A New Knowledge Sourcing Framework for Knowledge-Based Engineering: An aerospace industry case study

Quintana-Amate, S.,[2] Bermell-Garcia, P.,[2] Tiwari, A.,[1] Turner, C.J.[1]*

1: Cranfield University, Manufacturing department, School of Aerospace, Transport and Manufacturing (SATM), Cranfield University, Cranfield, Bedfordshire, UK

2: Airbus Group Innovations, 20A1 Ground South, Golf Course Lane, Filton, Bristol, United Kingdom

*Corresponding Author Email: c.j.turner@cranfield.ac.uk

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 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. The methodology proposed in this research is validated through the development and implementation of a case study involving the optimisation of wing design concepts at an Aerospace manufacturer. The results obtained proved the extended KBE capability for fast and effective knowledge sourcing. This evidence was provided by the experts working in the development of the case study through the implementation of structured quantitative and qualitative analyses.

Keywords: Knowledge-based engineering, Knowledge sourcing, Knowledge acquisition, Knowledge capture, Knowledge reuse

1. Introduction

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 limitations associated with some aspects of KBE that present a barrier to deliver the knowledge sourcing process requested by industry. To overcome these limitations this research proposes a new methodology for efficient knowledge capture and effective management of the complete knowledge life cycle.

The increasing competitiveness in the aerospace industry is forcing organisations to seek their profits beyond manufacturing. As a consequence, aerospace companies are changing from being providers of products to providers of Product Service Systems (PSS) (Selak et al. 2014). This new business model involves shifting efforts from manufacturing activities to those related to service systems such as Maintenance, Repair and Overhaul (MRO) (Stark et al., 2014), (Zhu et al., 2012).

This new trend adopted by aerospace organisations led the authors to the identification of two major risks for future manufacturing. The first risk is associated to the reduced availability of manufacturing 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 manufacturing 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 existing 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 traditionally been 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

2. Literature Review

Papers consulted as a background to this research were drawn from two key areas due to their relevance to Knowledge Based Engineering (KBE) and the remit of this work:

- Knowledge management in engineering design.
- Artificial intelligence for 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 (Pahl and Beitz, 1996), (Pugh, 1991), (Hubka and Eder, 1988), (Dieter, 2000), (Hales, 1993), (Pahl and Beitz, 1994). In the engineering design process there are four main steps (Pahl and Beitz, 1996), (Mendes et al., 2009) :

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 (Robinson, 2012).

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 economic feasibility are re-checked.

By observing 27 designers, while they were working, it was concluded that only 45% of the designers' time is dedicated to perform the steps of the engineering design process (Hales, 1986). Moreover, it was observed that 21% of designer's time was spent in tasks related with the management of engineering knowledge (Hales, 1986). These findings support the view of the engineering design process as a set of knowledge intensive activities (Chen et al., 2008).

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 (Chen et al., 2008). Knowledge management methods and tools have long been utilised to codify knowledge related to design issues (Hoegl & Schulze, 2005), (Huang, 2009), (Chapman & Pinfold, 1999). In the completion of this review a number of key concepts at the intersection between knowledge management and engineering design were identified and the requisite papers analysed in more detail. The concepts included: Knowledge capture; Knowledge modelling; Knowledge reuse.

In relation to the second key area of literature sought, artificial intelligence for knowledge management in engineering design, a number of papers were pertinent. 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 (Luger, 1999), (Deshpande, 2009), (Bellman, 1978), (Kurzweil, 1990), (Schalkoff, 1990). The role of AI in the knowledge management field can be summarised by the following two main contributions (Liebowitz, 2001):

- 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 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 efficiently, is the use of knowledge management methods combined with tools belonging to Artificial Intelligence (AI) (Guo et al. 2011), (Cheung et al., 2011), (Lau, 2009), (Palmer et al., 2011), (Ruiz et al., 2014), (Yang, 2002).

A further more focussed literature review allowed the authors to define the following functional roles linked to this research study:

- Capture of Expert Knowledge: Use of methods and tools to elicit knowledge from experts via interviews.

- Access to knowledge by KBE tool: implementations enabling the retrieval of information from an external data source that allows knowledge to be exploited by a KBE application.

- Automated knowledge extraction from data: Employment of tools able to reason over large datasets extracting engineering rules, constraints and other correlations from data.

- Advice on AI tool suitability: This refers to the use of applications to advise practitioners on the selection of AI tools to perform automated reasoning tasks.

- Knowledge lifecycle management: Presence of methods and tools to enable the holistic management of engineering knowledge throughout the life cycle of its use.

- 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).

The focussed literature search identified 63 research articles reporting on KBE implementations and classified within the functional roles defined by the authors.

Table 1 Research priorities analysis. (Quintana-Amate et al. 2015)

Functional roles	Priority	KBE experts assessment	No. of Papers	Literature Includes
Capture of Expert Knowledge:	Low	Not a well understood tool in the context of KBE; its use results in a costly process so automated approach would be beneficial.	62	(La Rocca, 2012), (Quintana-Amate et al., 2015), (Guo et al., 2011), (Chapman & Pinfold, 1999), (Cheung et al., 2011), (Lau et al., 2009), (Palmer et al., 2011), (Ruiz et al. 2014), (Yang, 2002), (Steenhuizen and van Tooren, 2012), (Yang et al., 2012), (Kochan, 1999), (Amadori et al. 2012), (Dettmar et al., 1998), (Dolšak and Novak, 2011), (Lou et al., 2004), (Naranje and Kumar, 2014), (Monticolo, et al. 2014)
Access to knowledge by KBE tool:	Med	Well understood tool by mainstream software development as interoperability. Realisation in KBE is limited.	57	(van der Elst and van Tooren, 2008), (Curi and Wang, 2013), (Jong et al. 2014), (Chapman and Pinfold, 2001), (Koini et al. 2009), (Naranje and Kumar, 2012), (Ko et al. 2007), (Choi, 2009), (Emberey and Milton, 2007), (Kumar and Singh, 2007), (La Rocca, 2012), (Quintana-Amate et al., 2015), (Guo et al., 2011), (Chapman & Pinfold, 1999)
Automated knowledge extraction from data:	High	A common tool in environments where a large amount of data exists but no formal knowledge models are available.	7	(Guo et al., 2011), (Cheung et al., 2011), (Lau et al., 2009), (Palmer et al., 2011), (Ruiz et al. 2014), (Yang, 2002), (Kumar, 2014)
Advice on AI tool suitability:	High	Choice of AI algorithms for engineering problems is based on the experience of experts.	0	
Knowledge lifecycle management:	Med	Changes to the knowledge used by KBE applications are common. Ability to modify engineering rule selection as an offline process has interest.	60	(La Rocca, 2012), (Quintana-Amate et al., 2015), (Guo et al., 2011), (Chapman & Pinfold, 1999), (Skarka, 2007), (Hunter et al., 2006), (Gardan and Gardan, 2003), (Wu and Shaw, 2011),

AI extracted knowledge ready for reuse:	High	Challenge in utilising extracted engineering knowledge.	(Lau et al., 2009), (Palmer et al., 2011), (Ruiz et al. 2014), (Yang, 2002), (Liao et al. 1999)

In order to support the research opportunities obtained through the analysis of the literature 6 experts from various domains working in different organisations were selected to assess the importance of the functional roles identified in this study. The outcome of the expert review is shown in 1. Special attention was paid by the authors to some of the papers included at this stage such as (La Rocca, 2012), (Bermell-Garcia, 2012), (Chapman & Pinfold, 1999) where the use of the three functional roles mentioned is clearly visible.

From Table 1 it can be seen that while two of the functional roles "Automated knowledge extraction from data" and "AI extracted knowledge ready for reuse" provided papers, no papers could be found to address the need for "Advice on AI tool suitability". Initially, a research gap on knowledge sourcing was identified as a response to the current limitations of KBE systems. The research gap found represented an opportunity to extend the capabilities of current KBE systems often used only as inference implementations. From literature it is clear that despite the benefits provided by KBE, it still fails to provide an efficient approach for expert knowledge capture. It was also evident that 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. Based on these findings, 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.

3. Knowledge Sourcing Platform

In the development of the knowledge sourcing platform three elements, identified from the literature review, form the core guidance for the architecture. The three elements are shown in Table 2.

Table 2 Core Elements of the Framework architecture

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

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. In this task the guidance of the publications (Segaran, 2007), (Hall et al., 2011), (Abu-Mostafa and Magdon-Ismail, 2012), (Bishop, 2007), (Murphy, 2012) proved valuable.

3.1 Machine learning methods

Two different ML libraries have been considered in this research: Weka and scikit-learn. The selection of the methods belonging to these libraries is based on:

- The accuracy of the results obtained in the learning process.
- The relevance of the rules 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. In this research only supervised methods were used due to the characteristics of the problems faced in the case study implemented. From the analysis of Table 3 shown below it is observed that explicit rules can be only extracted from the use of a single technique. The existence of methods to provide 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.

Table 3 Classification of common supervised machine learning methods.

Method	Solve Classification Problems	Solve Regression Problems	Explicit Rules: directly extracted	Level of rule's understanding	References
C4.5: Decision trees	Yes	No	Yes	Low / Medium	(Tso and Yau, 2007), (Elouedi et al. 2001)
Random Forest	Yes	Yes	No	-	(Chen and Ishwaran, 2012)
Naive Bayes	Yes	No	No	-	(Kang et al., 2012), (Pérez et al., 2009)
Logistic Regression	Yes	No	No	-	(Gusnanto et al., 2013), (Dreiseitl and Ohno-Machado, 2002)
Support Vector Machines	Yes	Yes	No	-	(Mountrakis and Ogole, 2011), (Barakat and Bradley, 2010)
Linear Regression	Yes	Yes	Yes	It depends on the problem's complexity	(Koç and Barkana, 2014)
Gaussian Processes	No	Yes	No	-	(Mackay, 1998)
Rule methods	No	Yes	Yes	Medium / High	(Holmes et al., 1999), (Quinlan,

(M5R)					1992), (Wang and Witten, 1996)
Neural Networks	Yes	Yes	No	-	(Barakat and Bradley, 2010), (Paliwal and Kumar, 2009)
Regression Trees (REP)	Yes	Yes	Yes	Low / Medium	(Portnoy and Koenker, 1997)

3.2 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 for the integration of KBE applications into engineering workflows, thus facilitating the adoption of the proposed framework in industry.

Over the last 20 years, several methodologies supporting the development and maintenance of Knowledge Based Systems (KBS) have been realised. In this context, three methodologies were analysed and one was selected as the most suitable to be adopted in this research. 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 cycle steps which are classified within 3 different stages as shown in Table 4.

Table 4 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

|	Formalise	elicitation methods and representation of the knowledge captured in a formal, consistent and standard manner.
Stage 3	Package	Design and implementation of the KBE system.
	Distribute	
	Introduce	
	Use	

MOKA and CommonKADS have been acknowledged by the research community as the most relevant methodologies when talking about KBE development (Verhagen, 2013). However there are some challenges remaining related to a number of limitations as described in Verhagen (2013). In this context, CommonKADS provides the user with a set of guidelines and templates to complete the tasks so improving repeatability. Nevertheless, the use of these guidelines and templates decreases its flexibility of the CommonKADS approach. Users also complain that there are not enough templates to cover most of their current tasks. The main drawback of MOKA is identified in Verhagen (2013) 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 inability to deal with knowledge change, namely being unable to account for knowledge origin and the repercussions in its transformation.

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 research 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 for this study is described below.

3.3. KNOMAD Methodology

As with MOKA and CommonKADS, KNOMAD (Knowledge Nurture for Optimal Multidisciplinary Analysis and Design) is a methodology created to support the development of KBE systems. 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 (detailed in Verhagen (2013)) and illustrated in Fig. 1) :

- Knowledge capture.
- Normalisation.
- Organisation.
- Modelling.
- Analysis.
- Delivery.
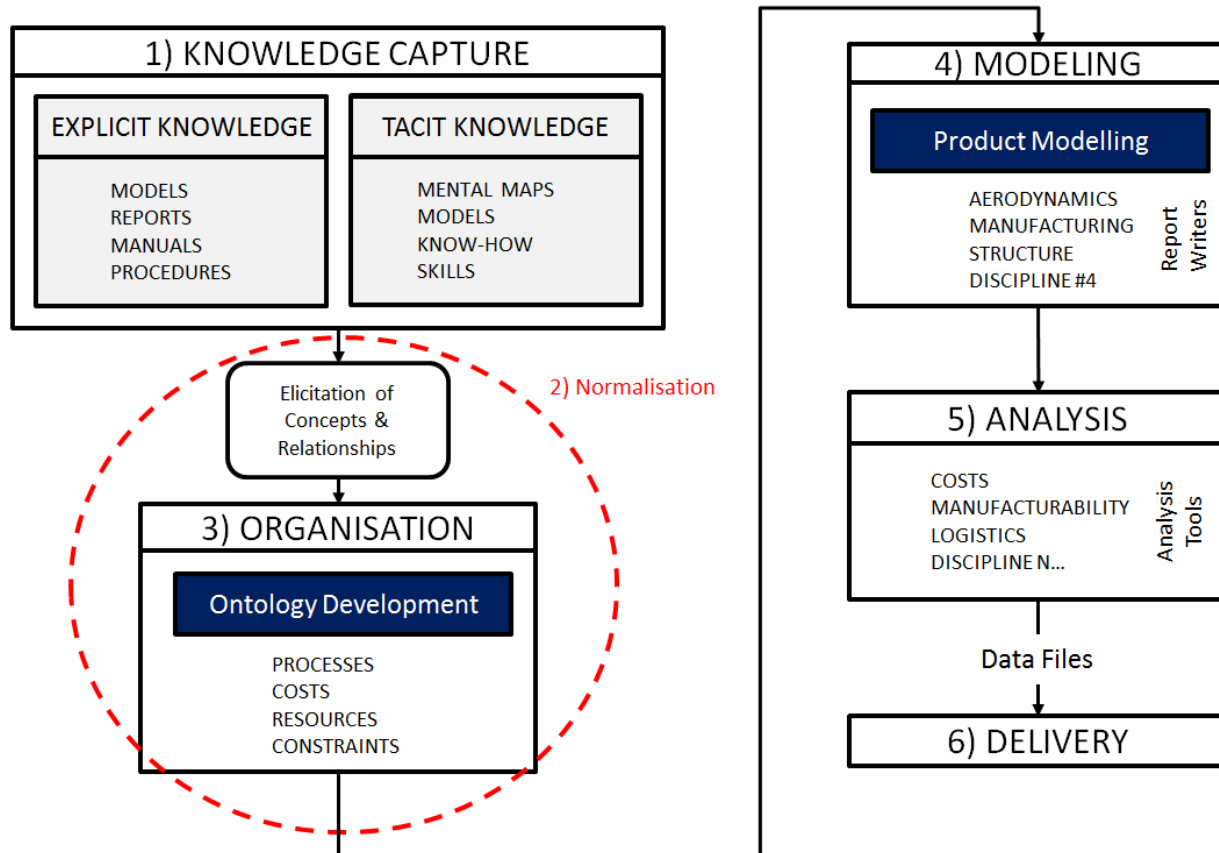
Fig 1. KNOMAD methodology by Verhagen (2013).

KNOMAD has been selected as the KBE methodology for use in this research as it integrates the best practices employed by MOKA and CommonKADS. KNOMAD uses an ontology to account for knowledge change, improving its applicability and transparency. There are two areas in which this methodology is appreciably superior to others:

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 enables machine learning algorithms to deliver predictions with increased accuracy.

To improve knowledge transparency and support knowledge change and traceability, conceptual models were used to create a KLC ontology. This ontology was created in Verhagen (2013) 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 similarities in research context.

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 Knowledge Sourcing Framework (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 captured knowledge.

- 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 required to automate the Weka and scikit-learn libraries respectively.

- Machine learning libraries. Two different ML libraries have been considered in this research: Weka and scikit-learn. 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 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.
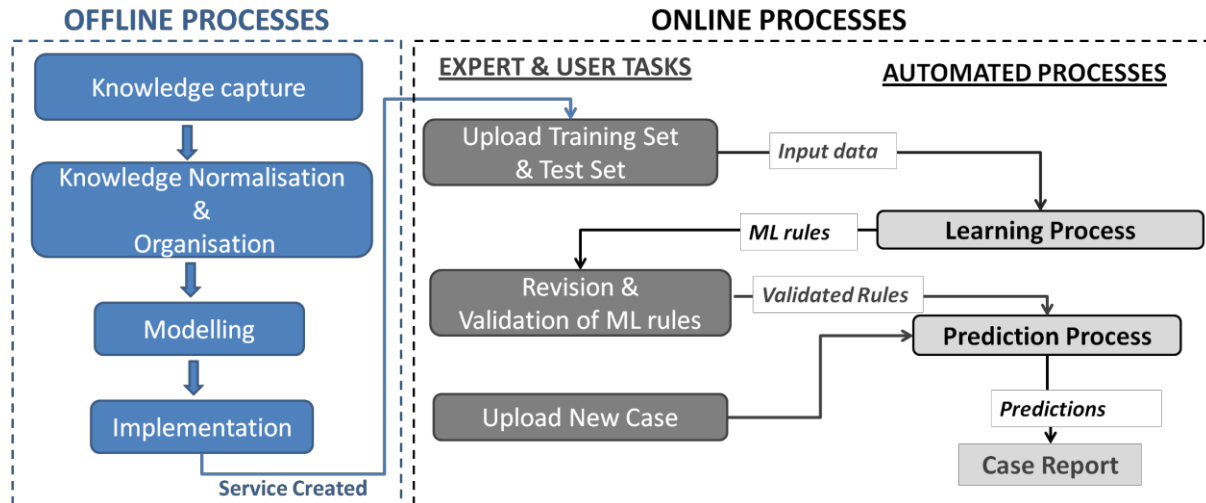
Fig 2. Service creation within KBE Platform: Process flow.

The framework proposed was developed using a web-based open source content management system built on top of a database. This system contains two different interfaces named as "administrator interface" and "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 (Fig. 2). 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 (this will be demonstrated later on in this paper with a manufacturing based case study drawn from the aerospace industry).

Offline and Online Phases review

In summary, the capture and pre-processing of the data together with the creation of the EKR 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.

In summary the main contributions of this framework are:

- Efficient creation of new knowledge. – AI algorithms are utilised in the capture of knowledge from experts in a more time efficient manner (in addition the framework enables the user to select the most appropriate algorithm for the problem to be solved)

- Enhanced KBE Reliability – as every aspect of the model generated by the tool may be reviewed the generation process may be fine-tuned by users

- Reduced reliance on experts – knowledge is captured once then reused many times with only limited time involvement of experts

- Effective knowledge reuse – the domain ontology utilised allows for the creation of knowledge packages which may be stored and reused at a later stage

- Integration of KBE tools into engineering workflows – the tool's integration into a content management system allows for its use across a wide range of engineering problems

4. Case study: Predicting Manufacturing Cycle Time in the Aerospace Industry

The case study has been developed in the context of an aerospace organisation with the aim of improving an existing cost modelling capability by realising a more efficient source of expert knowledge.

The cost modelling tool was built to provide Airbus with improved manufacturing systems supporting the conceptual design evaluation of wing designs; in summary, 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, composed of more than 300 driving parameters and in excess of 1000 intermediate calculations, resulted 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 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 (Fig. 3); converting 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.
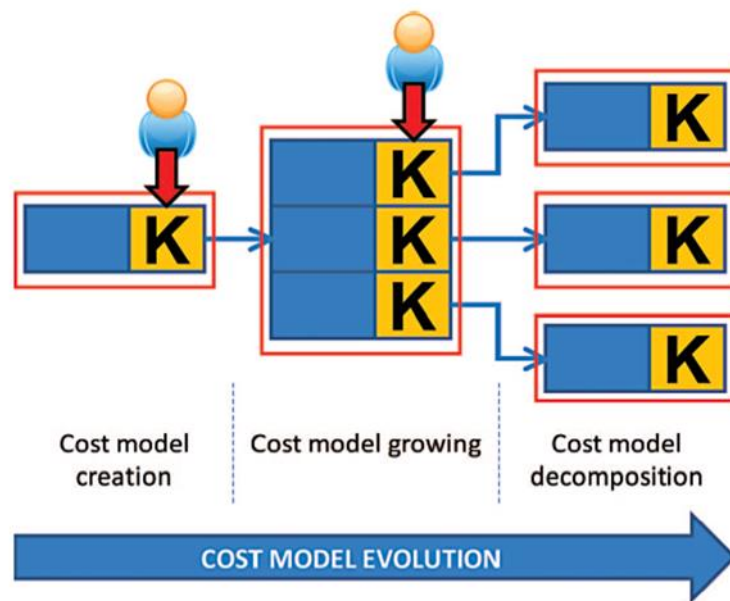
Fig. 3. Cost model evolution

From previous research into KBE within the aerospace industry the following limitations were identified as detailed Table 5.

Table 5 limitations regarding the identified KBE challenges

KBE challenge	Limitations associated with Case 1 (*Encoding expert knowledge*)	Limitations associated with Case 2 (Decoding knowledge from a "black box" capability)
Need for a more generic and reliable methodology	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	No expert involvement in the validation process. This characteristic is the main cause of the capability's low reliability.
A more efficient way of sourcing knowledge (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.	

4.1 Case Study: Manufacturing Cycle Time Prediction of Composite Wing Covers

This use case intends to carry out the design optimisation of wing covers made of Carbon Fibre Reinforced Plastics (CFRP). To do this, a Knowledge Sourcing platform has been developed enabling the design improvement of composite wing covers by predicting the manufacturing time of different wing design concepts. 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 informed decisions reducing the time required to select an optimal design.

From this perspective, the case study shows 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 case is the aerospace industry, and more precisely the lay-up process involved in the manufacture of CFRP wing covers.

The general context of this use case is the development of a new aircraft generation (Airbus A30X) that will substitute the current A320 model. Although manufacturing the A30X model will not commence for another 20 years, 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:

To obtain an optimal design in the shortest amount of time (decreasing the number of design iterations required to achieve it), decisions must be made 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.

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 scrap material 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 of 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 manufacturing 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.

In order to achieve an effective and more efficient sourcing of engineering knowledge, the solution to be developed must meet the following requirements:

- It must be simple enough to foster further use.
- Deliver fast and accurate predictions allowing the optimisation of a wing cover design.
- Enable the analysis of the knowledge and results obtained by the user to increase transparency and reliability.

4.2 The Application

The aim of this case study is to permit the user to efficiently perform the process of predicting the manufacturing cycle time of composite wing covers supporting the selection of an optimal design, saving time and costs. The approach proposed is realised in the form of a web-based platform, facilitating access to different teams based in multiple locations. The process of predicting the manufacturing cycle time of wing covers is modelled in Fig. 4 using IDEF0 representation. The illustration in Fig. 4 is focused on the

overall process of presenting the requirements to obtain the manufacturing cycle time predictions. These requirements are:

Input data from designs and simulations. This refers to design data corresponding to a set of design descriptors defined by experts, and the machine time output values provided by a simulation software application.

Resources in the form of expert involvement. Experts are required to initially define the parameters they believe are driving the manufacturing time. 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, selection of the machine learning process (including learning and prediction activities) and the advanced visualisation of the results provided by the knowledge sourcing capability.
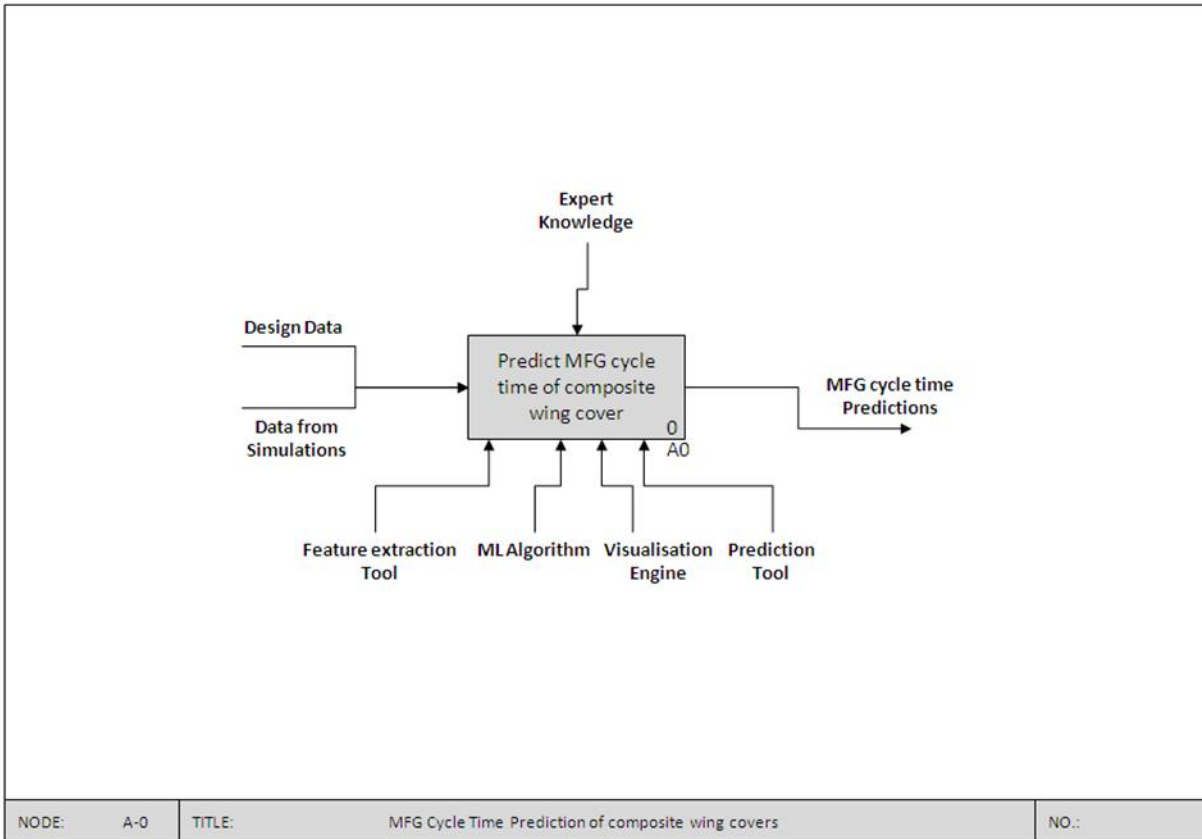
Fig. 4. IDEF0 - Main process.

The main process is divided in six sub-tasks as observed in Fig. 5. As a starting point the identification of the design descriptors that drive the manufacturing cycle times are required. To do that, experts must intervene using their experience and input data (simulation and design data) to understand and identify which parameters, from the wing cover design, affect the manufacturing time the most. As a result, a list of design descriptors driving manufacturing times were defined and stored within the Content Management System (CMS). A feature extraction tool then generated 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 manufacturing time values (output values) from simulation log files.

The "Training Set" file is employed by the machine learning algorithm to create a set of rules describing the system behaviour.
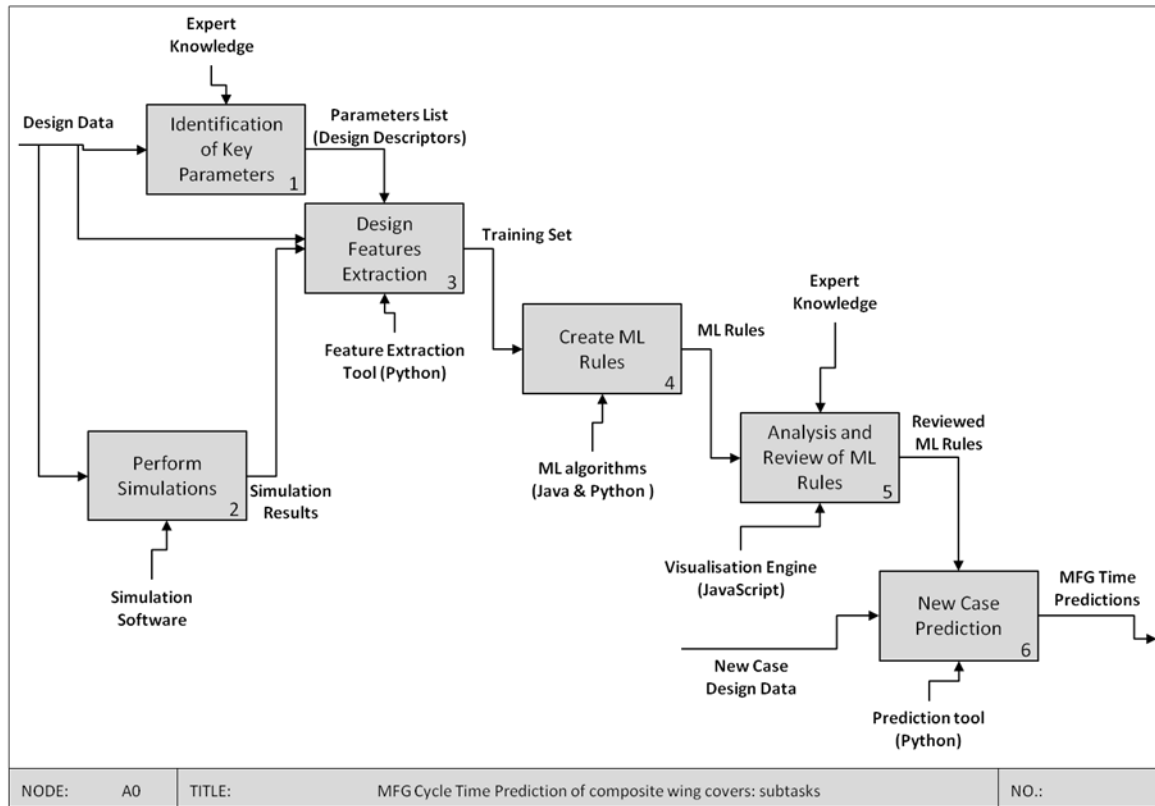
Fig. 5. IDEF0 Subtasks.

4.3 Adoption of KNOMAD methodology for the case study

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 to provide 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 Fig. 6.

```
┌─────────────────┐     ┌─────────────────┐     ┌─────────────────┐     ┌─────────────────┐     ┌─────────────────┐
│ 1) KNOWLEDGE    │     │ 2) NORMALISATION│     │ 3)ORGANISATION  │     │ 4) MODELLING /  │     │ 5) ANALYSIS AND │
│    CAPTURE      │     │                 │     │                 │     │  IMPLEMENTATION │     │    DELIVERY     │
├─────────────────┤     ├─────────────────┤     ├─────────────────┤     ├─────────────────┤     ├─────────────────┤
│ -SCOPE          │ ──▶ │ -DATA QUALITY   │ ──▶ │ -DOMAIN         │ ──▶ │ -EKR MODELS     │ ──▶ │ -SOLUTION       │
│ -JUSTIFICATION  │     │  CHECK          │     │  ONTOLOGY       │     │ -CMS MODELS     │     │  PERFORMANCE    │
│ -PARAMETERS     │     │                 │     │  DEVELOPMENT    │     │ -KNOWLEDGE      │     │  STUDY          │
│ -PROCESS SUBTASKS│    │ -INPUT AND OUTPUT│    │                 │     │  ARCHITECTURE   │     │ -QUANTITATIVE   │
│ -INPUTS         │     │  FORMAT         │     │                 │     │ -CMS REPORTS    │     │  AND QUALITATIVE│
│ -OUTPUTS        │     │  STANDARISATIONS│     │                 │     │                 │     │  ANALYSIS       │
└─────────────────┘     └─────────────────┘     └─────────────────┘     └─────────────────┘     └─────────────────┘
```

- 1) KNOWLEDGE CAPTURE
  - -SCOPE
  - -JUSTIFICATION
  - -PARAMETERS
  - -PROCESS SUBTASKS
  - -INPUTS
  - -OUTPUTS

- 2) NORMALISATION
  - -DATA QUALITY CHECK
  - -INPUT AND OUTPUT FORMAT STANDARISATIONS

- 3) ORGANISATION
  - -DOMAIN ONTOLOGY DEVELOPMENT

- 4) MODELLING / IMPLEMENTATION
  - -EKR MODELS
  - -CMS MODELS
  - -KNOWLEDGE ARCHITECTURE
  - -CMS REPORTS

- 5) ANALYSIS AND DELIVERY
  - -SOLUTION PERFORMANCE STUDY
  - -QUANTITATIVE AND QUALITATIVE ANALYSIS

- 3.1) DOMAIN ONTOLOGY: PRODUCT EKR
  - -PRODUCT STRUCTURE DEFINITION
  - -PRODUCT RELATIONS

- 3.2) DOMAIN ONTOLOGY: RESOURCE EKR
  - -RESOURCE STRUCTURE DEFINITION
  - -RESOURCE RELATIONS

- 3.3) DOMAIN ONTOLOGY: PROCESS EKR
  - -PROCESS STRUCTURE DEFINITION
  - -PROCESS RELATIONS

- 4.1) MODELLING
  - -EKR MODELS
  - -CMS MODELS

- 4.2) IMPLEMENTATION
  - -KNOWLEDGE ARCHITECTURE DEVELOPMENT
 - -CREATION OF CMS REPORTS

Fig. 6. KNOMAD - process flow

The knowledge used to obtain the manufacturing cycle time predictions is sourced by experts and Machine Learning (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 two experts with more than 15 years of experience on KBE systems.

Once the design descriptors are defined (Fig. 7), an application was developed to enable the extraction (from wing 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 manufacturing cycle time obtained from simulations. The ML algorithm searches for data correlations using the "Training Set" producing as a result a set of rules or an explicit model. Finally, the explicit model is captured and stored within the knowledge repository.
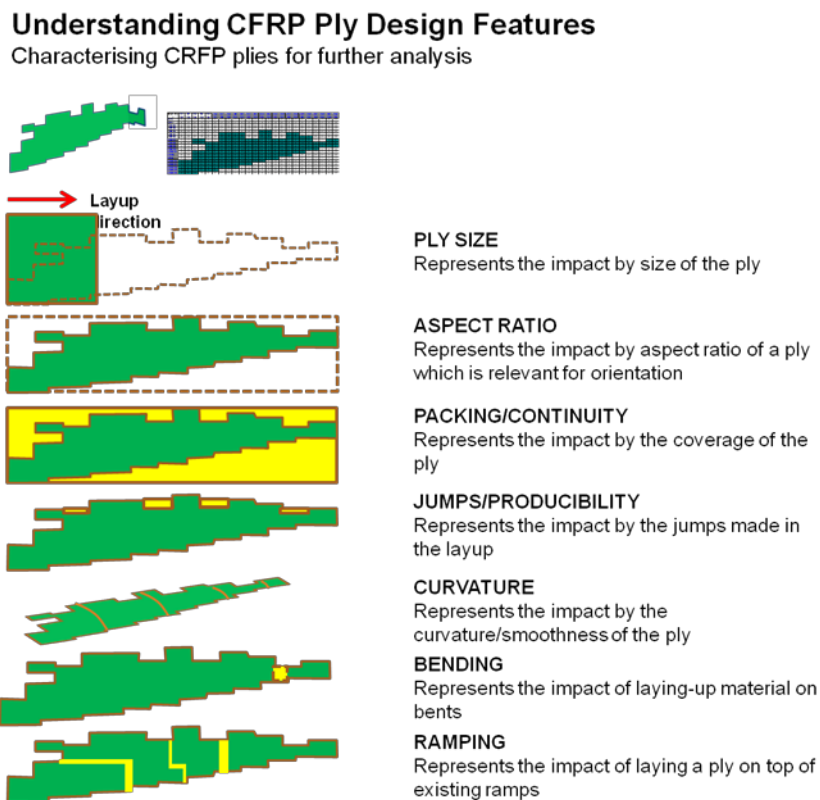
Fig. 7. List of Design Descriptors

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 KNOMAD ontology.

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

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 understandable by automated search applications, thus enhancing the accessibility and traceability of the knowledge stored while facilitating its update. KNOMAD fosters the use of an ontology to achieve the knowledge architecture. In this regard, a domain specific ontology is constructed in accordance with the main Knowledge Life Cycle (KLC) ontology 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".

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

Modelling: The EKR's created under the scope of this study have been annotated and modelled using the specific domain ontology. 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 wing cover 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.

4.4 Implementation

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. The elements belonging to the "Process" EKR class are integrated within the system architecture as shown in Fig. 8.
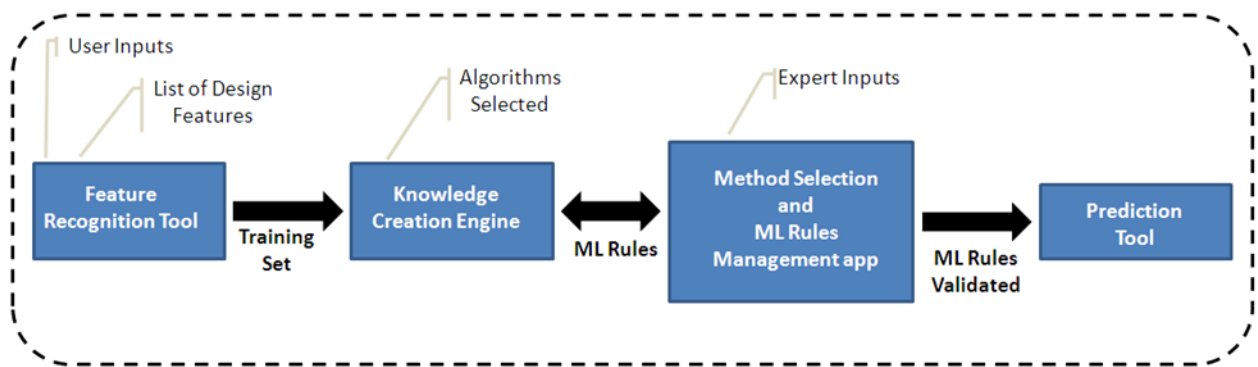
Fig. 8. KBE system architecture

5. Results and Discussion

Prior to obtaining 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 manufacturing time prediction of CFRP wing covers). 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.

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. In this research, the evaluation of the prediction accuracy of the ML algorithms follows a common criterion acknowledged by the research community (Tüfekci, 2014) (Kavaklioglu, 2011). 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 the problem domain reviewed the ML scoring values (MAE and RMSE) and the level of understanding of each explicit

model created. Table 6 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 6 Explicit models displayed by the KSF platform

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 these types of algorithm 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.

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 such a tool, employing charts and tables, allowed experts to:

- Identify trends in the 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 the following (and shown in Fig. 9):

- 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 Cross Validation (CV) process are acceptable from an engineering point of view. The meaningfulness of the model is also a key factor to be taken 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.
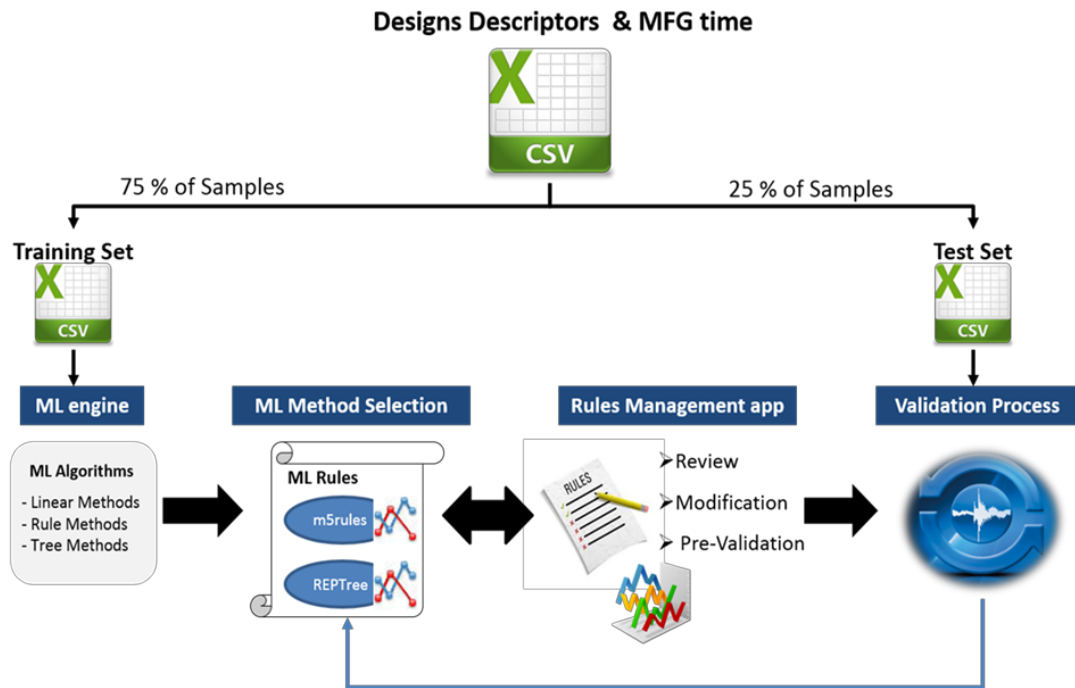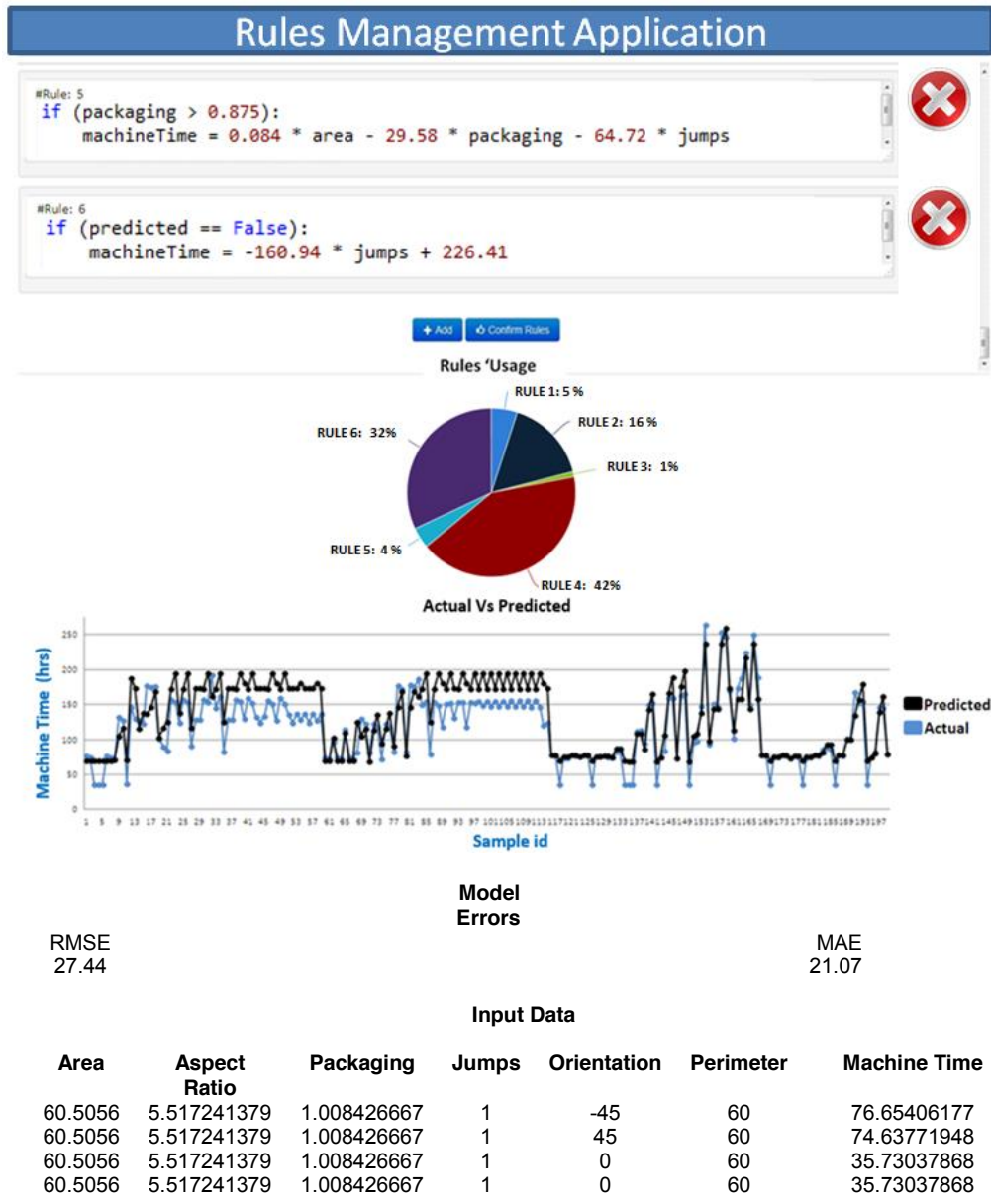
Fig. 9. Procedure followed to validate the model

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 7). 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 Fig. 10 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 Fig. 11 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.

Table 7. 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

Rules Management Application

```
#Rule: 5
if (packaging > 0.875):
    machineTime = 0.084 * area - 29.58 * packaging - 64.72 * jumps
```

```
#Rule: 6
if (predicted == False):
    machineTime = -160.94 * jumps + 226.41
```

+ Add   ↻ Confirm Rules

Rules 'Usage

RULE 1: 5 %
RULE 2: 16 %
RULE 3: 1%
RULE 4: 42%
RULE 5: 4 %
RULE 6: 32%

Actual Vs Predicted

■ Predicted
■ Actual

Model Errors

RMSE 27.44   MAE 21.07

Input Data

Area	Aspect Ratio	Packaging	Jumps	Orientation	Perimeter	Machine Time
60.5056	5.517241379	1.008426667	1	-45	60	76.65406177
60.5056	5.517241379	1.008426667	1	45	60	74.63771948
60.5056	5.517241379	1.008426667	1	0	60	35.73037868
60.5056	5.517241379	1.008426667	1	0	60	35.73037868

Fig. 10. R   esults generated in the CV process by non-reviewed rules
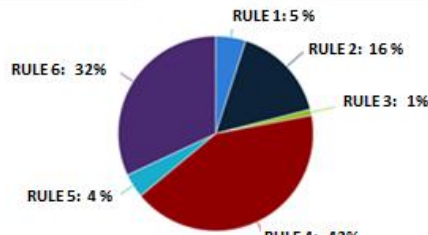
Rules Management Application

```
#Rule: 5
if (packaging > 0.875):
    machineTime = 0.084 * area - 29.58 * packaging - 64.72 * jumps

elif (area <3 and packaging < 0.21 and orientation != -45):
    machineTime = machineTime * 1.06

elif (area < 3 and packaging > 0.3 and orientation == -45):
    machineTime = machineTime * 0.84
```
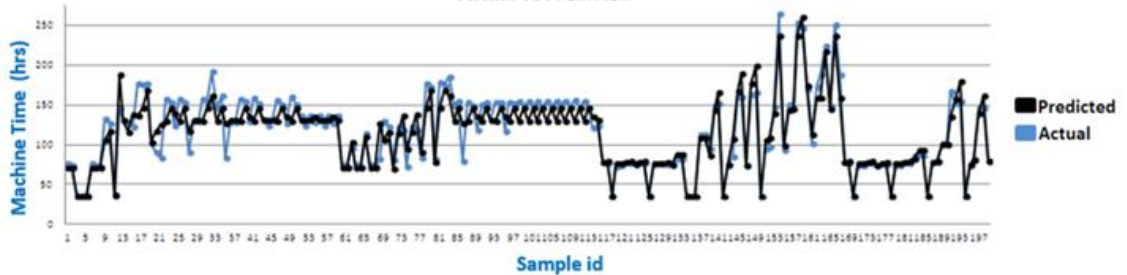
+ Add | **↻ Confirm Rules**

Rules 'Usage

RULE 1: 5 %
RULE 2: 16 %
RULE 3: 1%
RULE 4: 42%
RULE 5: 4 %
RULE 6: 32%

Actual Vs Predicted

Machine Time (hrs) vs Sample id

■ Predicted
■ Actual

Model Errors

RMSE	MAE
27.44	21.07

Input Data

Area	Aspect Ratio	Packaging	Jumps	Orientation	Perimeter	Machine Time
60.5056	5.517241379	1.008426667	1	-45	60	76.65406177
60.5056	5.517241379	1.008426667	1	45	60	74.63771948
60.5056	5.517241379	1.008426667	1	0	60	35.73037868
60.5056	5.517241379	1.008426667	1	0	60	35.73037868

Fig. 11. Results generated in the CV process using the rules reviewed and modified by experts

5.1 Key findings from the case study

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 with the aim of 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 (Mlynarski, 2006) (Alcala-Fdez, 2009).

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). 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 manufacturing cycle time prediction problem.

The models automatically generated in the Cross Validation (CV) process did not contain low error values as shown in Table 6. 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 increased the reliability of the model. MAE and RMSE values obtained using the "Test Set"

were lower than 20 units, taking into account the maximum allowance for MAE and RMSE (using a confidence level of 95%).

6. Conclusions

The creation of a service for fast prediction of manufacturing cycle times of wing composite structures has been presented in this paper. The work reported through the implementation of this service proves the foundations for the delivery of a more efficient approach for sourcing engineering knowledge.

Prior the development of this research, knowledge associated with the manufacturing cycle time estimation problem was kept in the minds of the experts or in documents locally stored. These practices provoked knowledge leaks in the case of experts leaving the organisation. Therefore, it became apparent that there was a need for an effective approach for managing engineering knowledge. In this regard, the research presented in this paper proved to be an effective solution for managing engineering knowledge by providing experts with a methodology fostering the systematic capture, retention and reuse of knowledge. Using the proposed framework, experts are now asked to store the knowledge used in their particular practice in a human readable format within independent knowledge packages placed in a centralised web-based platform. In doing so, knowledge becomes accessible for application to different engineering problems across the organisation.

In parallel, the previous approach used for time estimation lacks a model of the problem due to complexity issues. Much of the complexity in the process is due to the use of new technologies, with highly interdependent parts, 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. 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 capable of efficiently capturing relevant and meaningful knowledge from company data assets and experts. 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 manufacturing 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 work to generate the machine time estimates. Therefore, the knowledge sourcing framework developed is considered more efficient compared to current approaches as it provides faster manufacturing time estimations with accuracy levels considered as acceptable by experts in the domain.

In future work 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. Further work is also planned with the aim of facilitating further 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. The use of unstructured databases (in common use with big data sets) would allow the KSF to deal with larger amounts of data, composed of both structured and non-structured information sets such as the contents of manuals and emails respectively.

References

Abu-Mostafa, Y.S., Magdon-Ismail, M. & Lin, H.T., (2012). *Learning from data* (Vol. 4). Singapore: AMLBook.

Alcala-Fdez, J., Sanchez L., Garcia S., del Jesus M.J., Ventura S., Garrell J.M., Otero J., Romero C., Bacardit J., Rivas VM., & Fernandez J.C. (2009). KEEL: A software tool to assess evolutionary algorithms for data mining problems. *Soft Computing*, 13 (3), 307–318.

Amadori, K., Tarkian, M., Ölvander, J. & Krus, P., (2012), Flexible and robust CAD models for design automation. *Advanced Engineering Informatics*, 26 (2), 180–195.

Barakat, N., & Bradley, A. P. (2010). Neurocomputing Rule extraction from support vector machines: A review. *Neurocomputing*, 74(1–3), 178–190.

Bellman, R. E. (1978). *An introduction to artificial intelligence: Can computers think?* San Francisco: Boyd & Fraser Pub. Co.

Bermell-Garcia, P., Verhagen, W.J., Astwood, S., Krishnamurthy, K., Johnson, J.L., Ruiz, D., Scott, G. & Curran, R., (2012). A framework for management of Knowledge-Based Engineering applications as software services: Enabling personalization and codification. *Advanced Engineering Informatics*, 26 (2), 219–230.

Bishop, M. (2007). *Pattern Recognition and Machine Learning*. Berlin: Springer.

Chapman, C., & Pinfold, M., (1999). Design engineering—a need to rethink the solution using knowledge based engineering. *Knowledge-Based Systems*, 12 (5–6), 257–267.

Chapman C. B., & Pinfold, M. (2001). The application of a knowledge based engineering approach to the rapid design and analysis of an automotive structure. *Advances in Engineering Software*, 32 (12), 903–912.

Chen, X., & Ishwaran, H. (2012), Random forests for genomic data analysis, *Genomics*, 99 (6), 323–9.

Chen, Y.J., Chen, Y.M., Chu, H.C., & Kao, H.Y., (2008). On technology for functional requirement-based reference design retrieval in engineering knowledge management. *Decision Support Systems*, 44 (4), 798-816.

Cheung, C.F., Lee, W.B., Wang, W.M., Wang, Y., & Yeung, W.M., (2011). A multi-faceted and automatic knowledge elicitation system (MAKES) for managing unstructured information. *Expert Systems with Applications*, 38 (5), 5245–5258.

Choi, J., (2009). Architecture of a knowledge based engineering system for weight and cost estimation for a composite airplane structures. *Expert Systems with Applications*, 36 (8), 10828–10836.

Cui, J., & Wang, D. (2013). Application of knowledge-based engineering in ship structural design and optimization. *Ocean Engineering*, 72, 124–139.

Deshpande, N. (2009). *Artificial Intelligence*. 2nd Edition. Pune: Technical Publications.

Dettmar, H., Liu, X., Johnson, R., & Payne, A., (1998). Knowledge-based data generation. *Knowledge-Based Systems*, 11 (3–4), 167–177.

Dieter, E. (2000). *Engineering Design*. New York: McGraw-Hill.

Dolšak, B., & Novak, M., (2011). Intelligent decision support for structural design analysis. *Advanced Engineering Informatics*, 25 (2), 330–340.

Dreiseitl, S., & Ohno-Machado, L. (2002). Logistic regression and artificial neural network classification models: a methodology review. *Journal of biomedical informatics*, 35, (5–6), 352–359.

Elouedi, Z., Mellouli, K., & Smets, P., (2001). Belief decision trees: theoretical foundations. *International Journal of Approximate Reasoning*, 28, 91–124.

Emberey, C. L., & Milton, N. R. (2007). Application of Knowledge Engineering Methodologies to Support Engineering Design Application Development in Aerospace, In *Proceedings of 7th AIAA Aviation Technology, Integration and Operations Conference (ATIO)*, 18 - 20 September, Belfast, Northern Ireland.

Gardan, N., & Gardan, Y. (2003). An application of knowledge based modelling using scripts. *Expert Systems with Applications*, 25 (4), 555–568.

Guo, Y., Hu, J., & Peng, Y., (2011). Research on CBR system based on data mining. *Applied Soft Computing Journal*, 11 (8), 5006–5014.

Gusnanto, A., Ploner, A., Shuweihdi, F., & Pawitan, Y. (2013). Partial least squares and logistic regression random-effects estimates for gene selection in supervised classification of gene expression data. *Journal of biomedical informatics*, 46 (4), 697-709.

Hales, C. (1986). Analysis of the engineering design process in an industrial context, *Unpublished*. University of Cambridge, UK.

Hales, C. (1993). *Managing Engineering Design*. London: Longman Scientific & Technical.

Hall, M., Witten, I., & Frank, E., (2011). *Data Mining: Practical Machine Learning Tools and Techniques*, 3rd Edition. New York: Morgan Kaufmann.

Hoegl, M., & Schulze, A., (2005). How to Support Knowledge Creation in New Product Development: An Investigation of Knowledge Management Methods. *European Management Journal*, 23 (3), 263–273.

Holmes, G., Hall, M., & Prank, E. (1999). Generating Rule Sets from Model Trees, In *Advanced Topics in Artificial Intelligence*, *1747*, Berlin:Springer, 1 – 12.

Huang, J. J., (2009). The evolutionary perspective of knowledge creation – A mathematical representation. *Knowledge-Based Systems*, 22 (6), 430–438.

Hubka, V., & Eder, W. E. (1988). *Theory of technical systems: a total concept theory for engineering design*. Berlin: Springer-Verlag.

Hunter, R., Rios, J., Perez, J.M., & Vizan, A., (2006). A functional approach for the formalization of the fixture design process. *International Journal of Machine Tools and Manufacture*, 46 (6), 683–697.

Jong, W.R., Ting, Y.H., & Li, T.C. (2014). Application of knowledge-based engineering for automated slide design. *The International Journal of Advanced Manufacturing Technology*, 74 (5-8), 637-651

Kang, H., Yoo, S.J., & Han, D. (2012). Senti-lexicon and improved Naïve Bayes algorithms for sentiment analysis of restaurant reviews. *Expert Systems with Applications*, 39 (5), 6000-6010.

Kavaklioglu, K. (2011). Modeling and prediction of Turkey's electricity consumption using Support Vector Regression, *Applied Energy*, 88 (1), 368-375.

Ko, K., Pochiraju, H., K., & Manoochehri, S. (2007). An embedded system for knowledge-based cost evaluation of molded parts. *Knowledge-Based Systems*, 20 (3), 291–299.

Koç M., & Barkana, A. (2014). Application of Linear Regression Classification to low-dimensional datasets. *Neurocomputing*, 131, 331–335.

Kochan, A. (1999). Jaguar uses knowledge-based tools to reduce model development times, *Assembly Automation*, 19(2), 114-117.

Koini, G.N., Sarakinos, S.S., & Nikolos, I.K., (2009). A software tool for parametric design of turbomachinery blades. *Advances in Engineering Software*, 40 (1), 41–51.

Kumar, S. (2014). A knowledge based reliability engineering approach to manage product safety and recalls. *Expert Systems with Applications*, 41 (11), 5323–5339.

Kumar, S., & Singh, R. (2004). A low cost knowledge base system framework for progressive die design. *Journal of Materials Processing Technology*, 153–154, 958–964.

Kumar, S., & Singh, R. (2007). A knowledge-based system to automate the selection of progressive die components. *International Journal of Computational Materials Science and Surface Engineering*, 1, 85–96.

Kurzweil, R. (1990). *The age of intelligent machines*. Boston: MIT Press.

La Rocca, G. (2012). Knowledge based engineering: Between AI and CAD. Review of a language based technology to support engineering design. *Advanced Engineering Informatics*, 26 (2), 159–179.

Lau, H.C.W., Ho, G.T.S., Zhao, Y., & Chung, N.S.H. (2009). Development of a process mining system for supporting knowledge discovery in a supply chain network. *International Journal of Production Economics*, 122 (1), 176–187.

Liao, T.W., Zhan, Z.H., & Mount, C.R., (1999). An integrated database and expert system for failure mechanism identification: Part I, automated knowledge acquisition, *Engineering Failure Analysis*, 6 (6), 387-406.

Liebowitz, J. (2001). Knowledge management and its link to artificial intelligence, *Expert Systems with Applications*, 20 (1), 1–6.

Lou, Z., Jiang, H., & Ruan, X., (2004). Development of an integrated knowledge-based system for mold-base design. *Journal of Materials Processing Technology*, 150 (1–2), 194–199.

Luger, G. F. (1999). *Artificial Intelligence: Structures and Strategies for Complex Problem Solving*. Boston: Addison-Wesley Publishing Company.

Mackay, D. (1998). *Introduction to gaussian processes, in Neural Networks and Machine Learning*, Bishop, C. M. (Ed.) Berlin: Springer, 133–165.

Mendes, L.A., Back, N., & Oliveira, G.H. (2009). Designing automated test systems: An adapted methodology inspired on Pahl and Beitz ' s systematic approach. *Robotics and Computer-Integrated Manufacturing*, 25 (6), 945–950.

Mlynarski, R., Ilczuk, G., Wakulicz-Deja, A., & Kargul, W. (2006). A new method of data preparation for cardiological decision support, in *2006 Computers in Cardiology*, Valencia, 17-20 Sept. 2006, pp. 273-276.

Monticolo, D., Mihaita, S., Darwich, H., & Hilaire, V. (2014). An agent-based system to build project memories during engineering projects. *Knowledge-Based Systems*, 68, 88-102.

Mountrakis, G., Im, J., & Ogole, C. (2011). Support vector machines in remote sensing: A review. *ISPRS Journal of Photogrammetry and Remote Sensing*, 66 (3), 247-259.

Murphy, K. (2012). *Machine Learning - A Probabilistic Perspective*. Boston: MIT Press.

Naranje, V., & Kumar, S. (2012). A Knowledge Based System for Selection of Components of Deep Drawing Die. *American Journal of Intelligent Systems*, 2 (2), 1–11.

Naranje, V., & Kumar, S. (2014). A knowledge based system for automated design of deep drawing die for axisymmetric parts. *Expert Systems with Applications*, 41 (4), 1419–1431.

Pahl, G., & Beitz, W. (1994). *Engineering Design*. Design Council, London:Springer, 12, 221-226.

Pahl. G., & Beitz, W., (1996). Engineering Design. A Systematic Approach. Springer-Verlag,.

Paliwal, M., & Kumar, U. A. (2009). Neural networks and statistical techniques: A review of applications, *Expert Systems with Applications*, 36 (1), 2–17.

Palmer, C., Harding, J.A., Swarnkar, R., Das, B.P., & Young, R.I., (2011). Generating rules from data mining for collaboration moderator services. *Journal of Intelligent Manufacturing*, 24 (2), 313–330.

Pérez, A., Larrañaga, P., & Inza, I., (2009). Bayesian classifiers based on kernel density estimation: Flexible classifiers. *International Journal of Approximate Reasoning*, 50 (2), 341-362.

Portnoy, S., & Koenker, R. (1997). The Gaussian hare and the Laplacian tortoise: computability of squared-error versus absolute-error estimators. *Statistical Science*, 12 (4), 279 – 300.

Pugh, S. (1991). *Total design: Integrated methods for successful product engineering*. Boston: Addison Wesley Publishing Company.

Quinlan, R. (1992). Learning with continuous classes. In *Proceedings of the 5th Australian Joint Conference on Artificial Intelligence*, 16 -18 November, Hobart, Australia. 343 – 348.

Quintana-Amate, S., Bermell-Garcia, P. and Tiwari, A., 2015. Transforming expertise into Knowledge-Based Engineering tools: A survey of knowledge sourcing in the context of engineering design. Knowledge-Based Systems, 84, pp.89-97.

Robinson, M. A. (2012). How design engineers spend their time: Job content and task satisfaction. *Design Studies*, 33 (4), 391–425.

Ruiz, P.P., Foguem, B.K., & Grabot, B., (2014). Generating knowledge in maintenance from Experience Feedback. *Knowledge-Based Systems*, 68, 4–20.

Schalkoff, R. J. (1990). *Artificial intelligence: An engineering approach*. New York: McGraw-Hill.

Segaran, T. (2007). *Programming Collective Intelligence: Building Smart Web 2.0 Applications*, 1st edition. Sebastopol: O'Reilly Media.

Selak, L., Butala, P., & Sluga, A., (2014). Condition monitoring and fault diagnostics for hydropower plants. *Computers & Industrial Engineering*. 65 (6), 924–936.

Skarka, W. (2007). Application of MOKA methodology in generative model creation using CATIA. *Engineering Applications of Artificial Intelligence*, 20 (5), 677–690.

Stark, R., Grosser, H., Beckmann-Dobrev, B., Kind, S., & INPIKO Collaboration. (2014). Advanced Technologies in Life Cycle Engineering. *Procedia CIRP*, 22, 3–14.

Steenhuizen, D., & van Tooren, M. (2012). The implementation of a knowledge-based framework for the aerodynamic optimization of a morphing wing device. *Advanced Engineering Informatics*, 26 (2), 207–218.

Tso G. K. F., & Yau, K. K. W. (2007), Predicting electricity energy consumption: A comparison of regression analysis, decision tree and neural networks. *Energy*, 32 (9), 1761–1768.

Tüfekci, P. (2014). Prediction of full load electrical power output of a base load operated combined cycle power plant using machine learning methods. *International Journal of Electrical Power & Energy Systems*, 60, 126-140.

Van Der Elst, S. W. G., & Van Tooren, M. J. L. (2008). Application of a Knowledge Engineering Process to Support Engineering Design Application Development. In Curran, R. et al. (Eds.) *Collaborative Product and Service Life Cycle Management for a Sustainable World*, *Proceedings of the 15th Conference on Advanced Concurrent Engineering*, Berlin: Springer, 417-431.

Verhagen, C. (2013). *An Ontology-Based Approach for Knowledge Lifecycle Management within Aircraft Lifecycle Phases*, TU Delft, Delft University of Technology.

Wang, Y. & Witten, I. H. (1996). *Induction of model trees for predicting continuous classes*. (Working paper 96/23). Hamilton, New Zealand: University of Waikato, Department of Computer Science.

Wu, Y.-H., & Shaw, H.-J. (2011). Document based knowledge base engineering method for ship basic design. *Ocean Engineering*, 38 (13), 1508–1521.

Yang, H. Z., J. F. Chen, N. Ma., & D. Y. Wang. (2012). Implementation of knowledge-based engineering methodology in ship structural design. *Computer-Aided Design*, 44 (3), 196–202.

Yang, J.B. (2002). A rule induction-based knowledge system for retaining wall selection. *Expert Systems with Applications*, 23 (3), 273–279.

Zhu, H., Gao, J., Li, D., & Tang, D. (2012). A Web-based Product Service System for aerospace maintenance, repair and overhaul services. *Computers & Industrial Engineering*. 63, 338–348.