- Comments should be answered on the grey area of the tables.
- Questions only have one possible answer.

<u>Technical questions</u>

This questionnaire is intended to analyse the knowledge sourcing (KS) framework developed in this PhD study.

The KS platform developed in this study is a user friendly system fostering its wider use within the company.			
Strongly Agree ☐	Agree ☒ X	Disagree ☐	Strongly Disagree ☐
Comments:			

The KS framework is simple enough to be used by a non-programmer and its knowledge easy to keep up to date.			
Strongly Agree ☐	Agree ☒ X	Disagree ☐	Strongly Disagree ☐
Comments: The modification of the rules can be a little complex for people with non-programming skills.			

The system enables the user to easily review and validate the automated rules created.			
Strongly Agree ☐	Agree ☒ X	Disagree ☐	Strongly Disagree ☐
Comments: It is sufficient for an initial system like this but it would be beneficial to include additional features (e.g. easier access to previous versions of the rules).			

The system allows the knowledge to be stored separately from its application fostering the reuse of knowledge.			
Strongly Agree ☒ X	Agree ☐	Disagree ☐	Strongly Disagree ☐
Comments:			

An expert or group of experts could not have been able to derive the same rules obtained by the KSF.			
Strongly Agree ☐	Agree ☒ X	Disagree ☐	Strongly Disagree ☐
Comments: The use of advanced data mining techniques is quite useful for the understanding of processes involving new technologies.			

The ML algorithm could not have been able to obtain the same validated rules (obtained as a result of the interaction between ML algorithms and experts) without expert intervention.			
Strongly Agree ☒ X	Agree ☐	Disagree ☐	Strongly Disagree ☐
Comments: It would depend on the amount of data points and the selected features. AI experts claim that with enough information (data points) an advanced algorithm will provide a highly accurate model. However the need of the experts to define the features to be used in the model still would be required.			

The knowledge sourcing methodology enables to easily trace back the results generated.			
Strongly Agree ☐	Agree ☒ X	Disagree ☐	Strongly Disagree ☐
Comments: The advanced features delivered by the existing visualisations facilitate the data traceability activity. Data points can be easily linked to the values in the “inputs” table or to their corresponding rule.			

The use of a well-established methodology (KNOMAD) to develop the knowledge sourcing framework delivers a generic capability which enables the systematic capture, retaining and reuse of engineering knowledge.				
Strongly Agree ☐	Agree ☒ X	Disagree ☐	Strongly Disagree ☐	N/A ☐
Comments: The use of KNOMAD + ontologies enhances knowledge maintain and update. This increases users’ engagement.				

B.2 Questionnaires of case study 2

Questionnaire 2.1

<u>Interviewed id:</u> EXPERT_4

<u>Information about the questionnaire</u>

What is the aim of the questionnaire?

The objective of this questionnaire is to qualitatively validate the knowledge sourcing framework developed in this research to predict lightning strike effects.

What is the purpose of this questionnaire?

The aim of this questionnaire is to validate and consolidate the assumption that a effective and more efficient way of source engineering knowledge can be realised by integrating artificial intelligence knowledge-based tools within a knowledge Sourcing framework.

How the proposed methodology is validated?

The validation of the methodology was planned to be achieved by the development of two use cases. Therefore, the validation of these two case studies will consequently validate the framework proposed.

The validation of the use case encompasses three main tasks as shown in Figure 62. The first one is linked to the process of reviewing and modification the ML rules obtained as a result of training the ML algorithm with 76 samples. In this case study M5R was selected due to the meaningfulness of its rules (assessed by experts) and the accuracy of the results provided.

Once the explicit model created was validated by experts, the last version of the ML rules (pre-validated model) were used to predict the MFG times of 33 samples corresponding to the “Test Set” file. This file contains data of two new design configurations which are not used in the learning process. The values of MAE and

RMSE obtained in the validation process are 0.9 and 1.36 respectively which were considered by the experts as acceptable from and engineering point of view.

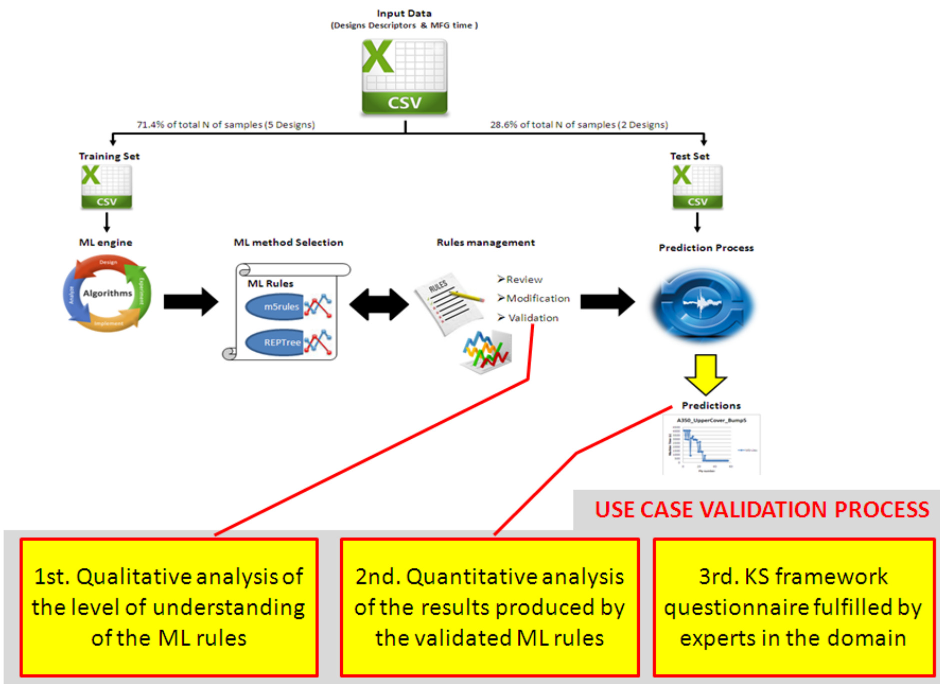


Figure 65. Use case validation process.

Finally, the last step to validate the use case was achieved by performing a qualitative analysis of the KS framework. This was realised by asking the experts to fulfil this questionnaire.

Instructions to complete the questionnaire.

- Questions should be answered ticking the correspondent box on the white area of the tables.
- Comments should be answered on the grey area of the tables.
- Questions only have one possible answer.

<u>Technical questions</u>

This questionnaire is intended to analyse the knowledge sourcing (KS) framework developed in this PhD study.

The KS platform developed in this study is a user friendly system fostering its wider use within the company.			
Strongly Agree ☐	Agree ☒	Disagree ☐	Strongly Disagree ☐
Comments: This framework enables to gather data in one place what allows instant opportunity for processing. The relationships between parameters are complex and often non-linear. Many test campaigns are similar with sometimes subtle variations. The platform can take advantage of this by expanding the dataset.			

The KS framework is simple enough to be used by a non-programmer and its knowledge easy to keep up to date.			
Strongly Agree ☒	Agree ☐	Disagree ☐	Strongly Disagree ☐
Comments: The user interface means that no programming skills are required.			

The system enables the user to easily review and validate the automated rules created.			
Strongly Agree ☐	Agree X	Disagree ☐	Strongly Disagree ☐
Comments: There is facility to review the rules that are automatically generated by the algorithms and manipulate the rules if expert experience is beneficial. The rules can then be validated against a new dataset with statistical outputs of the error.			

The system allows the knowledge to be stored separately from its application fostering the reuse of knowledge.			
Strongly Agree ☐	Agree X	Disagree ☐	Strongly Disagree ☐
Comments: Yes, the rules/model generated are stored in a repository and can be accessed/edited at a later date.			

An expert or group of experts could not have been able to derive the same rules obtained by the KSF.			
Strongly Agree ☐	Agree ☒ X	Disagree ☐	Strongly Disagree ☐
Comments: The relationships between parameters are complicated and often non-linear. Perhaps with sufficient time an expert could derive a model that is greatly beneficial to be supported by machine learning algorithms.			

The ML algorithm could not have been able to obtain the same validated rules (obtained as a result of the interaction between ML algorithms and experts) without expert intervention.			
Strongly Agree ☐	Agree ☒ X	Disagree ☐	Strongly Disagree ☐
Comments: For the same reason as above (complex and non-linear relationships between			

parameters) it is beneficial for the expert to influence rules in order to increase accuracy. For example, the expert may know when a certain parameter exceeds a given threshold, the response is very different.

The knowledge sourcing methodology enables to easily trace back the results generated.			
Strongly Agree ☐	Agree ☒ X	Disagree ☐	Strongly Disagree ☐
Comments:			

The use of a well-established methodology (KNOMAD) to develop the knowledge sourcing framework delivers a generic capability which enables the systematic capture, retaining and reuse of engineering knowledge.				
Strongly Agree ☐	Agree ☒ X	Disagree ☐	Strongly Disagree ☐	N/A ☐

Comments:

The methodology is the most powerful part of the system because it allows for easy storage of data from many tests campaigns which simplifies the process of analysing differences between campaigns.

Questionnaire 2.2

Interviewed id: EXPERT_5

Information about the questionnaire

What is the aim of the questionnaire?

The objective of this questionnaire is to qualitatively validate the knowledge sourcing framework developed in this research to predict lightning strike effects.

What is the purpose of this questionnaire?

The aim of this questionnaire is to validate and consolidate the assumption that a effective and more efficient way of source engineering knowledge can be realised by integrating artificial intelligence knowledge-based tools within a knowledge Sourcing framework.

How the proposed methodology is validated?

The validation of the methodology was planned to be achieved by the development of two use cases. Therefore, the validation of these two case studies will consequently validate the framework proposed.

The validation of the use case encompasses three main tasks as shown in Figure 62. The first one is linked to the process of reviewing and modification the ML rules obtained as a result of training the ML algorithm with 76 samples. In this case study M5R was selected due to the meaningfulness of its rules (assessed by experts) and the accuracy of the results provided.

Once the explicit model created was validated by experts, the last version of the ML rules (pre-validated model) were used to predict the MFG times of 33 samples corresponding to the “Test Set” file. This file contains data of two new design configurations which are not used in the learning process. The values of MAE and RMSE obtained in the validation process are 0.9 and 1.36 respectively which were considered by the experts as acceptable from and engineering point of view.

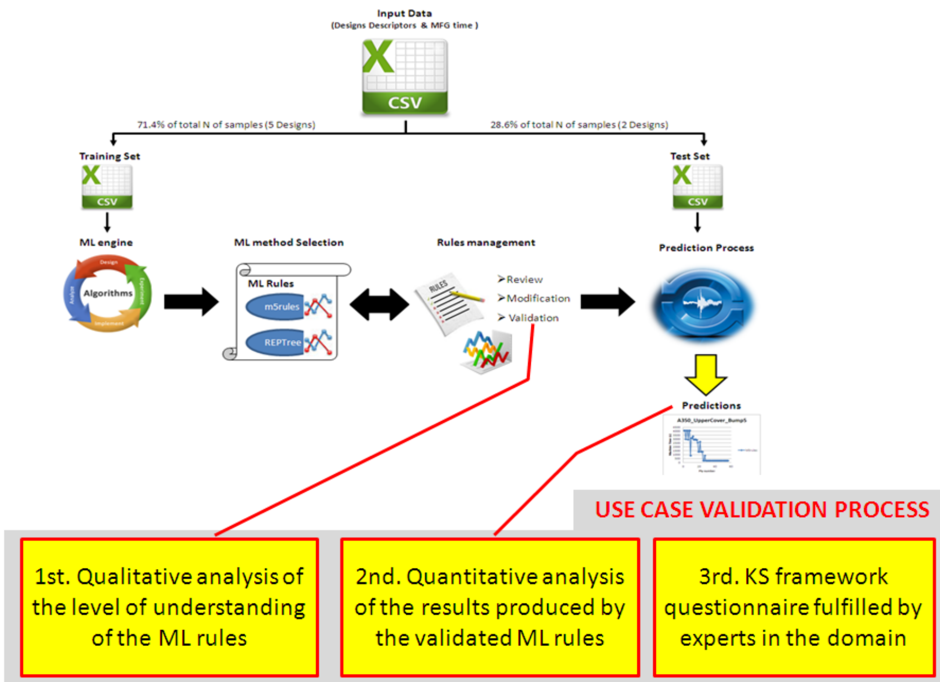


Figure 66. Use case validation process.

Finally, the last step to validate the use case was achieved by performing a qualitative analysis of the KS framework. This was realised by asking the experts to fulfil this questionnaire.

Instructions to complete the questionnaire.

- Questions should be answered ticking the correspondent box on the white area of the tables.
- Comments should be answered on the grey area of the tables.
- Questions only have one possible answer.

Technical questions

This questionnaire is intended to analyse the knowledge sourcing (KS) framework developed in this PhD study.

The KS platform developed in this study is a user friendly system fostering its wider use within the company.			
Strongly Agree ☒	Agree ☐	Disagree ☐	Strongly Disagree ☐
Comments: The platform allows quick and easy access to visual analytical tools to manipulate data. It is especially useful for users no used-to or confident with manipulating large or complex data sets.			

The KS framework is simple enough to be used by a non-programmer and its knowledge easy to keep up to date.			
Strongly Agree ☐	Agree ☒	Disagree ☐	Strongly Disagree ☐
Comments: The nature of our tests and samples (extremely complex) requires expert			

intervention to keep knowledge up-to-date.

The system enables the user to easily review and validate the automated rules created.			
Strongly Agree ☐	Agree ☒ X	Disagree ☐	Strongly Disagree ☐
Comments: The system to review and validate the rules is quick and easy to use. Validation requires some expert knowledge to ensure it is reasonable and grounded in reality.			

The system allows the knowledge to be stored separately from its application fostering the reuse of knowledge.			
Strongly Agree ☒ X	Agree ☐	Disagree ☐	Strongly Disagree ☐
Comments: The system allows users to easily reuse/re-analyse the data.			

An expert or group of experts could not have been able to derive the same rules obtained by the KSF.			
Strongly Agree ☐	Agree ☐	Disagree X	Strongly Disagree ☐
Comments: With enough time and resources a group of experts using other advanced tools/techniques may be able to produce similar sets of validated rules. This would probably not be as quick or efficient as using the KSF proposed in this research.			

The ML algorithm could not have been able to obtain the same validated rules (obtained as a result of the interaction between ML algorithms and experts) without expert intervention.			
Strongly Agree X	Agree ☐	Disagree ☐	Strongly Disagree ☐
Comments: Extremely complex nature of tests and samples means expert			

intervention/knowledge is essential.

The knowledge sourcing methodology enables to easily trace back the results generated.			
Strongly Agree X	Agree ☐	Disagree ☐	Strongly Disagree ☐
Comments: Data/results are easy to trace using the knowledge sourcing platform.			

The use of a well-established methodology (KNOMAD) to develop the knowledge sourcing framework delivers a generic capability which enables the systematic capture, retaining and reuse of engineering knowledge.				
Strongly Agree ☐	Agree X	Disagree ☐	Strongly Disagree ☐	N/A ☐
Comments: The systems seems to be a generic approach for knowledge sourcing. However, it would be beneficial to try with more campaigns.				

Questionnaire 3

<u>Interviewed id: </u>EXPERT_6

<u>Information about the questionnaire</u>

What is the aim of the questionnaire?

The objective of this questionnaire is to qualitatively validate the knowledge sourcing framework developed in this research to predict lightning strike effects.

What is the purpose of this questionnaire?

The aim of this questionnaire is to validate and consolidate the assumption that a effective and more efficient way of source engineering knowledge can be realised by integrating artificial intelligence knowledge-based tools within a knowledge Sourcing framework.

How the proposed methodology is validated?

The validation of the methodology was planned to be achieved by the development of two use cases. Therefore, the validation of these two case studies will consequently validate the framework proposed.

The validation of the use case encompasses three main tasks as shown in Figure 62. The first one is linked to the process of reviewing and modification the ML rules obtained as a result of training the ML algorithm with 76 samples. In this case study M5R was selected due to the meaningfulness of its rules (assessed by experts) and the accuracy of the results provided.

Once the explicit model created was validated by experts, the last version of the ML rules (pre-validated model) were used to predict the MFG times of 33 samples corresponding to the “Test Set” file. This file contains data of two new design configurations which are not used in the learning process. The values of MAE and RMSE obtained in the validation process are 0.9 and 1.36 respectively which were considered by the experts as acceptable from and engineering point of view.

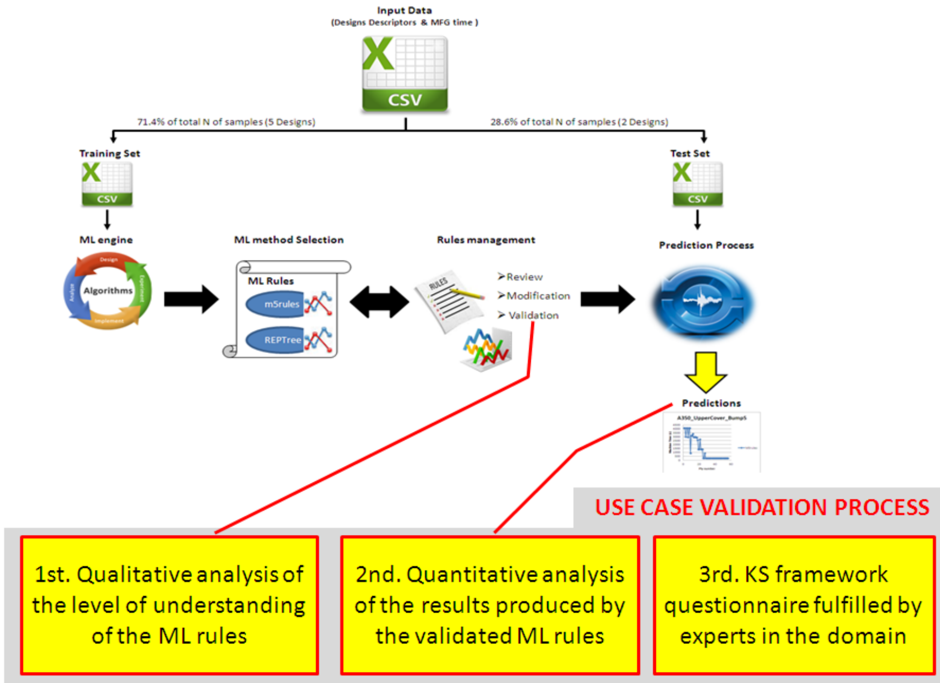


Figure 67. Use case validation process.

Finally, the last step to validate the use case was achieved by performing a qualitative analysis of the KS framework. This was realised by asking the experts to fulfil this questionnaire.

Instructions to complete the questionnaire.

- Questions should be answered ticking the correspondent box on the white area of the tables.
- Comments should be answered on the grey area of the tables.
- Questions only have one possible answer.

Technical questions

This questionnaire is intended to analyse the knowledge sourcing (KS) framework developed in this PhD study.

The KS platform developed in this study is a user friendly system fostering its wider use within the company.			
Strongly Agree X	Agree ☐	Disagree ☐	Strongly Disagree ☐
Comments: For an initial system is very easy to use. Wider use of the tool will largely depend on their mind-set as much as the quality of the tool.			

The KS framework is simple enough to be used by a non-programmer and its knowledge easy to keep up to date.			
Strongly Agree ☐	Agree X	Disagree ☐	Strongly Disagree ☐
Comments: There are some aspects such as manipulation of the rules which be off-putting to a non-programmer but this is only a small part of the tool. Most of it is easy to use.			

Knowledge kept up-to-date is very good for such a low TRL tool.

The system enables the user to easily review and validate the automated rules created.			
Strongly Agree ☐	Agree ☒	Disagree ☐	Strongly Disagree ☐
Comments:			

The system allows the knowledge to be stored separately from its application fostering the reuse of knowledge.			
Strongly Agree ☐	Agree ☒	Disagree ☐	Strongly Disagree ☐
Comments:			

An expert or group of experts could not have been able to derive the same rules obtained by the KSF.			
Strongly Agree ☐	Agree ☒ X	Disagree ☐	Strongly Disagree ☐
Comments: Agree, within such a short timeframe. Given a large time to complete the task, based on the fact the rules should follow physics, it is plausible that an expert would generate them eventually.			

The ML algorithm could not have been able to obtain the same validated rules (obtained as a result of the interaction between ML algorithms and experts) without expert intervention.			
Strongly Agree ☐	Agree ☒ X	Disagree ☐	Strongly Disagree ☐
Comments: Decisions need to be made on physical basis of rules and engineering tolerance that is appropriate in results.			

The knowledge sourcing methodology enables to easily trace back the results generated.			
Strongly Agree ☐	Agree ☒ X	Disagree ☐	Strongly Disagree ☐
Comments:			

The use of a well-established methodology (KNOMAD) to develop the knowledge sourcing framework delivers a generic capability which enables the systematic capture, retaining and reuse of engineering knowledge.				
Strongly Agree ☐	Agree ☒ X	Disagree ☐	Strongly Disagree ☐	N/A ☐
Comments:				

APPENDIX C. Framework Storyboard

Storyboard is a well-known methodology in which a set of drawings are created to permit the visualisation and understanding of a sequence of events [204]. In this case, the selected format is the comic-strip where the use of illustrations to tell the story is exploited. To facilitate the understanding, the storyboard has been divided into two main points: background and solution suggested.

C.1 Background

The context of this storyboard is the aerospace industry and more particularly the conceptual engineering design process where a wide variety of design variations are analysed with the aim of selecting an optimal design.

The storyboard starts with a common industrial scenario where a designer ("D" in the figures) sends data of a new design to an expert in design for manufacturing ("E" in the figures) in order to get it manufactured (Figure 68). Often designers are not aware of the constraints belonging to the manufacturing processes such as if the part can be produced or how long is going to take to be built. Therefore, having this information at the conceptual design stage supports the achievement of an optimal design.

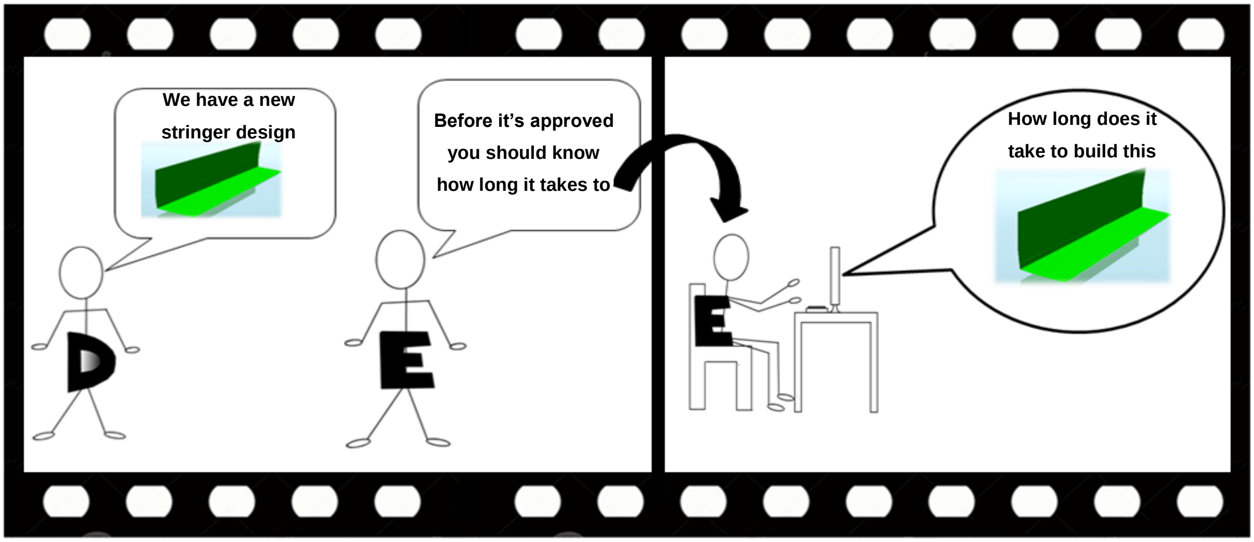


Figure 68. Designer asking to an expert to estimate the time required to manufacture a new design

Once the expert in design for manufacturing receives the data about a new design he/she will try to estimate how long it will take to build it. This will enable to work out how many parts can be manufactured in one year and how much it will cost. To do this, the expert will create a set of equations that calculate the manufacturing time. The MFG time will depend on certain parameters (area, curvature, etc) and a few coefficients. A limitation found following this procedure is caused as a result of the use of immature technologies involved and the lack of theory about the process. For instance, after interviewing experts in the context of the aerospace industry it was observed that there is not yet enough knowledge to accurately predict the manufacturing cycle time of parts made of carbon fibre reinforced plastic (CFRP) manufactured using technologies such as Automated Tape Laying (ATL) and Advanced Fibre Placement (AFP).

Figure 69 shows a situation where the expert is trying to create an explicit model that defines the system behaviour. Moreover, in this picture it is shown how the expert is lacking knowledge about the process caused by the immature technologies involved. A traditional approach to solve this problem has been the use of simulation software tools that helps engineers to find an accurate estimation for a specific engineering problem.

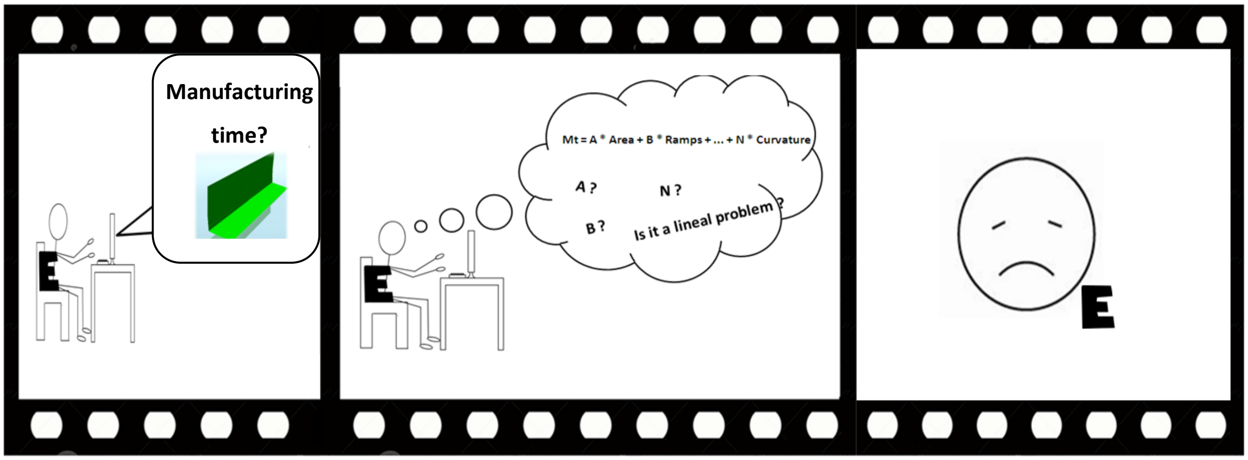


Figure 69. Approach 1: Expert trying to create a model of the problem.

In the following scenario described in Figure 70 some limitations are pointed since the time scales to obtain an answer (MFG time estimation) can be highly dependent on the technologies implicated, in some cases being longer than a week (e.g. simulation software used to estimate MFG KPI's related to the manufacture of carbon fibre reinforced plastics). In the conceptual design process, the use of software tools could be unaffordable if the time required used to simulate the manufacturing process of a part is too high and the MFG time estimation of many design variations is required.

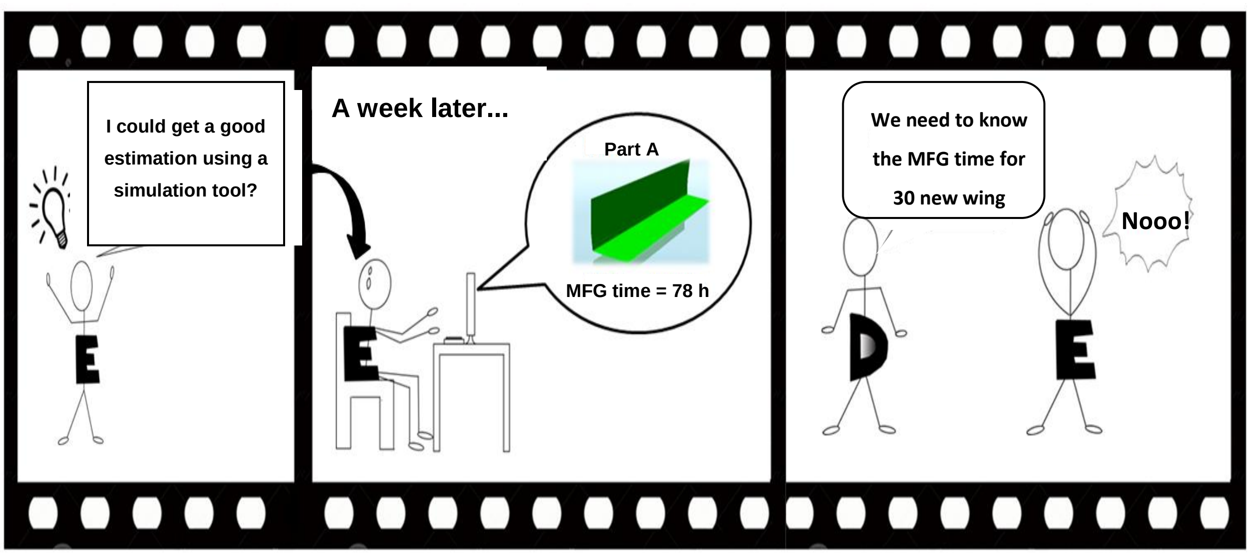


Figure 70. Approach 2: Estimation of the time required to build a part using a specific simulation tool.

A different approach is the use of new methods and tools that allow fast estimations based on company historical data. These techniques are widely used in areas such as medicine and software engineering [205][206][207][208][209][210]**.** However, a common problem found when using these methods is the lack of reliability due to the difficulty to trace back the results obtained [211][212][213][166] (Figure 71).

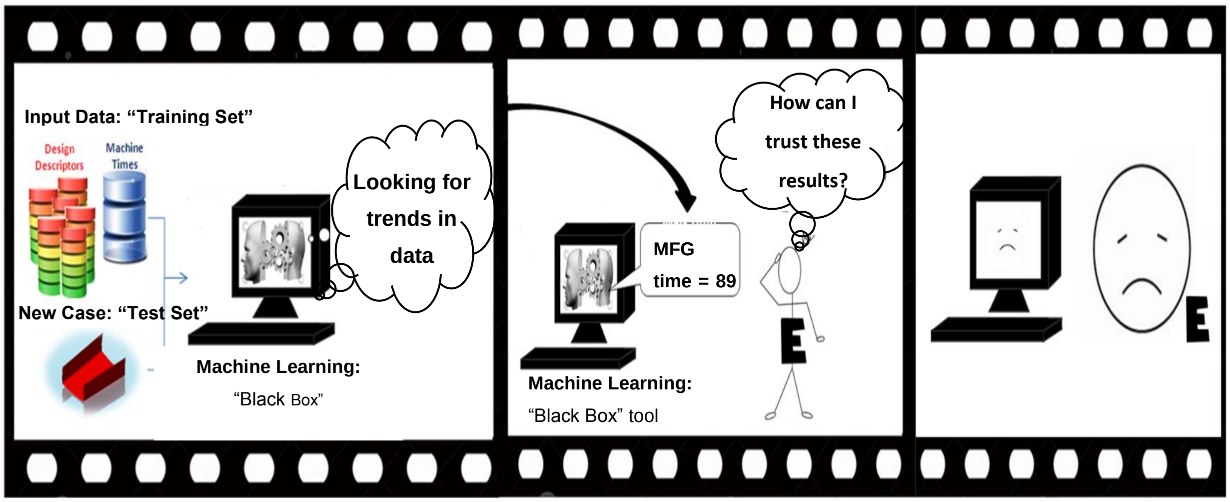


Figure 71. Approach 3: Time estimation using machine learning methods ("black box" applications).

C.2 Suggested Solution

The knowledge sourcing approach proposed in this work provides a solution to the challenges identified in section 1.1 by combining expert knowledge and machine learning algorithms (Figure 72). In this context, the framework is divided in two main stages: knowledge sourcing and required methodological support.

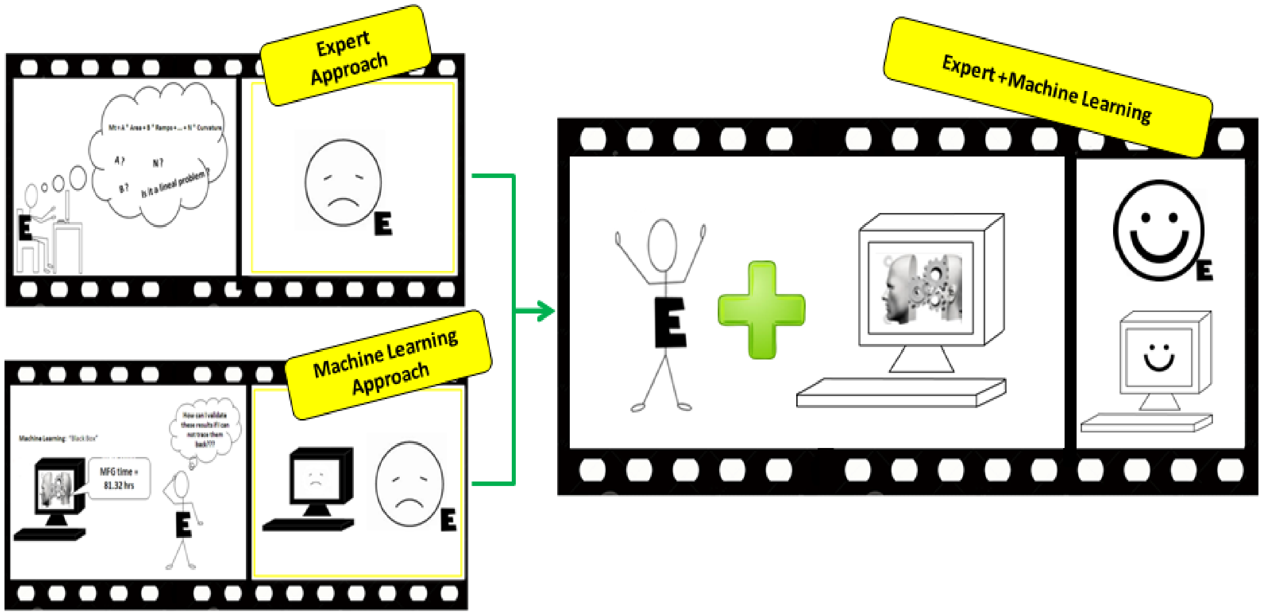


Figure 72. Suggested solution: Combination of expert knowledge and machine learning algorithms.

Stage 1: Knowledge Sourcing

This stage encompasses three phases: data analysis and knowledge creation, knowledge retention and knowledge execution.

Phase 1.1 Data analysis and knowledge creation

After the definition of the problem objective, it is required to realise a pre-process of the data captured with the aim of obtaining a dataset ("Training Set") from which a sensible explicit model can be automatically extracted (Figure 73). The parameter reduction technique is common approach employed to reduce the data noise, thus increasing the quality of a dataset [211][212][213][166]. Therefore, it is important to analyse the data captured in order to identify a list of parameters which experts believe are driving the output. The processed data is later used together the output values to create a "Training Set" which is the input file used by the machine learning tool.

The pre-process of the data captured permits to obtain more accurate results by the AI algorithms exploited. The "Training Set" generated is employed by a machine learning algorithm in a procedure named as "learning process" where an explicit model or set of rules is automatically generated. The possibility of understanding the knowledge generated by the algorithm influences the acceptance of the machine learning techniques by the users [214][215][216]. In this direction, this approach only uses algorithms delivering explicit knowledge which enables experts to trace back the results.

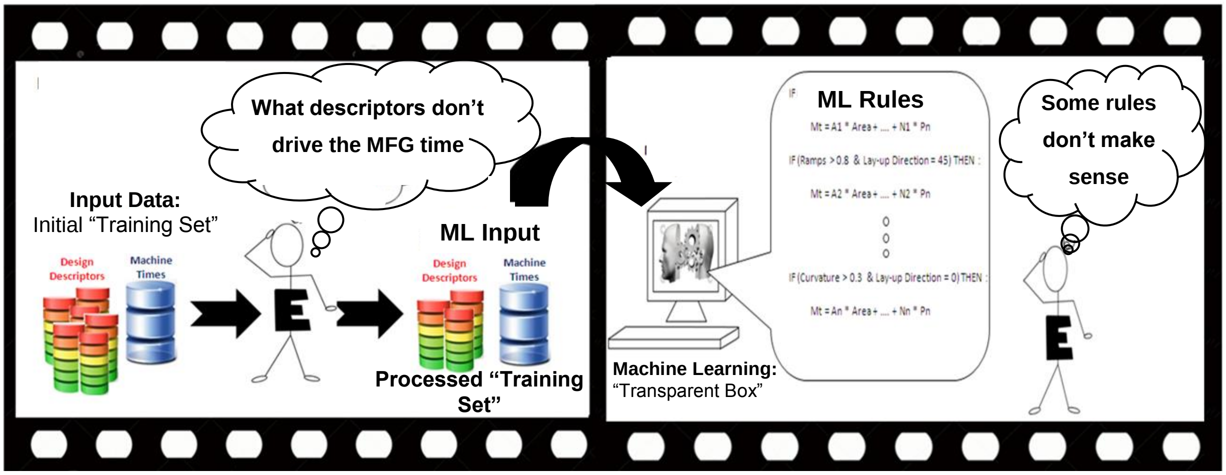


Figure 73. Stage 1: Parameters' impact analysis and "Training Set" creation.

Phase 1.2: Knowledge retention

The use of an AI tool which generates an explicit model of the problem fosters knowledge reuse and enhances the knowledge retention and maintenance. In this context, to rely on the rules created by the machine learning algorithm, a review and validation process performed by a group of experts was realised. These experts are in charge of analysing the knowledge created by the machine learning algorithm and modify it or add new knowledge (Figure 74). Once the rules modelling the problem are validated, these rules are ready to be used in the prediction of a new design configuration (similar to the ones used in the learning process).

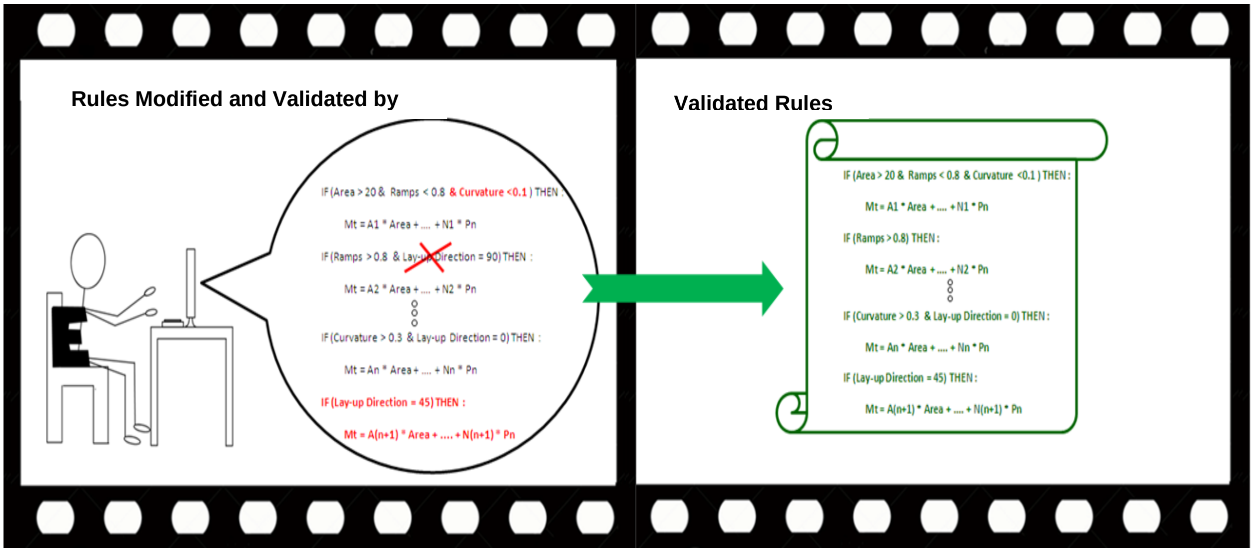


Figure 74. Stage 2: Rules' management.

Phase 1.3: Knowledge execution

In the knowledge execution stage the knowledge stored is exploited into KBE applications. This means that the knowledge (rules) created by the machine learning algorithm becomes executable and can be reused across different engineering problems. Having a set of validated rules allows fast predictions of new design configurations which are similar to previous designs Figure 75. Therefore, this approach delivers a quick and reliable solution to problems poor in theory or new problems where experts require additional knowledge to create the rules that model the system behaviour.

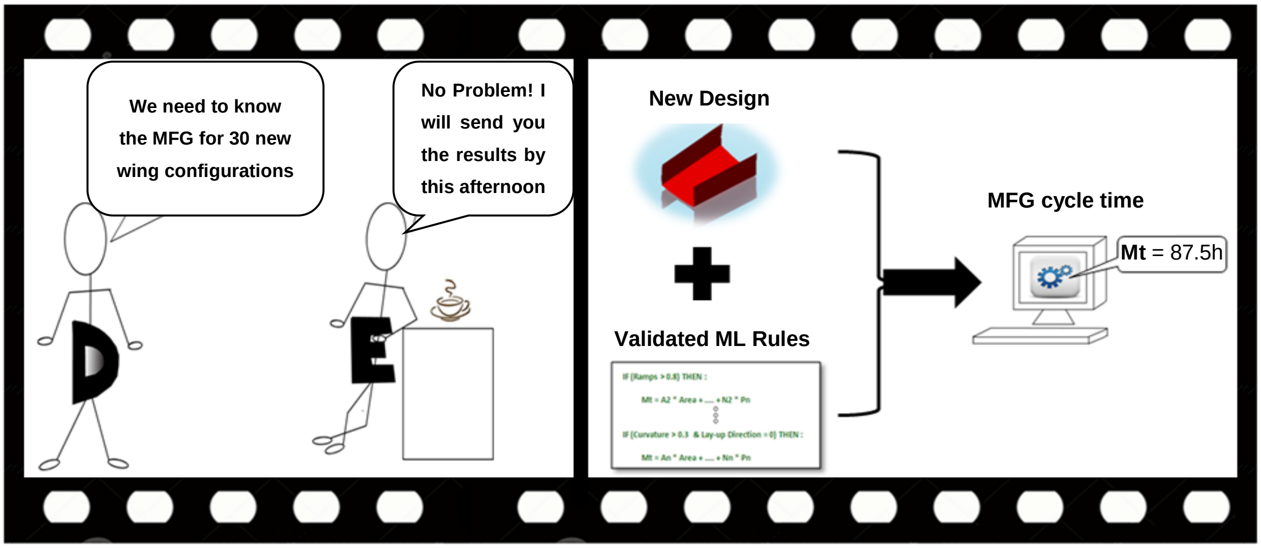


Figure 75. Stage 3: Prediction of new design configuration using ML rules.

Stage 2. Methodological support.

The use of a well-established methodology enabling the systematic capture, transfer and reuse of the knowledge created has been realised in this research. Moreover, the use of this methodological support allows the integration of KBE applications into engineering workflows.

The workflow of the KBE application developed under the scope of the first use case of this research is summarised in Figure 76. It starts with the data capture process (left side of the figure) and it is followed by the creation and management of machine learning rules. Finally, the manufacturing time prediction of a new design configuration is generated using a set of validated machine learning rules.

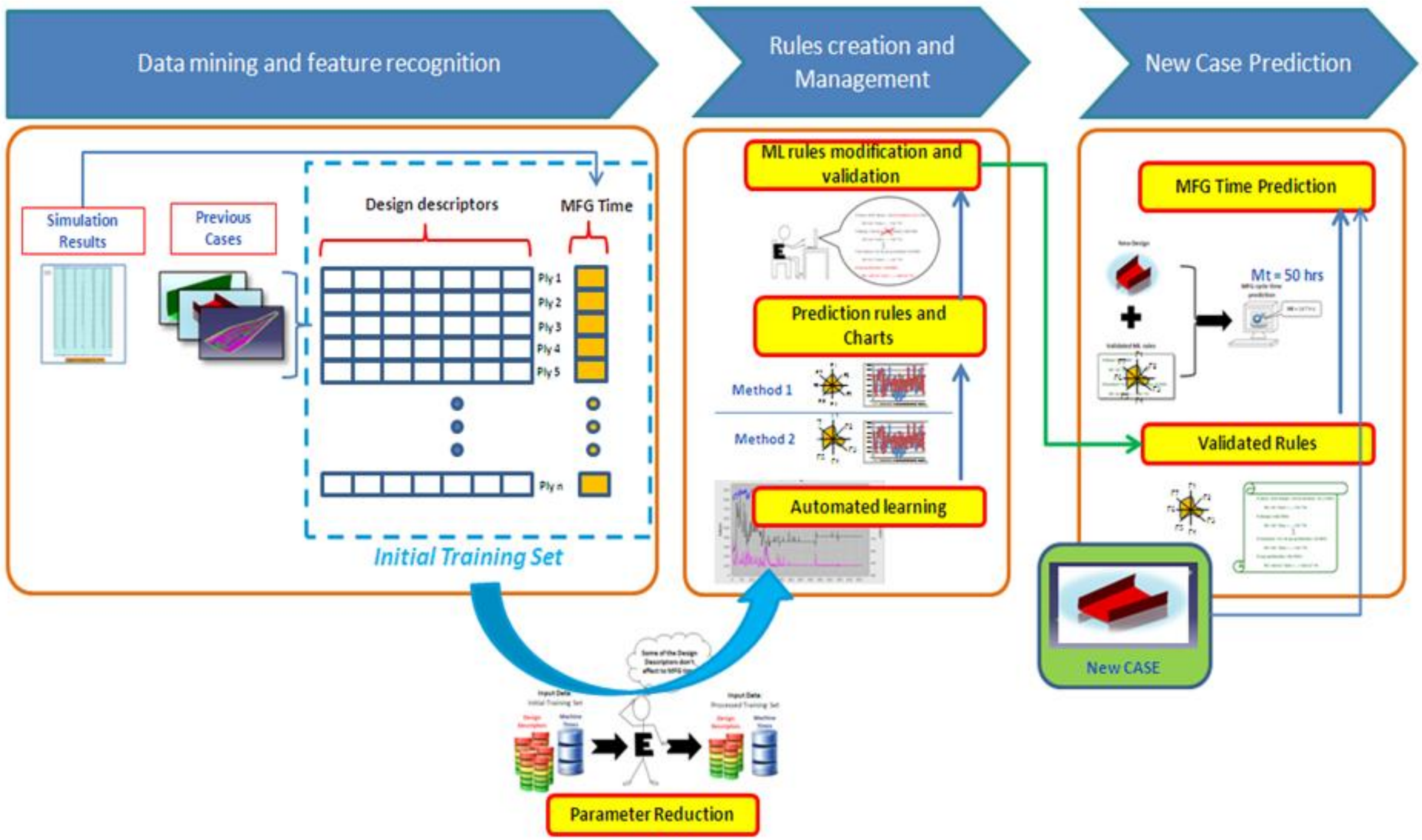


Figure 76. Methodology workflow.

APPENDIX D. Expert Interview Approach

The aim of this interview was to define the parameters the experts believe are driving the output values. The method followed to realise the interviews was the same for the two use cases developed in the context of this research.

Prior carrying out the interview, the experts provided the interviewer with a detailed explanation of the problem and its objective. This detailed description enabled the later creation of the problem domain ontology and data structure.

In first place, experts were asked to score a set of parameters regarding their impact on the output (e.g. machine cycle time in the case of the first use case) as shown in table below. This initial analysis permitted to reduce the number of parameters under study and focus on those the experts believe are driving the output.

Table 28. Parameter impact on manufacturing cycle time estimation for wing covers

Parameter	Score
Length	1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐
Width	1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐
Colour	1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐
Lay-up direction	1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐
Thickness	1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐
Curvature	1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐
Ramps	1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐

After the list of parameters having an impact on the output was reduced, experts were asked to analyse the possible dependencies of the filtered parameters. This process enabled to reduce even more the resulting list by grouping various parameters into a single one (e.g. “length” and “width” were grouped into “aspect ratio” in the first case study).

Additionally to the direct parameters identified, experts were asked to think about contextual or inferred parameters. These parameters (e.g. “packaging” and “jumps” in the first case study) supported the machine learning to build a more accurate model.

Finally, values corresponding to the direct and inferred parameters where merged with their respective output values creating the input file (“Training Set”) used by the machine learning algorithm to generate the model of the problem (Figure 77).

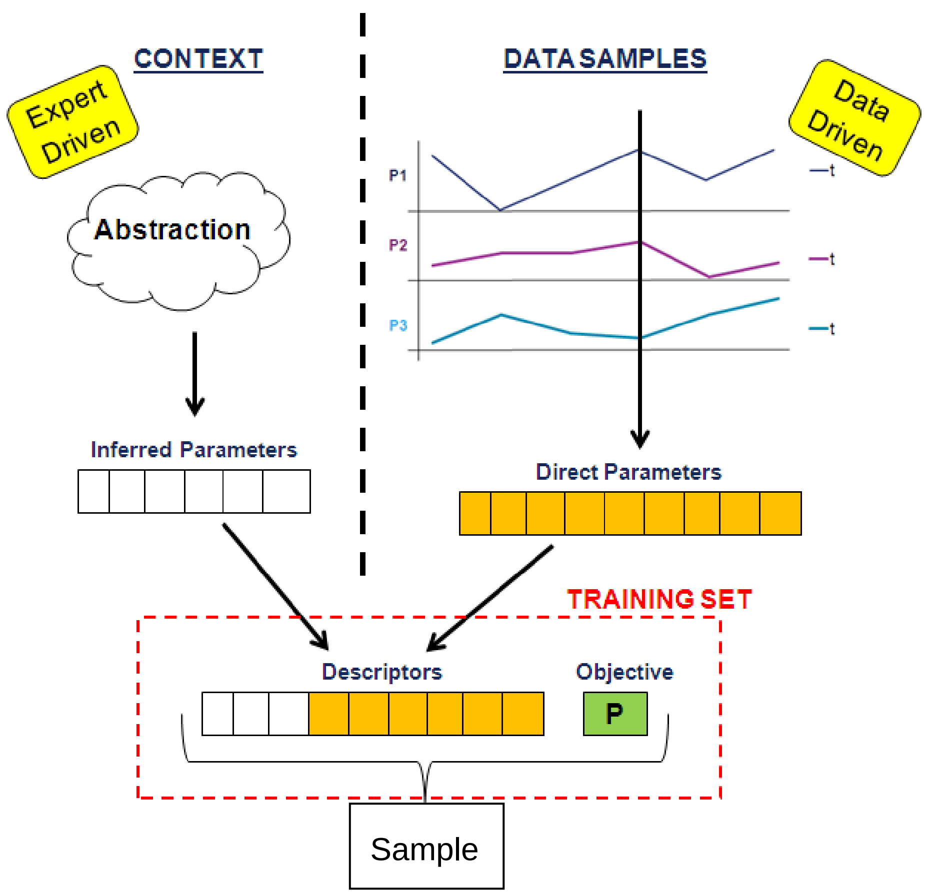


Figure 77. “Training Set” creation process

APPENDIX E. Researches classification matrix

							GENERAL AREA								AI TOOLS			APPLICATIONS				ASPECTS OF KBE					
							KM																				
Id	Year	Author	R	Title	Location	Organisation	KBE	KBS	KA	KC	P	C	AI	ED	ML	Other	What technique?	Context	KBE App.	Shows Implem. ?	What App.?	Capture of Expert Knowledge	Access to knowledge by the tool	Automated knowledge extraction from data	Advice on AI tool suitability	Knowledge lifecycle management	AI extracted knowledge ready for reuse
th_1	2010	Urška Sancin, Mitja Dobravc, Bojan Dolšak	1	Human cognition as an intelligent decision support system for plastic products Design	Slovenia	University	x											n/a									
th_2	2012	Gianfranco La Rocca	4,5	Knowledge based engineering: Between AI and CAD. Review of a language based technology to support engineering design	Holland	Delft University	x					x		x				Aerospace / Design	x		Multi Method Generator of Aircraft models	x	x			x	

							GENERAL AREA								AI TOOLS			APPLICATIONS				ASPECTS OF KBE					
							KM																				
Id	**Year**	**Author**	**R**	**Title**	**Location**	**Organisation**	**KBE**	**KBS**	**KA**	**KC**	**P**	**C**	**AI**	**ED**	**ML**	**Other**	**What technique?**	**Context**	**KBE App.**	**Shows Implem. ?**	**What App.?**	Capture of Expert Knowledge	Access to knowledge by the tool	Automated knowledge extraction from data	Advice on AI tool suitability	Knowledge lifecycle management	AI extracted knowledge ready for reuse
th_4	2012	Rob H. Bracewell	4.5	The retrieval of structured design rationale for the re-use of design knowledge with an integrated representation	UK	University	x							x				n/a									
th_5	2007	McMahon	5	Knowledge management in engineering design: personalization and codification	UK	University					x	x						n/a									
th_6	2012	Bermell P. et. al	4.5	A critical review of Knowledge-Based Engineering: An identification of research challenges	Filton, Holland	EADS IW , Delf University	x											n/a									

							GENERAL AREA								AI TOOLS			APPLICATIONS				ASPECTS OF KBE					
							KM																				
Id	Year	Author	R	Title	Location	Organisation	KBE	KBS	KA	KC	P	C	AI	ED	ML	Other	What technique?	Context	KBE App.	Shows Implem. ?	What App.?	Capture of Expert Knowledge	Access to knowledge by the tool	Automated knowledge extraction from data	Advice on AI tool suitability	Knowledge lifecycle management	AI extracted knowledge ready for reuse
th_8	2010	Llewellyn C.M. Tanga, et.al	1.5	A study of the information evaluation practice between aerospace and construction industries	UK	Universities					x	x						n/a									
th_9	1997	DAVID F. YOUNG, DONALD E. BROWN & WILLIAM T. SCHERER	3	A Polynomial Induction Network Approach to Engineering Design	Virginia	University							x	x	x		Polynomial induction network --> Artificial NN	n/a									
th_10	2011	Mogens Myrup Andreasen	1.5	45 Years with design methodology	Denmark	University								x				n/a									

							GENERAL AREA								AI TOOLS			APPLICATIONS				ASPECTS OF KBE					
							KM																				
Id	Year	Author	R	Title	Location	Organisation	KBE	KBS	KA	KC	P	C	AI	ED	ML	Other	What technique?	Context	KBE App.	Shows Implem. ?	What App.?	Capture of Expert Knowledge	Access to knowledge by the tool	Automated knowledge extraction from data	Advice on AI tool suitability	Knowledge lifecycle management	AI extracted knowledge ready for reuse
th_12	2000	Patrick Mejasson, et.al	2	Intelligent design assistant IDA : a case base reasoning system for material and design	London	University							x	x	x		CBR	n/a									
th_13	2012	Kristian Amadori	2	Flexible and robust CAD models for design automation	Sweden	University	x							x				Aerospace,Robotics / Design	x		transport aircraft, industrial robots, and micro air vehicles	x	x			x	
th_14	2012	Durk Steenhuizen, Michel van Tooren	2	The implementation of a knowledge-based framework for the aerodynamic optimization of a morphing wing device	Holland	Delft University	x							x				Aerospace	x		Airfoil	x	x			x	

							GENERAL AREA								AI TOOLS			APPLICATIONS				ASPECTS OF KBE					
							KM																				
Id	Year	Author	R	Title	Location	Organisation	KBE	KBS	KA	KC	P	C	AI	ED	ML	Other	What technique?	Context	KBE App.	Shows Implem. ?	What App.?	Capture of Expert Knowledge	Access to knowledge by the tool	Automated knowledge extraction from data	Advice on AI tool suitability	Knowledge lifecycle management	AI extracted knowledge ready for reuse
th_16	2004	Crispin Hales	3	Managing engineering design	n/a	n/a								x				n/a									
th_17		Ahmed S.		The Language of Design: Theory and Computation	n/a	n/a								x				n/a									
th_18	2005	Martin Hoegl, Anja Schulze	2	How to Support Knowledge Creation in New Product Development	Germany / Switz..	Universities				x								n/a									

							GENERAL AREA								AI TOOLS			APPLICATIONS				ASPECTS OF KBE					
							KM																				
Id	**Year**	**Author**	**R**	**Title**	**Location**	**Organisation**	**KBE**	**KBS**	**KA**	**KC**	**P**	**C**	**AI**	**ED**	**ML**	**Other**	**What technique?**	**Context**	**KBE App.**	**Shows Implem. ?**	**What App.?**	Capture of Expert Knowledge	Access to knowledge by the tool	Automated knowledge extraction from data	Advice on AI tool suitability	Knowledge lifecycle management	AI extracted knowledge ready for reuse
th_20	2006	Silvio Popadiuk, Chun Wei Choo	1.5	Innovation and knowledge creation:How are these concepts related?	Brazil / Canada	University				x								n/a									
th_21	2006	A.M.M. Sharif Ullah*, Khalifa H. Harib	4.5	A human-assisted knowledge extraction method for machining operations	Arabs Emirates	University							x			x	Mathematical approach	MFG			Machinery parameters	x	x	x			
th_22	2006	Weiming Shena, et. Al	2.5	Applications of agent-based systems in intelligent manufacturing: An updated review	Canada	University and Tech. Centre							x					n/a									

							GENERAL AREA								AI TOOLS			APPLICATIONS				ASPECTS OF KBE					
							KM																				
Id	Year	Author	R	Title	Location	Organisation	KBE	KBS	KA	KC	P	C	AI	ED	ML	Other	What technique?	Context	KBE App.	Shows Implem. ?	What App.?	Capture of Expert Knowledge	Access to knowledge by the tool	Automated knowledge extraction from data	Advice on AI tool suitability	Knowledge lifecycle management	AI extracted knowledge ready for reuse
th_24	2009	Leila S. Shafti*, Eduardo Pe ´rez	2	Evolutionary multi-feature construction for data reduction: A case study	Madrid	University							x					n/a									
th_25	2009	Oliver Bandte	1	A broad and narrow approach to interactive evolutionary design—An aircraft design example	USA	Company							x					n/a									
th_26	2009	Lionel C. Briand, et al	1	Using machine learning to refine Category-Partition test specifications and test suites	Norway and Canada	Universities							x					n/a									

							GENERAL AREA								AI TOOLS			APPLICATIONS				ASPECTS OF KBE					
							KM																				
Id	Year	Author	R	Title	Location	Organisation	KBE	KBS	KA	KC	P	C	AI	ED	ML	Other	What technique?	Context	KBE App.	Shows Implem. ?	What App.?	Capture of Expert Knowledge	Access to knowledge by the tool	Automated knowledge extraction from data	Advice on AI tool suitability	Knowledge lifecycle management	AI extracted knowledge ready for reuse
th_28	2011	Ying-Han Wu, , Heiu-Jou Shaw	3.5	Document based knowledge base engineering method for ship basic design	Taiwan	University	x					x		x				Ship / Design	x		Ship Design	x	x			x	
th_29	2011	B.M. Li, S.Q. Xie,Xu	1	Recent development of knowledge-based systems, methods and tools for one kind of production	New Zealand	University		x		x		x						n/a									
th_30	2010	Michel J.L. van Tooren, Ton. H. van der Laan	3	A multidisciplinary implementation methodology for knowledge based engineering: KNOMAD	Holland	Delft University	x							x				Aerospace / MFG	x		Riveting and Friction Stir Welding	x	x			x	

							GENERAL AREA								AI TOOLS			APPLICATIONS				ASPECTS OF KBE					
							KM																				
Id	Year	Author	R	Title	Location	Organisation	KBE	KBS	KA	KC	P	C	AI	ED	ML	Other	What technique?	Context	KBE App.	Shows Implem. ?	What App.?	Capture of Expert Knowledge	Access to knowledge by the tool	Automated knowledge extraction from data	Advice on AI tool suitability	Knowledge lifecycle management	AI extracted knowledge ready for reuse
th_32	2003	Nicolas Gardan*, Yvon Gardan	2.5	An application of knowledge based modelling using scripts	France	CAD computer science lab	x				x	x		x				n/a	x		Filling system	x	x			x	
th_33	2011	Bojan Dolšak . , Marina Novak	5	Intelligent decision support for structural design analysis	Slovenia	University	x				x	x	x	x	x		Inductive Logic Programming (ILP) algo-rithm CLAUDIEN	Finite Element Analysis	x		Mesh Creation	x	x	x		x	x
th_34	2008	Leonid Sheremetov	3.5	Knowledge-based collaborative engineering of pipe networks	Mexico	Mexican Petroleum Institute	x					x	x	x				Design	x		Pipeline Design	x	x			x	

							GENERAL AREA								AI TOOLS			APPLICATIONS				ASPECTS OF KBE					
							KM																				
Id	Year	Author	R	Title	Location	Organisation	KBE	KBS	KA	KC	P	C	AI	ED	ML	Other	What technique?	Context	KBE App.	Shows Implem. ?	What App.?	Capture of Expert Knowledge	Access to knowledge by the tool	Automated knowledge extraction from data	Advice on AI tool suitability	Knowledge lifecycle management	AI extracted knowledge ready for reuse
th_36	2006	R. Hunter, J. Rios, J.M. Perez, A. Vizan	1	A functional approach for the formalization of the fixture design process	Madrid	University	x					x		x				n/a	x		Fixtures	x	x			x	
th_37	2011	Yuan Guo, Jie Hu*, Yinghong Peng	3	Research on CBR system based on data mining	Shangai	University	x(1)						x	x	x		NN + CBR	Design	x		Power transformer	x	x	x		x	
th_38	##	Noel Leon	3	The future of computer-aided innovation	Mexico	Tech. Centre	x						x	x				n/a									

							GENERAL AREA								AI TOOLS			APPLICATIONS				ASPECTS OF KBE					
							KM																				
Id	Year	Author	R	Title	Location	Organisation	KBE	KBS	KA	KC	P	C	AI	ED	ML	Other	What technique?	Context	KBE App.	Shows Implem. ?	What App.?	Capture of Expert Knowledge	Access to knowledge by the tool	Automated knowledge extraction from data	Advice on AI tool suitability	Knowledge lifecycle management	AI extracted knowledge ready for reuse
th_40	2007	Wojciech Skarka	2.5	Application of MOKA methodology in generative model using CATIA	Poland	University	x				x	x		x				Design	x		CAT Parts	x	x			x	
th_41	2009	Jin-Woo Choi	3	Architecture of a knowledge based engineering system for weight and cost estimation for wing composite structures	Australia	University	x					x		x				Aerospace / Design	x		wings	x	x			x	
th_42	2012	Carlos Alberto Costa , et.al	1	Assessment of a Product Range Model concept to support design reuse using RBS and CBR	Brazil, UK	Universities	x					x	x					Design			Injection moulding and friction materials						

							GENERAL AREA								AI TOOLS			APPLICATIONS				ASPECTS OF KBE					
							KM																				
Id	Year	Author	R	Title	Location	Organisation	KBE	KBS	KA	KC	P	C	AI	ED	ML	Other	What technique?	Context	KBE App.	Shows Implem. ?	What App.?	Capture of Expert Knowledge	Access to knowledge by the tool	Automated knowledge extraction from data	Advice on AI tool suitability	Knowledge lifecycle management	AI extracted knowledge ready for reuse
th_44	2002	Byounggu Choi, Heeseok Lee	1	Knowledge management strategy and its link to Knowledge Creation	S. Korea	Institute Science and Technology				x		x						n/a									
th_45	2010	Yuh-Jen Chen	1.5	Development of a method for ontology-based empirical knowledge representation and reasoning	Taiwan	University												n/a									
th_46	2001	J. Liebowitz	3.7	Knowledge management and its link to arti®cial intelligence	Baltimore (USA)	University				x			x					n/a									

							GENERAL AREA								AI TOOLS			APPLICATIONS				ASPECTS OF KBE					
							KM																				
Id	Year	Author	R	Title	Location	Organisation	KBE	KBS	KA	KC	P	C	AI	ED	ML	Other	What technique?	Context	KBE App.	Shows Implem. ?	What App.?	Capture of Expert Knowledge	Access to knowledge by the tool	Automated knowledge extraction from data	Advice on AI tool suitability	Knowledge lifecycle management	AI extracted knowledge ready for reuse
th_48	1999	Balasubramaniam Ramesh, Amrit Tiwana	4	Supporting Collaborative Process Knowledge Management in NPD Teams	Atlanta	University					x	x						n/a									
th_49	1999	T. Hansen	3	Tierney, What's your strategy for managing knowledge?, Harvard Business Review, March–01/04/1999	USA	Hardvard Review					x	x						n/a									
th_50	2008	Xuan F. Zha, et.al	1	collaborative decision support for design processes: A hybrid decision support model and agent	USA	Tech. Centre								x				n/a									

							GENERAL AREA								AI TOOLS			APPLICATIONS				ASPECTS OF KBE					
							KM																				
Id	Year	Author	R	Title	Location	Organisation	KBE	KBS	KA	KC	P	C	AI	ED	ML	Other	What technique?	Context	KBE App.	Shows Implem. ?	What App.?	Capture of Expert Knowledge	Access to knowledge by the tool	Automated knowledge extraction from data	Advice on AI tool suitability	Knowledge lifecycle management	AI extracted knowledge ready for reuse
th_52	2001	Robin Cowan	2	Expert systems: aspects of and limitations to the codifiability of knowledge	Holland	University						x	x					n/a									
th_53	2000	A.Fowler	3.5	The role of AI-based technology in support of the knowledge management value activity cycle	Newcastle	University	x						x		x		CBR , NN	n/a									
th_54	2012	C.H.Y. Lam, K.L. Choy, G.T.S. Ho, S.H. Chung	1	A hybrid case-GA-based decision support model for warehouse operation in fulfilling cross-border orders	Hong Kong	University							x		x		CBR, GA	n/a			Warehouse operation planning						

							GENERAL AREA								AI TOOLS			APPLICATIONS				ASPECTS OF KBE					
							KM																				
Id	Year	Author	R	Title	Location	Organisation	KBE	KBS	KA	KC	P	C	AI	ED	ML	Other	What technique?	Context	KBE App.	Shows Implem. ?	What App.?	Capture of Expert Knowledge	Access to knowledge by the tool	Automated knowledge extraction from data	Advice on AI tool suitability	Knowledge lifecycle management	AI extracted knowledge ready for reuse
th_56	2009	Chun Ching Lee, C. Ou-Yang	2	A neural networks approach for forecasting the supplier's bid prices in supplier selection negotiation process	Taiwan	University							x		x		NN	n/a			Supplier selection						
th_57	2010	Jui-Sheng Chou. Et al	2	Predicting the development cost of TFT-LCD manufacturing equipment with artificial intelligence models	Taiwan	University							x		x		MRA, ANNs, CBR, Hi	n/a			Predicting the development cost of mfg equipment						
th_58	2009	H.C.W. Lau, G.T.S. Ho, Y. Zhao, N.S.H. Chung	3.5	Development of a process mining system for supporting knowledge discovery in a supply chain network	Hong Kong	University	x(3)						x			x(4)	i-PM algorithm	Suppliers chain (knowledge discovery)	x		Customer satisfaction and errors detection	x		x			x

							GENERAL AREA								AI TOOLS			APPLICATIONS				ASPECTS OF KBE					
							KM																				
Id	Year	Author	R	Title	Location	Organisation	KBE	KBS	KA	KC	P	C	AI	ED	ML	Other	What technique?	Context	KBE App.	Shows Implem. ?	What App.?	Capture of Expert Knowledge	Access to knowledge by the tool	Automated knowledge extraction from data	Advice on AI tool suitability	Knowledge lifecycle management	AI extracted knowledge ready for reuse
th_60	2000	Craig B. Chapman, Martyn Pynfold	3	The application of a knowledge based engineering approach to the rapid design analysisof an automotive structure	Coventry	University	x	x			x	x		x				Automotive/ Design	x		Car Structures	x	x			x	
th_61	2007	Gianluca Colombo, Alessandro Mosca, Fabio Sartori	3.5	Towards the design of intelligent CAD systems: An ontological approach	Milan	University	x	x					x	x				MFG / Design	x		Dies for car bodies	x	x			x	
th_62	2011	S. Kumar, R. Singh	3.6	An automated design system for progressive die	India	University		x					x	x		x	Not specified (rules encoded in LISP)	MFG/Design			Progressive die	x	x			x	

							GENERAL AREA								AI TOOLS			APPLICATIONS				ASPECTS OF KBE					
							KM																				
Id	Year	Author	R	Title	Location	Organisation	KBE	KBS	KA	KC	P	C	AI	ED	ML	Other	What technique?	Context	KBE App.	Shows Implem. ?	What App.?	Capture of Expert Knowledge	Access to knowledge by the tool	Automated knowledge extraction from data	Advice on AI tool suitability	Knowledge lifecycle management	AI extracted knowledge ready for reuse
th_64	2000	J.A. Penoyer, et al.	2	Knowledge based product life cycle systems:principles of integration of KBE and C3P	USA	Companies	x											n/a									
th_65	1995	UTPAL ROY	2.5	[illegible] symbolic and numeric analysis tools required for the development of an integrated cad system	USA	University	x								x		Heuristic-depth-first (HDF)	FEA	x		Automated Fixture Design	x	x			x	
th_66	2012	H.Z. Yang, J.F. Chen, N. Ma, D.Y. Wang	3	Implementation of knowledge-based engineering methodology in ship structural design	China	University	x						x	x	x		CBR	Design	x		Ship design	x	x			x	

							GENERAL AREA								AI TOOLS			APPLICATIONS				ASPECTS OF KBE					
							KM																				
Id	Year	Author	R	Title	Location	Organisation	KBE	KBS	KA	KC	P	C	AI	ED	ML	Other	What technique?	Context	KBE App.	Shows Implem. ?	What App.?	Capture of Expert Knowledge	Access to knowledge by the tool	Automated knowledge extraction from data	Advice on AI tool suitability	Knowledge lifecycle management	AI extracted knowledge ready for reuse
th_68	2004	Z. Lou, H. Jiang, X. Ruan	4	Development of an integrated knowledge-based system for mold-base design	China	University	x(6)	x					x	x	x		CBR,NN,Rule Frame	Mould Design	x		Mould Design	x	x			x	x
th_69	2011	Wei Wang, Rong Mo,Qingming Fan	2.5	Knowledge extraction from aerodynamic simulationdata of compressor rotor	China	University									x		Not relevant	n/a									
th_70	2007	R. Curran, et al.	3	Integrated digital design for manufacture for reduced life cycle cost	UK (Belfast)	University	x											n/a									

							GENERAL AREA								AI TOOLS			APPLICATIONS				ASPECTS OF KBE					
							KM																				
Id	Year	Author	R	Title	Location	Organisation	KBE	KBS	KA	KC	P	C	AI	ED	ML	Other	What technique?	Context	KBE App.	Shows Implem. ?	What App.?	Capture of Expert Knowledge	Access to knowledge by the tool	Automated knowledge extraction from data	Advice on AI tool suitability	Knowledge lifecycle management	AI extracted knowledge ready for reuse
th_72	1999	Ilona Jagielska, et al	4.5	networks,fuzzy logic, genetic algorithms, and rough sets to automatedknowledge acquisition for classiPcation	Australia	University			x				x		x		GA,NN, Fuzzy	Classification				x		x	x		
th_73	1998	T. Warren Liao, et al	5.5	expert system for failure mechanism identication: Part automated knowledge acquisition	Lousiana USA	University			x			x	x		x		GA,	Failure detection				x	x	x		x	x
th_74	1997	Bingchiang Jeng, et al	1	FILM: a fuzzy inductive learning method for automated knowledge acquisition	Taiwan	University			x				x		x		GA	n/a									

							GENERAL AREA								AI TOOLS			APPLICATIONS				ASPECTS OF KBE					
							KM																				
Id	**Year**	**Author**	**R**	**Title**	**Location**	**Organisation**	**KBE**	**KBS**	**KA**	**KC**	**P**	**C**	**AI**	**ED**	**ML**	**Other**	**What technique?**	**Context**	**KBE App.**	**Shows Implem. ?**	**What App.?**	Capture of Expert Knowledge	Access to knowledge by the tool	Automated knowledge extraction from data	Advice on AI tool suitability	Knowledge lifecycle management	AI extracted knowledge ready for reuse
th_76	2009	Se´bastien Gebus*, Kauko Leiviska	3	Knowledge acquisition for decision support systems on an electronic assembly line	Finland	University			x			x						n/a									
th_77	2002	Jyh-Bin Yang	5	A rule induction-based knowledge system for retaining wall selection	Taiwan	University		x	x		x	x	x	x	x		Rule Induction (RI)	Construction design	x		Rule generation and case prediction	x	x	x		x	x
th_78	1998	Z. Chen	3.7	Knowledge acquisition assisted by CLIPS programming	USA	University			x				x	x	x		Learning algorithm, rough set	General			diagnostics expert system for air-craft	x (*)	x	x		x	x

							GENERAL AREA								AI TOOLS			APPLICATIONS				ASPECTS OF KBE					
							KM																				
Id	Year	Author	R	Title	Location	Organisation	KBE	KBS	KA	KC	P	C	AI	ED	ML	Other	What technique?	Context	KBE App.	Shows Implem. ?	What App.?	Capture of Expert Knowledge	Access to knowledge by the tool	Automated knowledge extraction from data	Advice on AI tool suitability	Knowledge lifecycle management	AI extracted knowledge ready for reuse
th_80	2007	Baxter, David Gao, James Case, Keith Harding, Jenny Young, Bob	3	An Engineering design knowledge reuse Methodology Using Process Modelling	Cranfield	University	x		x		x	x		x				Semiconductor	x		producing high vacuum pumps	x	x			x	
th_81	2007	Emberey, C L Milton, N R	3	Engineering Methodologies to Support Engineering Design Application Development in Aerospace	UK & Netherlands	n/a	x		x			x		x				Aerospace	x		Automation of prepeg plies	x	x			x	
th_82	2010	Elgh, Fredrik Cederfeldt, Mikael	3	Management of Product Knowledge in a System for Automated Variant Design – A Case Study	Sweden	University	x		x			x		x				n/a	x		Automated variant design	x	x			x	

							GENERAL AREA								AI TOOLS			APPLICATIONS				ASPECTS OF KBE					
							KM																				
Id	Year	Author	R	Title	Location	Organisation	KBE	KBS	KA	KC	P	C	AI	ED	ML	Other	What technique?	Context	KBE App.	Shows Implem. ?	What App.?	Capture of Expert Knowledge	Access to knowledge by the tool	Automated knowledge extraction from data	Advice on AI tool suitability	Knowledge lifecycle management	AI extracted knowledge ready for reuse
th_84	2007	Choi, Jin-woo Kelly, Don Raju, John	3	A knowledge-based engineering tool to estimate cost and weight of composite aerospace structures at the conceptual stage of the design process	Australia	University	x		x			x		x				Plastics	x		Automated estimation of weight for composite structures	x	x			x	
th_85	2009	Corallo, Angelo Laubacher, Robert Margherita, Alessandro Turrisi, Giuseppe	3	Enhancing product development through knowledge-based engineering	Italy & USA	University	x		x			x		x				Aerospace	x		Semi-automated new product development	x	x			x	
th_86	2010	F.Elgh	3	Supporting Management and Analysis of Quotations in a Design Automation Approach to Customization	n/a	n/a	x		x			x		x				Generic Design	x		efficient generation of product variants	x	x			x	

							GENERAL AREA								AI TOOLS			APPLICATIONS				ASPECTS OF KBE					
							KM																				
Id	**Year**	**Author**	**R**	**Title**	**Location**	**Organisation**	**KBE**	**KBS**	**KA**	**KC**	**P**	**C**	**AI**	**ED**	**ML**	**Other**	**What technique?**	**Context**	**KBE App.**	**Shows Implem. ?**	**What App.?**	Capture of Expert Knowledge	Access to knowledge by the tool	Automated knowledge extraction from data	Advice on AI tool suitability	Knowledge lifecycle management	AI extracted knowledge ready for reuse
th_88	2005	Kulon, J. Broomhead, P. Mynors, D. J.	3	Applying knowledge-based engineering to traditional manufacturing design	n/a	n/a	x		x			x		x				Metal Stamping	x		Decision support system for generation of forgeable geometry	x	x			x	
th_89	2006	Kulon, J. Mynors, D.J. Broomhead, P.	3	A knowledge-based engineering design tool for metal forging	UK	University	x		x			x		x				Metal Stamping	x		Decision support system for generation of forgeable geometry	x	x			x	
th_90	2006	Mendikoa, Iñigo Sorli, Mikel Barbero, Jose I Carrillo, Ana	3	Distributed Product Design and Manufacturing Based on KBE	Spain	Tech. Centre	x		x			x		x				Value chain management	x		Manages the distributed design and manufacturing process through the internet	x	x			x	

Id	Year	Author	R	Title	Location	Organisation	GENERAL AREA								AI TOOLS			APPLICATIONS				ASPECTS OF KBE					
							KM																				
							KBE	KBS	KA	KC	P	C	AI	ED	ML	Other	What technique?	Context	KBE App.	Shows Implem. ?	What App.?	Capture of Expert Knowledge	Access to knowledge by the tool	Automated knowledge extraction from data	Advice on AI tool suitability	Knowledge lifecycle management	AI extracted knowledge ready for reuse
th_92	2006	Shehab, E M Abdalla, H S	3	A cost-effective knowledge-based reasoning system for design for automation	Cranfield (UK)	University	x		x			x		x				Manufacturing	x		Automation and early cost modelling	x	x			x	
th_93	2008	Van Der Elst, S. Van Tooren, M.	3	Application of a Knowledge Engineering Process to Support Engineering Design Application Development	n/a	n/a	x		x			x		x				Aerospace	x		Application supporting the design and manufacturing of aircraft wiring harnesses	x	x			x	
th_94	2008	Van der Laan, A.H.	3	Knowledge based engineering support for aircraft component design	Delft	University	x		x			x		x				Aerospace	x		Aided components design	x	x			x	

							GENERAL AREA								AI TOOLS			APPLICATIONS				ASPECTS OF KBE					
							KM																				
Id	Year	Author	R	Title	Location	Organisation	KBE	KBS	KA	KC	P	C	AI	ED	ML	Other	What technique?	Context	KBE App.	Shows Implem. ?	What App.?	Capture of Expert Knowledge	Access to knowledge by the tool	Automated knowledge extraction from data	Advice on AI tool suitability	Knowledge lifecycle management	AI extracted knowledge ready for reuse
th_96	2013	Cui, Jin-ju Wang, De-yu	3	Application of knowledge-based engineering in ship structural design and optimization	China	University	x		x			x		x				Ship	x		Management of expert knowledge for optimised structural design	x	x			x	
th_97	2014	Jong, Wen-Ren Ting, Yu-Hung Li, Tai-Chih	3	Application of knowledge-based engineering for automated slide design	China	University	x		x			x		x				Plastics	x		KBE to integrate knowledge and CAD software to identify the undercut feature	x	x			x	
th_98	2014	Chandima Ratnayake, R.	3.5	KBE development for criticality classification of mechanical equipment: A fuzzy expert system	Norway	University	x		x			x	x	x	x		Fuzzy Logic	Maintenance Engineering	x		Use of KBE and Fuzzy Logic to prioritise maintenance activities	x	x			x	

							GENERAL AREA								AI TOOLS			APPLICATIONS				ASPECTS OF KBE					
							KM																				
Id	Year	Author	R	Title	Location	Organisation	KBE	KBS	KA	KC	P	C	AI	ED	ML	Other	What technique?	Context	KBE App.	Shows Implem. ?	What App.?	Capture of Expert Knowledge	Access to knowledge by the tool	Automated knowledge extraction from data	Advice on AI tool suitability	Knowledge lifecycle management	AI extracted knowledge ready for reuse
th_100	2011	Palmer, C. Harding, J. a. Swarnkar, R. Das, B. P. Young, R. I. M.	3	Generating rules from data mining for collaboration moderator services	Leicestershire (UK)	University	x		x			x		x	x		Several algorithms (Poly Analyst platform)	Generic	x		Semi-automatic methodology to generate rules by inserting relationships discovered as a result of data mining into a generic template	x	x	x		x	x
th_101	2014	Monticolo, Davy Mihaita, Simona Darwich, Hind Hilaire, Vincent	3	An agent-based system to build project memories during engineering projects	France	University	x		x			x		x				Generic Design	x		KBE for capitalize and reuse knolwedge depending on the user role	x	x			x	
th_102	2014	Potes Ruiz, Paula Kamsu Foguem, Bernard Grabot, Bernard	3	Generating knowledge in maintenance from Experience Feedback	France	University	x		x			x		x	x		Association rules	Maintenance Engineering	x		Decision support system for maintenance activities	x	x	x		x	x

							GENERAL AREA								AI TOOLS			APPLICATIONS				ASPECTS OF KBE					
							KM																				
Id	**Year**	**Author**	**R**	**Title**	**Location**	**Organisation**	**KBE**	**KBS**	**KA**	**KC**	**P**	**C**	**AI**	**ED**	**ML**	**Other**	**What technique?**	**Context**	**KBE App.**	**Shows Implem. ?**	**What App.?**	Capture of Expert Knowledge	Access to knowledge by the tool	Automated knowledge extraction from data	Advice on AI tool suitability	Knowledge lifecycle management	AI extracted knowledge ready for reuse
th_3	2012	Bermell P.	4,5	of Knowledge-Based Engineering applications as software services: Enabling personalization and codification	Filton	Airbus Group	x				x	x		x				Aerospace / Design	x		Mpda	x	x			x	
th_7	2000	T.W. Liao, et.al	2	A case-based reasoning system for identifying failure mechanisms	Louisiana	University							x		x		CBR + GA	n/a			Failure mechanisms						
th_11	1999	C.B Chapman*, M. Pinfold	3.5	Design engineering—a need to rethink the solution using knowledge based engineering	Coventry	University	x	x						x				Automotive/ Design	x		Mesh and Structures design	x	x			x	

							GENERAL AREA								AI TOOLS			APPLICATIONS				ASPECTS OF KBE					
							KM																					
Id	Year	Author	R	Title	Location	Organisation	KBE	KBS	KA	KC	P	C	AI	ED	ML	Other	What technique?	Context	KBE App.	Shows Implem. ?	What App.?	Capture of Expert Knowledge	Access to knowledge by the tool	Automated knowledge extraction from data	Advice on AI tool suitability	Knowledge lifecycle management	AI extracted knowledge ready for reuse
th_15	2007	J. NACSA, R. BUENO, A. ALZAGA and G. L. KOVA´ CS	1.8	Knowledge management support for machine tool designers using expert enablers	Hungary & Spain	n/a	x							x				MFG / Design	x		machinery engineering	x	x			x	
th_19	2009	Jih-Jeng Huang	1.5	The evolutionary perspective of knowledge creation – A mathematical representation	Taiwan	University				x								n/a									
th_23	2007	M.R. Gholamian*, S.M.T. Fatemi Ghom	2	Meta knowledge of intelligent manufacturing: An overview of state-of-the-art	Iran	University							x					n/a									

							GENERAL AREA								AI TOOLS			APPLICATIONS				ASPECTS OF KBE					
							KM																				
Id	Year	Author	R	Title	Location	Organisation	KBE	KBS	KA	KC	P	C	AI	ED	ML	Other	What technique?	Context	KBE App.	Shows Implem. ?	What App.?	Capture of Expert Knowledge	Access to knowledge by the tool	Automated knowledge extraction from data	Advice on AI tool suitability	Knowledge lifecycle management	AI extracted knowledge ready for reuse
th_27	1996	Yoram Reich and Nahum Travitzky	1	Machine learning of material behaviour knowledge from empirical data	Israel	University							x					n/a									
th_31	2012	Christian van der Velden	2	methodology for automation application development. Choose between DA adn KBE app depending o characteristics	Australia	University	x				x	x		x				Aerospace / Design	x		Wiring Path Design	x	x			x	
th_35	1999	J.A. Penoyer, G. Burnett, D.J. Fawcett, S.-Y. Liou	1.8	Knowledge based product life cycle systems...	USA	Companies	x						x					n/a									

							GENERAL AREA								AI TOOLS			APPLICATIONS				ASPECTS OF KBE					
							KM																				
Id	Year	Author	R	Title	Location	Organisation	KBE	KBS	KA	KC	P	C	AI	ED	ML	Other	What technique?	Context	KBE App.	Shows Implem. ?	What App.?	Capture of Expert Knowledge	Access to knowledge by the tool	Automated knowledge extraction from data	Advice on AI tool suitability	Knowledge lifecycle management	AI extracted knowledge ready for reuse
th_39	2011	W.L. Chena, S.Q. (Shane) Xie*, F.F. Zeng, B.M. Li	1	A new process knowledge representation approach using parameter flow chart	N.Zealand, China	Universities	x				x			x				n/a									
th_43	2012	Olivier Kuhn, et. Al	1	Framework for the support of knowledge-based engineering template update	France and Germany	Universities	x					x						n/a									
th_47	2012	José Luis López-Cuadrado, et. al	1	SABUMO: Towards a collaborative and semantic framework for knowledge sharing	Leganes	University							x					n/a									

							GENERAL AREA								AI TOOLS			APPLICATIONS				ASPECTS OF KBE					
							KM																				
Id	**Year**	**Author**	**R**	**Title**	**Location**	**Organisation**	**KBE**	**KBS**	**KA**	**KC**	**P**	**C**	**AI**	**ED**	**ML**	**Other**	**What technique?**	**Context**	**KBE App.**	**Shows Implem. ?**	**What App.?**	Capture of Expert Knowledge	Access to knowledge by the tool	Automated knowledge extraction from data	Advice on AI tool suitability	Knowledge lifecycle management	AI extracted knowledge ready for reuse
th_51	2001	Patrick Cohendet, Frieder Meyer-Krahmer	3	The theoretical and policy implications of knowledge codification	France, Germany	University and Tech. Centre						x						n/a									
th_55	2006	Se´bastien Thomassey T, Antonio Fiordaliso	3	A hybrid sales forecasting system based on clustering and decision trees	France, Belgium	Universities							x		x		Decission trees and clustering	n/a			Sales forecasting						
th_59	1996	Tong-Seng Quah et al.	1	Towards integrating rule-based expert systems and neural networks	Singapore	University							x		x		RBS and NN	n/a									

							GENERAL AREA								AI TOOLS			APPLICATIONS				ASPECTS OF KBE					
							KM																				
Id	Year	Author	R	Title	Location	Organisation	KBE	KBS	KA	KC	P	C	AI	ED	ML	Other	What technique?	Context	KBE App.	Shows Implem. ?	What App.?	Capture of Expert Knowledge	Access to knowledge by the tool	Automated knowledge extraction from data	Advice on AI tool suitability	Knowledge lifecycle management	AI extracted knowledge ready for reuse
th_63	2006	J.Kulon et al.	3	A knowledge-based engineering design tool for metal forging	UK	University												n/a									
th_67	2010	Wen-Chin Chen, et al.	3	A systematic optimization approach for assembly sequence planning using Taguchi method, DOE, and BPNN	Taiwan	University	x						x		x		NN	Assembly	x		Assembly Sequence	x	x			x	
th_71	2011	C.F. Cheung, W.B. Lee, W.M. Wang, Y. Wang, W.M. Yeung	4	A multi-faceted and automatic knowledge elicitation system (MAKES) for managing unstructured information	China	Research centre	x		x	x	x	x	x		x		Self-associated concept mapping (SACM) algorithm	Knowledge retrieval	x		Knowledge retrieval	x	x	x		x	

							GENERAL AREA								AI TOOLS			APPLICATIONS				ASPECTS OF KBE					
							KM																				
Id	Year	Author	R	Title	Location	Organisation	KBE	KBS	KA	KC	P	C	AI	ED	ML	Other	What technique?	Context	KBE App.	Shows Implem. ?	What App.?	Capture of Expert Knowledge	Access to knowledge by the tool	Automated knowledge extraction from data	Advice on AI tool suitability	Knowledge lifecycle management	AI extracted knowledge ready for reuse
th_75	2002	W.P. Wagner, et al.	2.5	Knowledge acquisition for expert systems in accounting and financial problem domains	USA	University			x									n/a									
th_79	1997	Harvey Dettmar, et al	3.5	Knowledge-based data generation	London	University		x	x					x	x		NN	Chemistry	x		chemical structures	x	x	x		x	x
th_83	2010	Pokojski, Jerzy Szustakiewicz, Karol Gil, Maciej	3	Engineering Knowledge Modeling in Design	Warsaw	University	x		x					x				Generic Design	x			x	x			x	

							GENERAL AREA								AI TOOLS			APPLICATIONS				ASPECTS OF KBE					
							KM																				
Id	Year	Author	R	Title	Location	Organisation	KBE	KBS	KA	KC	P	C	AI	ED	ML	Other	What technique?	Context	KBE App.	Shows Implem. ?	What App.?	Capture of Expert Knowledge	Access to knowledge by the tool	Automated knowledge extraction from data	Advice on AI tool suitability	Knowledge lifecycle management	AI extracted knowledge ready for reuse
th_87	2007	Ko, K.H. Pochiraju, K. Manoochehri, S.	3	An embedded system for knowledge-based cost evaluation of molded parts	Korea	Science Institute	x		x			x		x				Generic	x		Estimate MFG time of molded parts	x	x			x	
th_91	2010	Ruschitzka, Margot Suchodolski, Adam Wróbel, Jerzy	3	Ontology-Based Approach in Hybrid Engineering Knowledge Representation for Stamping Die Design	n/a	n/a	x		x			x		x				Metal Stamping	x		Semi-automated stamping die design	x	x			x	
th_95	2010	Kuhn, Olivier Ghodous, Parisa Dusch, Thomas Collet, Pierre	3.5	Collaboration for Knowledge-Based Engineering Templates Update	France & Germany	n/a	x		x			x		x				Generic	x		Update of knowledgebased engineering templates in the computer-aided design field	x	x			x	

							GENERAL AREA								AI TOOLS			APPLICATIONS				ASPECTS OF KBE					
							KM																				
Id	Year	Author	R	Title	Location	Organisation	KBE	KBS	KA	KC	P	C	AI	ED	ML	Other	What technique?	Context	KBE App.	Shows Implem. ?	What App.?	Capture of Expert Knowledge	Access to knowledge by the tool	Automated knowledge extraction from data	Advice on AI tool suitability	Knowledge lifecycle management	AI extracted knowledge ready for reuse
th_99	2014	Naranje, V. Kumar, S.	3	A knowledge based system for automated design of deep drawing die for axisymmetric parts	India	University	x		x			x		x				Manufacturing	x		Automated design of deep drawing dies	x	x			x	x
th_103	2002	Shehab, E. Abdalla, H.	3	An intelligent knowledge-based system for product cost modelling	Leicester (UK)	University	x		x			x		x				Manufacturing	x		Automated cost modeling of machining and injection mouling components	x	x			x	