Concepts for translation and transformation of information to knowledge

Written by:

Ahmed Bufardi, EPFL

Paul Folan, CIMRU

Kathryn Cormican, CIMRU

DELIVERABLE NO	DR7.1: Concepts for translation and transformation of information to knowledge
DATE	5th August, 2005
WORK PACKAGE NO	WP R7
VERSION NO.	Final Version 7.0
ELECTRONIC FILE CODE	DR71_v7.0.doc
CONTRACT NO	507100 PROMISE A Project of the 6th Framework Programme Information Society Technologies (IST)
ABSTRACT:	The document outlines concepts for the translation and transformation of information to knowledge. The 11 application scenarios are compared and contrasted to find distinct and common elements in their development of raw data to knowledge.

STATUS OF DELIVERABLE		
ACTION	**BY**	**DATE**
SUBMITTED (author(s))	Ahmed Bufardi, Paul Folan	13.07.2005
VU (WP Leader)	Lutz Rabe	11.08.2005
APPROVED (QIM)	Dimitris Kiritsis	21.08.2005

Revision History

Date (dd.mm.yyyy)	Version	Author	Comments
18.03.2005	0.1	Ahmed Bufardi	First draft of DR7.1 plan
21.03.2005	0.2	Paul Folan	Minor corrections of first draft
08.04.2005	0.3	Ahmed Bufardi	Detailed version of DR7.1 plan
15.04.2005	0.4	Paul Folan	Addition of content – first draft
22.04.2005	0.5	Paul Folan	Deletion of some content; suggestions for further material
27.04.2005	0.6	Ahmed Bufardi	Corrections and addition of content
29.05.2005	0.7	Ahmed Bufardi	Addition of content
16.06.2005	0.8	Paul Folan	Addition of application scenario material; major revision of structure
20.06.2005	0.9	Ahmed Bufardi	Modification of structure and addition of Section 4 describing the approach
30.06.2005	1.0	Ahmed Bufardi	Addition of content
06.07.2005	1.1	Ahmed Bufardi	Addition of content
08.07.2005	1.2	Paul Folan	Addition of content
11.07.2005	1.3	Paul Folan	Corrections and delivery of final draft
11.07.2005	1.4	Ahmed Bufardi	Corrections and final checking before review
16.07.2005	1.5	Ahmed Bufardi	Integration of input regarding the application scenario A10 from Markus Frey

Author(s)' contact information

Name	Organisation	E-mail	Tel	Fax
Ahmed Bufardi	EPFL	Ahmed.bufardi@epfl.ch	+ 412169338 68	+ 41216933509
Paul Folan	CIMRU	Paul.folan@nuigalway.ie	+ 35391750414	+ 35391562894

Table of Contents

1 Purpose of Deliverable R7.1

In many companies there is quite a substantial history of collection and analysis of product lifecycle data. However, the primary purpose of the use of this data has been to implement it in a managerial, marketing or logistical framework. The transformation of this data into appropriate knowledge can be useful for many other purposes such as predictive maintenance, improving design and production, determining the best end-of-life scenarios for retired products, etc. This type of use of product field data significantly contributes to one of the main objectives of PROMISE which is the closing of the product lifecycle information loops.

The purpose of Deliverable R7.1 is the analysis of the different product life cycle phases, for the translation and transformation of the related information into knowledge. However, as what is available as input for this transformation process is mainly the product field data, then the issue of transforming this data into information is a *prerequisite* for the translation and transformation of information into knowledge. That is why we will consider in this deliverable both the transformation of data into information and the transformation of information into knowledge.

There are three product life cycle phases under analysis:

- Beginning of Life (BOL) and particularly the production phase—whereby data about the different quality control points in the production line are to be considered;
- Middle of Life (MOL)—whereby data related to the status and functional condition, maintenance, reliability, availability, maintainability, life cycle cost (LCC), environmental and safety characteristics of a product are accounted for; and
- End of Life (EOL)—where data focusing upon the second life of products or product components and the behaviour during the upgrading, recycling and a second, third or more use of a product are examined.

From Browne et al. (1995) it can be deduced that there is a shift in focus from production to design, MOL and EOL, see Figure 1 (Kiritsis et al., 2003).

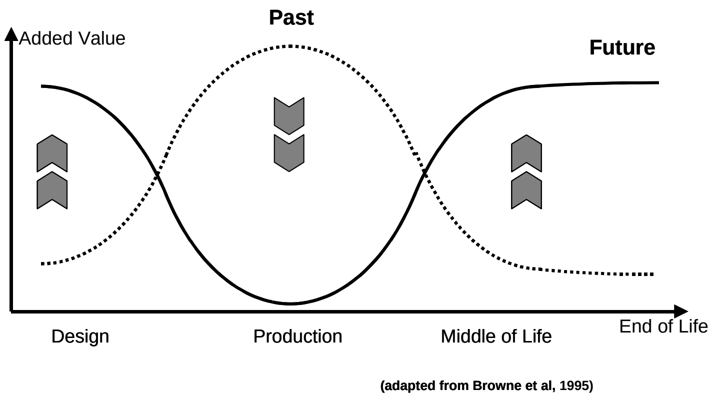


Figure 1. Shift in focus from production to design, MOL and EOL (Kiritsis et al., 2003)

Consequently, the generation of knowledge to support design, MOL and EOL are in line with the new trends expressed in Figure 1.

Also the knowledge to be generated from the product field data gathered at different lifecycle phases is intended to be used for different purposes and in different lifecycle phases:

- BOL:
  - Design—to improve some DFX aspects such as RAM, LCC, safety, environment, etc.;
  - Production—to adapt production processes;
- MOL—for example in predictive maintenance;
- EOL—for example in decommissioning and selection of relevant EOL options.

These three product life cycle phases are outlined in Figure 2; note the various types of material and information flows that feed information and material from EOL to BOL and MOL again. In the BOL stage, the design and production of the product and process are considered; quality is an important consideration—whether it is the quality of new or secondary materials, or the quality of the manufacturing process. In the MOL stage, when the product is with the customer, its success must be assessed upon its performance for reliability, availability, maintainability, environmental and safety factors—it is important for these information types to flow back to the designer and producer to enable them to correct and improve the product offering; at the MOL stage, the producer's responsiveness to the customer is key. In EOL a product may be re-introduced into the product life cycle completely, or, more likely, components from it be re-introduced; four main material flows may be outlined for this re-introduction: reuse (whereby components are removed and sold to the secondary market); remanufacturing (whereby components are brought up to the quality standards of new components and then reused); recycling (whereby *materials*—not components—from used products are reused again); and disposal (the removal of remaining waste products by incineration with or without energy recovery, landfilling etc.) (Thierry et al., 1995). Each of these re-introduction paths carry their own type of information flow back to the producer and designer of the product, which can act as a new input into BOL, and close the information loop.

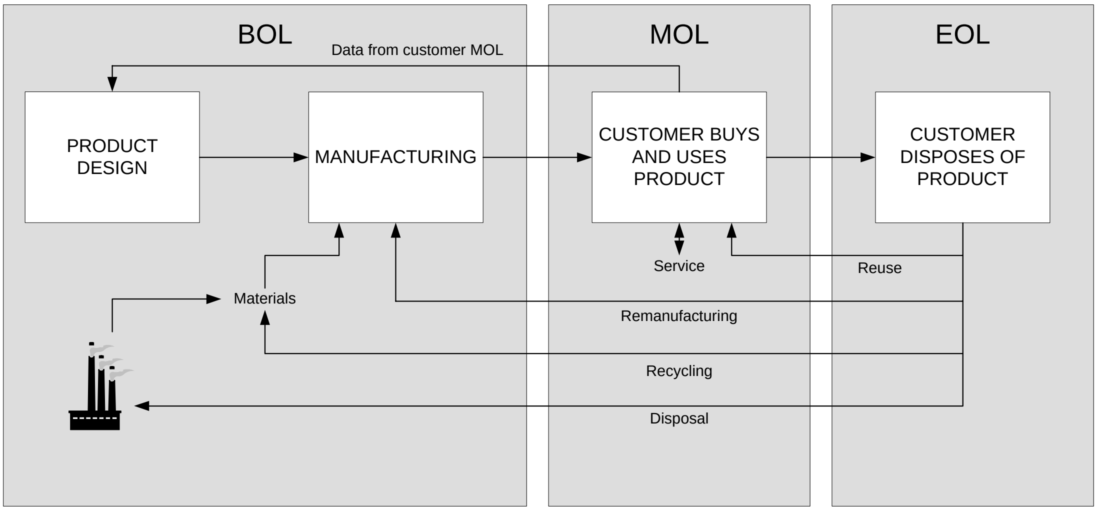


Figure 2: BOL, MOL, and EOL with reuse, remanufacturing, recycling and disposal

In this deliverable an efficient analysis of each of these product lifecycle phases will be carried out to ascertain the information flows from each, and how they may be utilised to capture the relevant knowledge for the maintenance personnel to support them in performing the maintenance operations, the product producer and designer to respectively improve the production and the design of similar products, etc.

The overall concept of PROMISE is to enable and exploit the seamless flow, tracing and updating of information about a product, after its delivery to the customer and up to its final destiny (deregistration, decommissioning) and back to the designer and producer; this will allow information flow management to go beyond the customer, to close the product lifecycle information loops, and to enable the seamless e-Transformation of Product Lifecycle Information to Knowledge. To perform these operations, Deliverable R7.1 will analyse the 11 application scenarios that are to be performed in PROMISE to determine their knowledge requirements, and their available information and data capabilities and propose concepts and models for the translation and transformation of data into knowledge on the basis of a thorough examination of the existing concepts for the translation and transformation of information into knowledge. When necessary, adaptation, extension or modification of existing concepts will be carried out.

2 Introduction to Deliverable R7.1

The main objective of this deliverable is to develop concepts for translation and transformation of information to knowledge in the context of the PROMISE project to be used in the three product life cycle steps: beginning of life (BOL) both in design and production, middle of life (MOL) and end-of-life (EOL). A prerequisite for the transformation and translation of information to knowledge is the transformation of the product field data into relevant information. Indeed, the starting point of all PROMISE application scenarios is to generate the required knowledge from the product field data. That is why we will consider in this deliverable all the steps that are necessary to go through in order to generate the required knowledge from the product field data.

The concepts to be developed should apply for the 11 application scenarios of the PROMISE project. However, the data to be considered from the different application scenarios, plus the knowledge derived from this data will be applied in a variety of different purposes.

The approach to be followed to define the concepts for transforming and translating product field data into information and information into knowledge will be based on a thorough analysis of the requirements of the 11 application scenarios regarding this issue. The determination of the common and specific requirements of the 11 application scenarios will guide us in defining the different concepts to consider. The 11 application scenarios are classified according to the main step where the required knowledge is to be applied as shown in Table 1. It is worth noticing that some application scenarios involve the application of the generated knowledge in more than one step.

Table 1: Classification of the application Scenarios A1 – A11 according to the main step where the required knowledge is to be applied

A1	A2	A3	A4	A5	A6	A7	A8	A9	A10	A11
EOL	EOL	EOL	MOL	MOL	MOL	MOL	MOL	MOL	BOL-design	BOL-production

The knowledge to be generated requires various product field data at different steps as explained below:

- At the BOL step and mainly at the production phase, the example of data to consider is the one related to the different quality control points in the production line;
- At the MOL step and mainly during the maintenance and service, the example of data to consider is the one related to the evolution of the status and operating conditions of the product and the reliability, availability, maintainability, environment and safety characteristics of the product;
- At the EOL step, the example of data to consider is the one related to the second life of products or their components and their behaviour during upgrading, recycling and a further use of a product.

Three related notions are addressed in this deliverable: data, information and knowledge. There are no common agreed definitions of these terms (Meadow and Yuan, 1997). One explanation for this is the fact that they are widely used and in so many different contexts that their meaning and use can vary from one context to another. Indeed, data, information and knowledge are polymorphic concepts that cannot be defined by a classical definition i.e. as a set of necessary and sufficient features that are universally valid (Aamodt and Nygård, 1995). The meaning of a polymorphic (non-classical) concept should be understood within a particular context i.e. in relation to some purpose or intended use, and seen from a certain perspective (Compton and Jansen, 1989).

In this deliverable, we will adopt the definitions of data, information and knowledge that are compliant with the PROMISE context. The common belief among researchers is that data is something less than information and information is in its turn less than knowledge and it is also assumed that data is needed before the information is created and only when information is available the knowledge can emerge (Tuomi, 1999). The representation of the links between data, information and knowledge in the PROMISE context is shown in Figure 3 (adapted from (Clark, 2005)). Note the context of R7.1 in the figure: the focus is upon a move from data to knowledge without necessarily being concerned with the origin of the data, or with the successive results of generating knowledge—wisdom.

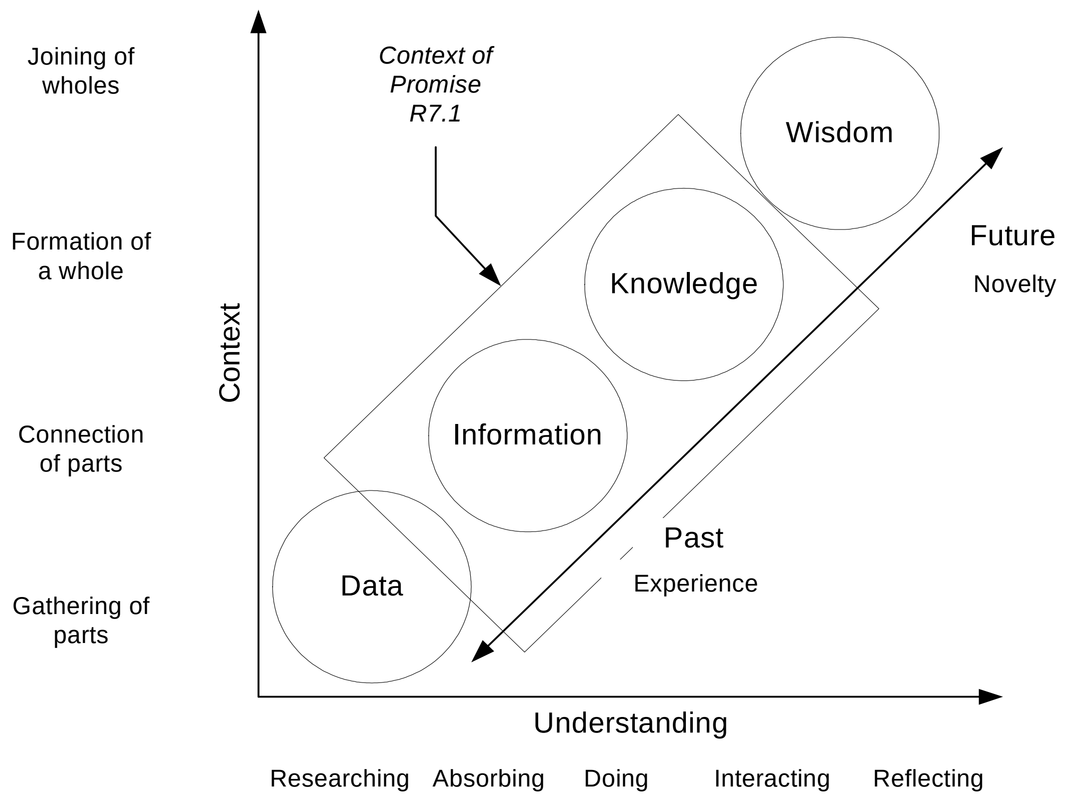


Figure 3: Links between data, information and knowledge in the PROMISE context (adapted from (Clark, 2005))

A general overview of the main steps through which data should go in order to generate the required knowledge is shown in Figure 4. It is worth noticing that not all the PROMISE application scenarios will consider all the steps shown in Figure 4 and also they may use different tools and methods at each step.

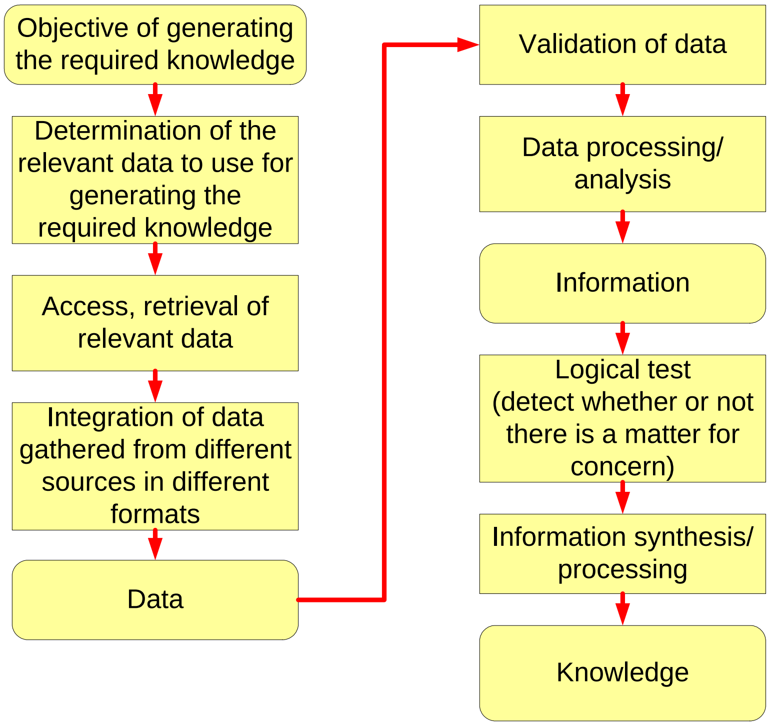


Figure 4: Main steps of transforming product field data into knowledge

The literature review captures the state of the art concerning the translation and transformation of data to information, and information into knowledge; with especial concentration upon a clear discussion upon the meaning and definition of these terms, as they are used generically, and—later—as they are to be applied in PROMISE. It is important that one set of definitions is adapted for the PROMISE context, as multiple definitions may cause considerable confusion. Subsequent sections introduce each of the application scenarios briefly, before outlining the knowledge requirements of each, and the available means to capture these; this is developed by an examination of the existing literature concerning each application scenario. Specifically, each application scenario will be analysed as to its processes of data access, retrieval and integration; data validation; processing / analysis of data; testing; and the synthesis / processing of information from this.

The individual analysis of each application scenario as to its knowledge requirements, and available capabilities to achieve these, ultimately allows for a comparison with the state of the art. Using the state of the art as a template for operations, specific concepts for the translation and transformation of information to knowledge, for each application scenario, may be outlined. Many of these, it will be found, are of a similar nature, enabling the development of generic knowledge development strategies based upon clusters of common application scenario requirements. Minor changes and amendments may have to be performed upon individual application scenarios to achieve this generic template, and these are considered here also. Ultimately, conclusions summarise the main achievements of the deliverable, and accounts for the extent to which the fixed objectives are fulfilled.

The deliverable is organized as follows:

- Section 3: Review of the related literature (generic)—the objective of this section is to summarize the state of the art concerning the translation and transformation of data to information and information to knowledge and also the definitions of data, information, knowledge and related terms. A brief overview of the classifications of knowledge, information and data will be provided in this section.

* Section 4: The approach to follow in order to generate the knowledge generator concepts is described.
* Section 5: In Sub-section 5.1, the process that describes how data is transformed to knowledge is introduced by detailing the main steps through which the product field data should go through in order to generate the required knowledge. This process is the key to the subsequent parts of the report, in that it is successively applied to each of the 11 application scenarios. In Sub-section 5.2, the knowledge application scenario checksheet used to gather the knowledge generation requirements of the 11 application scenarios is introduced. Sub-section 5.3 summarizes the main knowledge generation requirements of the 11 application scenarios. The results of the analysis of the requirements of application scenarios are provided in Sub-section 5.4.
* Section 6: Further analysis and a generalised summary of the 11 application scenarios are presented.
* Section 7: Conclusions.

3 Review of Related Literature

3.1 Introduction

In this deliverable we consider the three concepts of data, information and knowledge in the order of evolution from data into information and from information into knowledge. The inverse order also exists as data and information can be extracted from knowledge and data can be extracted from information but these issues are not considered in this deliverable.

The differences between data, information and knowledge are a matter of degree and a clear-cut definition of these terms is very difficult (Davenport and Pursak, 1998). However, knowledge can be better understood with reference to data and information because they are more familiar than knowledge. Data is needed to understand information and information in its turn is needed to understand knowledge. Consequently, the transformation of data into information and the transformation of information to knowledge are key issues in the generation of knowledge from product field data.

According to Davenport and Prusak (1998) knowledge can be considered as refined information to which human cognition has added value. Consequently, knowledge can be generated from information, other knowledge or both. In the framework of Deliverable R7.1, knowledge is mainly generated from information using eventually some auxiliary knowledge.

Data becomes information as soon as it is given a meaning (Davenport and Pursak, 1998). As most PROMISE application scenarios aim at generating knowledge from the product field data then, the transformation of data into information is a prerequisite for the transformation of information to knowledge.

According to Drucker (1993), innovation is the application of knowledge to produce new knowledge. It requires systematic efforts and a high degree of organisation. As we enter the knowledge society, ownership of knowledge and information as a source of competitive advantage is becoming increasingly important. In other words, organisations depend more on the development, use and distribution of knowledge based competencies. This is particularly relevant in knowledge intensive processes such as product innovation. Consequently, research and development (R&D) organisations are paying more attention to the concept of managing their knowledge base in order to increase competitive advantage, through effective decision making and increased innovation (Nonaka, 1991; Davenport et al, 1996; Sveiby, 1997). Knowledge is a key resource that must be managed if improvement efforts are to succeed and businesses are to remain competitive in a networked environment (Gunasekaran, 1999). In particular, the two major challenges that face organisations are; (a) ensuring that they have the knowledge to support their operations and (b) ensuring that they optimise the knowledge resources available to them.

This section is devoted to the review of literature related to data, information, knowledge, transformation of data into information and transformation of information to knowledge. Sub-section 3.2 deals with the links between data, information and knowledge; sub-section 3.3 discusses the dynamics of enterprise knowledge, which includes a comprehensive definition of knowledge and knowledge work; and sub-section 3.4 deals with knowledge typologies and the knowledge process for transforming data to knowledge. Finally, sub-section 3.5 summarises the definitions of data, information and knowledge and the knowledge process as used in the deliverable.

3.2 Understanding Knowledge, Information and Data

Understanding the key concepts of data, information and knowledge is important for setting the scope of this study. Many authors have noted that there is a difference between these concepts (Knock et al 1997; Wilson 1996; Bohn 1994). However, this difference is difficult do define clearly. Data is characterised as a set of discrete facts about events and the world. According to Davenport and Prusak (1998) data describes only a part of what happened; it provides no judgment or interpretation and no sustainable basis of action. Therefore, there is no inherent meaning in data.

Glazer (1991) contends that information is "*data that have been organised or given structure – that is placed in context – and thus endowed with meaning*". In other words, information is the outcome of capturing and providing context to experiences and ideas. For example, reliability, failure rate, percentile and mean time to failure, etc. are information about the life characteristics of a product (Oh and Bai, 2001).

Knowledge on the other hand is composed of tacit experiences, ideas, insights, values and judgements of individuals (Bohn, 1994). It is dynamic and can only be accessed through direct collaboration and communication with experts who have the knowledge. According to Wilson (1996) by selecting and analysing data, we can produce information and by selecting and combining information we can generate knowledge. The processing hierarchy of data, information and knowledge is illustrated below (see figure 5).

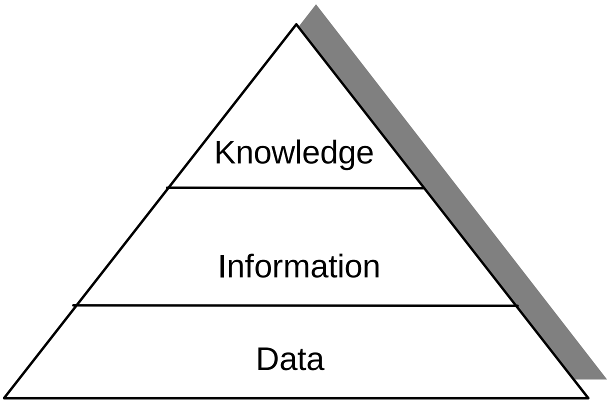


Figure 5: Hierarchy of Knowledge Assets

It is important to note that information technologies can help translate data into information. Information, on the other hand, is converted into knowledge through the social human process of shared understanding and sense making at both the personal level and the organisational level. According to Churchman (1971) "*to conceive of knowledge as a collection of information systems seems to rob the concept of all its life……Knowledge resides in the user and not in the collection. It is how the user reacts to the collection of information that matters*". Therefore, managing knowledge is about creating an environment that fosters the continuous creation, aggregation, use and reuse of both organisational and personal knowledge in the pursuit of new business value.

The objective of this section is to summarize the state of the art concerning the translation and transformation of data to information and information to knowledge and also the definitions of data, information, knowledge and related terms.

A brief overview of the classifications of knowledge, information and data will be provided in this section.

3.3 Knowledge and related issues

3.3.1 Defining Knowledge & Knowledge Work

Knowledge is an elusive concept and therefore it is important to define it in context in order to understand it. The term is used in several different ways in the literature. For example, Nonaka and Takuechi (1995) two of the early researchers in this field adopt a philosophical angle and define knowledge as "*justified true belief*". In this view, knowledge is an opinion, idea or theory that has been verified empirically and agreed upon by a community. According to Wilson (1996), knowledge at the most basic level is "*that which is known*". Quinn et al (1996) liken knowledge with professional intellect where professional intellect in organisations centres on know-what, know-why, know-how and self motivated creativity. Stewart (1997) also considers knowledge in terms of intellectual capital. On the other hand, Bohn (1994) examines knowledge in terms of a company's processes. He believes that an organisation's knowledge about its processes may range from total ignorance about how they work to very complex and formal mathematical models. According to Davenport et al (1998), knowledge is information combined with experience, context, interpretation and reflection. It is a high value form of information that is ready to apply to decisions and actions. Simply put, knowledge can be defined as the integration of ideas, experience, intuition, assertions, skills and lessons learned that have the potential to create value for a business by informing decisions and improving performance. In this view, knowledge is a key enabler to organisational success. However, in order for knowledge to be useful it must be available, accurate, effective and accessible.

The specialisation of work leads to an increasing need for knowledge workers. For example, as product innovation becomes increasingly complex there is a greater need for specialised workers. In this environment what flows most between knowledge workers is information and data as opposed to physical material. However, unlike knowledge relatively few researchers have attempted to define knowledge work. The nature of knowledge work is *ad hoc*, demand driven and creative (Harris, 1999). Davenport et al (1996) contend that knowledge work focuses on the acquisition, creation, packaging or application of knowledge. In this view, it is complex and diverse and it is performed by professional or skilled workers with a high level of expertise and competence. According to Harris (1999), a knowledge worker is formally defined as one who gathers, analyses, adds value and communicates information to empower decision making. A knowledge worker's job entails doing work for which there is no finitely determined process. Their tasks are not prescribed in advance, but are determined just in time in response to issues, opportunities or problems as they arise. Each event may require a customised unique content and collaboration with a different group of people.

According to Laudon and Laudon (1999) not only do knowledge workers use their knowledge to interpret incoming information, but they also create new knowledge as well. Knowledge work processes include such activities as research and development, product development and professional services such as software development, law, accounting and consulting (Davenport et al, 1996). Knowledge workers hold expertise composed of competence and skills and they are typically more productive and better paid than non-experts. Knowledge workers value is acquired through formal education. Such people understand how to learn and will continue to learn throughout their productive lives. What is learned and how it is applied will determine competitive success. According to Takeuchi (1998) knowledge workers now constitute up to 35-40% of the workforce and these will become the leading social group. Therefore, organisations' core competencies will focus on managing knowledge and knowledge workers. Furthermore, industrial growth and productivity gains will depend heavily on improvements in knowledge work.

Drucker (1993) believes that the great management task of this century will be to make knowledge work productive. Davenport et al (1998) also state that organisations' core competencies will centre on managing knowledge and knowledge workers in the future. They add that industrial growth and productivity gains will depend heavily on improvements in knowledge work. Thus, a viable approach is critically needed for improving knowledge work. However, managing knowledge is intricate, complex and often very difficult and consequently companies are finding it difficult to implement knowledge-based practices. Wiig (1995) provides a list of knowledge related problems found in organisations. These include:

- Knowledge is not managed as a valuable asset.
- There is insufficient knowledge at the point of action.
- Learning opportunities are often missed or not exploited.
- Knowledge transfer is confined.
- There is often an unnecessary division of tasks and decisions.

There is little evidence (anecdotal, empirical or otherwise) to suggest that adequate provision is made for promoting, capturing, sharing and disseminating knowledge in organisations. Also, as knowledge management initiatives and systems are just beginning to appear in organisations, there is little research and field data to guide the development and implementation of such systems or to guide the expectations of the potential benefits of such systems. Upon analysis it seems that these deficits must be addressed.

3.3.2. Knowledge Typologies

Many types or classifications of knowledge have been suggested in the literature. However, this section explores knowledge typologies, which relate to product innovation. For example, distinctions are made between experiential knowledge and reported knowledge as well as intimate knowledge and declared knowledge (Wikstrom et al, 1994). Carlsen and Skaret (1998) speak of individual and collective knowledge i.e. whether knowledge resides in individuals, groups of individuals or the company as a whole. Ruggles (1997) proposes a broad typology of knowledge, which describes knowledge in terms of what it is about. This typology incorporates; (a) process knowledge, such as methods for doing things well; (b) factual knowledge, which is basic information about people and things; (c) catalogue knowledge, which refers to knowing where things are; and (d) cultural knowledge, which comprises understanding the values, rules and norms in an organisation. Stewart (1997) speaks of four levels of professional intellect. They are cognitive knowledge (know what); advanced skills (know how); systems understanding (know why); and self motivated creativity (care why). He believes the value of intellect increases as one moves up the intellectual scale from cognitive knowledge to self-motivated creativity.

Considerable attention has been paid to the distinction between tacit (implicit) knowledge and explicit (codified) knowledge. The term tacit was originally coined by Michael Polanyi (Polanyi, 1966). According to Wilson (1996) tacit knowledge is personal knowledge, which consists of highly subjective insights, intuitions and instincts. Tacit or implicit knowledge has a personal quality that makes it hard to formalise and communicate. It is deeply rooted in action and involved in a specific context. Explicit or codified knowledge refers to knowledge that can be communicated in formal systematic language. It is worth noticing that in this deliverable, we are concerned with explicit knowledge that is obtained through processing and synthesizing of information. Joseph Badaracco (1991) uses another term to describe explicit knowledge when it is captured as formulae, designs, manuals or books or in pieces of machinery. He calls it migratory knowledge because it can move out of the organisation very quickly. Conversely, embedded knowledge is the organisational knowledge, which cannot be owned and used in isolation by an individual. It is likened to the culture of the organisation in that it exists in norms, attitudes and relationships among individuals or groups. These typologies are summarised in table 2. While this classification is by no means exhaustive it does however provide some indication of the research undertaken in this area.

Table 2: Typologies of Knowledge

CLASSIFICATION	REPORTED BY
Tacit knowledge Explicit knowledge	Polanyi (1966)
Migratory knowledge Embedded knowledge	Badaracco (1991)
Experiential knowledge Reported knowledge Intimate knowledge Declared knowledge	Wikstrom et al (1994)
Cognitive knowledge (know what) Advanced skills (know how) Systems understanding (know why) Self motivated creativity (care why)	Stewart (1997)
Process knowledge Factual knowledge Catalogue knowledge Cultural knowledge	Ruggles (1997)
Individual knowledge Collective knowledge	Carlsen and Skaret (1998)

Knowledge is of little use unless it can be applied. In other words, it must be translated into creating some observable product or service. Therefore, the focus of attention must shift from the individual to organisational knowledge management. Manufacturing enterprises must be able to develop an environment, which can facilitate the creation or generation of tacit knowledge for product innovation while simultaneously be capable of converting individual skills and competencies into corporate knowledge and know how. In order to do this, knowledge management initiatives (such as those aimed at enhancing product innovation initiatives) must be put in place. The following section focuses on the concept of knowledge management in an industrial setting.

3.3.3. Knowledge Management Definition

The central problem of knowledge management is its lack of an absolute definition. For example, many information technology journals define knowledge management in terms of understanding the relationships of data, identifying and documenting rules for managing data and assuring that data are accurate and maintain integrity (Malhorta, 1998). However, knowledge should not be viewed simply as data or information that can be stored in the computer as it also involves emotions, values or hunches (Takeuchi, 1998). Malhorta (1998) also believes that interpreting knowledge management in terms of rules and procedures embedded in technology does not reflect the dynamically changing business environment.

Many researchers and industrialists postulate that knowledge management centres on the creation or generation of knowledge (Nonaka, 1991, Stewart, 1997). Others believe that knowledge management should focus less on knowledge creation and more its capture and integration (Martin, 1995; Grant, 1996; Alavi and Leidner, 1997). However, most agree that knowledge management encompasses all of these activities, that is, the creation or generation, codification, storage, dissemination and implementation of knowledge in the organisation. For example, Bassi (1998) defines knowledge management as the process of creating, capturing and using knowledge to enhance organisational performance. Blake (1998) believes it as "*… the process of capturing a company's collective expertise wherever it resides and distributing it to wherever it can help produce the biggest payoffs*". Parlby (1997) also believes that knowledge management is the discipline of capturing knowledge based competencies, storing and disseminating them for the benefit of the organisation as a whole. Ruggles (1998) considers knowledge management as, "*an approach to adding or creating value by more actively leveraging the know how, experience and judgement resident within, and in many cases, outside the organisation*." Taking these definitions into consideration knowledge management can be considered to be a systematic and organised attempt to use knowledge within a company to transform its ability to generate, store and use knowledge in order to improve performance. In short, the overriding purpose of enterprise knowledge management is to make knowledge accessible and reusable to the organisation.

3.3.4. Knowledge Management Goals

According to Neef (1997), enterprise growth depends upon innovation and innovation depends on knowledge. Therefore, knowledge management not only acts as a catalyst for activities such as product innovation, but also provides the means, by which innovative ideas can be captured, shared and leveraged leading to new ideas. This increased recognition of knowledge as a core competence combined with recent advances in information technology such as intranets and the world wide web, has enhanced organisations' interest in the topic of knowledge management. According to Drew (1999) some features of knowledge management include:

- Holism and humanism: the priority is to make better use of human potential rather that downsize it.
- A concern with growth and new possibilities by developing new knowledge.
- Support of creative management practices, which result in new competencies.
- Making good use of important technological developments such as networks.

The principle goals of knowledge management are twofold. Firstly, such initiatives help to foster innovation and secondly, they facilitate better decision making. It is worth noticing that the main goal of the knowledge to be generated in this deliverable is to use it to support decision making. According to Carayannis (1999) the aims and objectives of knowledge management initiatives can be summarised as follows:

- To develop and foster new and promising areas of collaborative, inter-disciplinary, and cross-functional knowledge work.

- To catalyse the creation of cross-disciplinary and cross-functional knowledge clusters across teams and organisations.
- To enable better utilisation of resources by reducing/eliminating redundancies and identifying weaknesses and anticipating opportunities for change.
- To provide a more responsive information technology infrastructure supporting knowledge workers, being able to design products and services which are in line with current and emerging market needs.

Unfortunately, there is no simple means of generating and transferring knowledge effectively. However, by understanding the knowledge process and the steps involved managers will be in a better position to actively manage the intellectual capital of the firm. Therefore, the following section examines the knowledge process.

3.3.5. The Knowledge Process

Wikstrom et al (1994) contend that knowledge processes of various kinds are a constant and stable part of company life. Many researchers have examined knowledge in terms of its process (see Ruggles 1998; Wikstrom et al, 1994; Davenport et al, 1996; Bohn, 1994; Kotnour et al, 1997; Nonaka and Takeuchi, 1995). In other words, they look at the sequence of steps or stages that are involved in the knowledge process. Bohn (1994) defines a process as "*any repetitive system for producing a product or service, including the people, machines, procedures, and software, in that system*". In this view, the process is defined in terms of the outputs generated. Other researchers focus on the sequence of events or the decisions by which an innovation is introduced. For example, Davenport (1993) defines a process as "*a specific ordering of work activities across time and place, with a beginning, an end and clearly identified inputs and outputs: a structure for action*". Here a process has inputs, outputs as well as variables that characterise what is happening inside it. This implies that the knowledge process involves the effective management of many different activities.

3.4 Transformation of data into information and information into knowledge

The process of moving from a situation of data acquisition into a procedure for deducing knowledge via a process of developing the data into information is outlined by Hicks et al. (2002) in Figure 6. As can be seen from the figure, data is given a context or a surrounding descriptor in order to enable the development of information; the information itself may be used in a decision-making process (and new information may be obtained thereby), or it may be used to ultimately produce knowledge. By combining specific information to gain an increased understanding or perspective of the objective that we wish to proceed towards, results in knowledge generation. Note that a backward step is possible whereby knowledge is transformed into information via a semantic interpretation of the knowledge element itself—a simple example of this would be where an author writes a book from their own *knowledge*, but this acts as *information* for a new reader of the book. The knowledge produced ultimately results in a decision—based upon the knowledge—being reached and pursued; this decision-making process may also result in new knowledge in itself, enriching the previous knowledge with new insights.

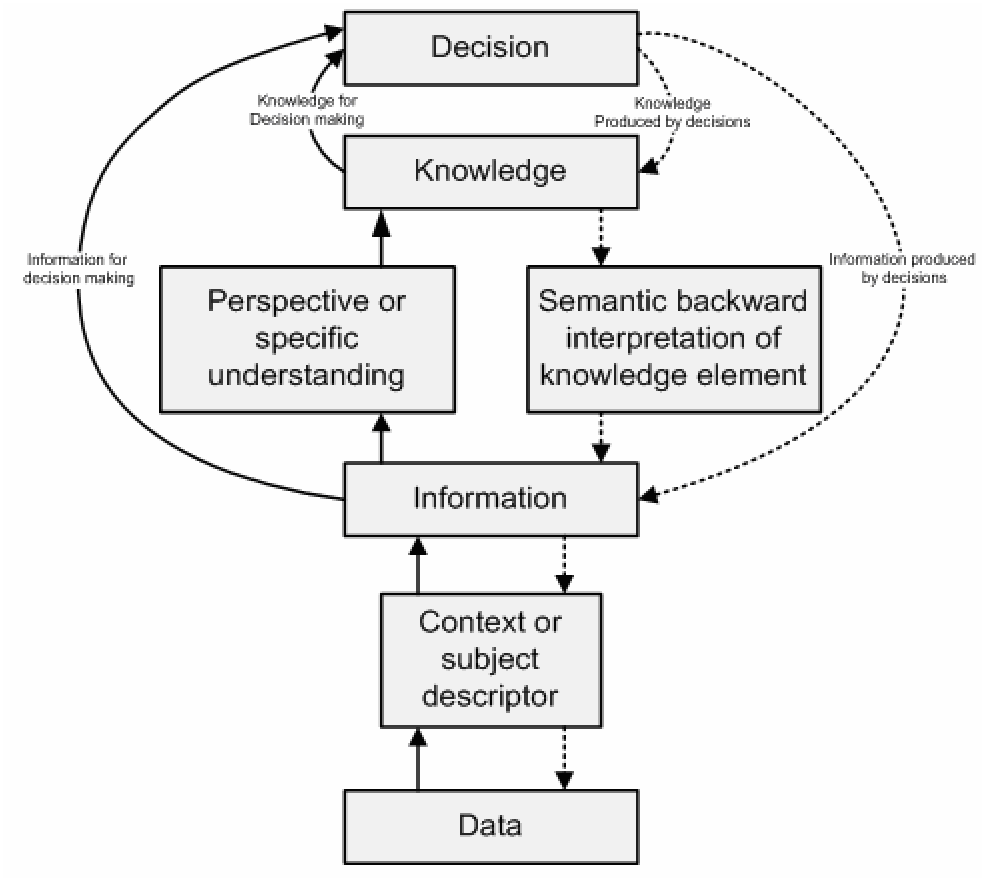


Figure 6: Bi-directional info. & knowledge transformation processes for decision making (Hicks et al., 2002)

A first step in the process of generating knowledge from field data is to transform this data into information. According to Davenport and Pursak (1998) there are various methods that allow transforming data into information:

- **Contextualization**: Determining for which purpose the data is gathered;
- **Categorization**: Determining the units of analysis or key components of the data;
- **Calculation**: Mathematical or statistical analysis of the data;
- **Correction**: Removal of errors from the data;
- **Condensation**: Summarization of the data in a more concise form.

According to Hicks et al. (2002) knowledge generation involves two main aspects: (i) the knowledge element and (ii) the knowledge process. The knowledge process is the procedures utilized by the individuals to infer the knowledge element information, or other knowledge elements or a combination of both; the knowledge elements are inferred from one or more elements of the information.

The second step is to move from information to knowledge. According to Davenport and Pursak (1998) there are various methods that allows transforming information into knowledge:

- **Comparison**: Comparison of the information about a given situation to other known situations;
- **Consequences**: Determination of the implications that the information have for decisions and actions;
- **Connections**: Determination of the relation of a bit of knowledge to others;

- **Conversation**: Determination of the opinion of other people about the information.

The knowledge produced must ultimately be compared against the objectives held at the beginning to see if the results produced match those that are required from the knowledge. The knowledge will probably be used as the basis upon which decisions will be made, prompting the development of new knowledge and information from the results of the decisions reached. In this way, the knowledge process continually accumulates new knowledge from existing knowledge.

3.5 Summary and definitions

In Deliverable R7.1 the following definitions are used:

- Data—all that is collected by individuals (such as maintenance personnel) or devices (such as sensors, PEIDs, RFIDs, etc.) concerning the behaviour/status/function of the systems under consideration in the 11 application scenarios; these data can be provided in different formats such as symbols, numbers, graphs, figures, text, etc. The data is without a specific context.
- Information—data worked out so as to situate it in a context and give them a meaningful format or structure.
- Knowledge—information selected and combined, plus the influence of tacit experiences, ideas, insights, values and judgements of individuals.

The knowledge process that is adopted in the rest of this report is depicted in figure 7. This figure will be explained in section 5.

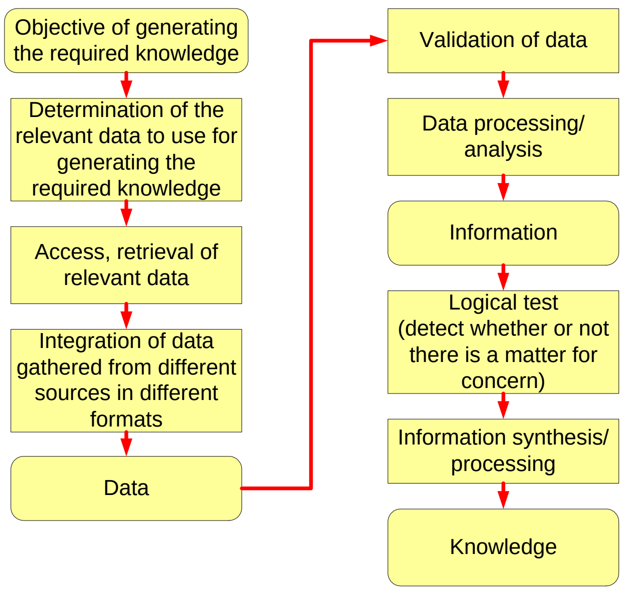


Figure 7: Main steps of transforming product field data into knowledge

4 Approach to developing concepts for transformation of field data to knowledge

The approach followed in this deliverable to define the concepts for the transformation of product field data into the required knowledge is based on the analysis of the requirements regarding the transformation of product field data into knowledge of the 11 application scenarios considered in PROMISE.

The approach is composed of three main steps (Figure 8):

Step 1: Determination of the requirements regarding the transformation of product field data into the required knowledge for each application scenario. This goal is achieved through the use of a knowledge application scenario template (Table 6 in Sub-section 5.2) consisting of 8 areas (fixing the main objectives of obtaining the required knowledge, determining the data that is relevant for generating the required knowledge, accessing and retrieving the relevant data, integrating the data that is gathered from different sources, validating the data, analysing/ processing the data, detecting whether there is a matter of concern or not, synthesising/processing information) which the participant had to fill in for each application scenario in order to provide an overview of the knowledge issues involved.

Step 2: Classification of the knowledge generation requirements of the 11 application scenarios into specific and common requirements. It is worth noticing that the notions of specific and common are relative since a common requirement can be related to two, three or more application scenarios.

Step 3: Definition of the concepts for the generation of knowledge from product field data by exploiting the set of common and specific requirements of the 11 application scenarios. The following conditions can aid in defining the number of concepts to consider:

i) **completeness** i.e. all the knowledge generation requirements of the 11 applications scenarios should be covered by the concepts;

ii) **non redundancy** i.e. the concepts should be distinct enough in the sense that these should exist some features that allow distinguishing between any two different concepts; and

iii) **minimality** i.e. the number of concepts should be kept to a minimum while preserving the two other conditions.

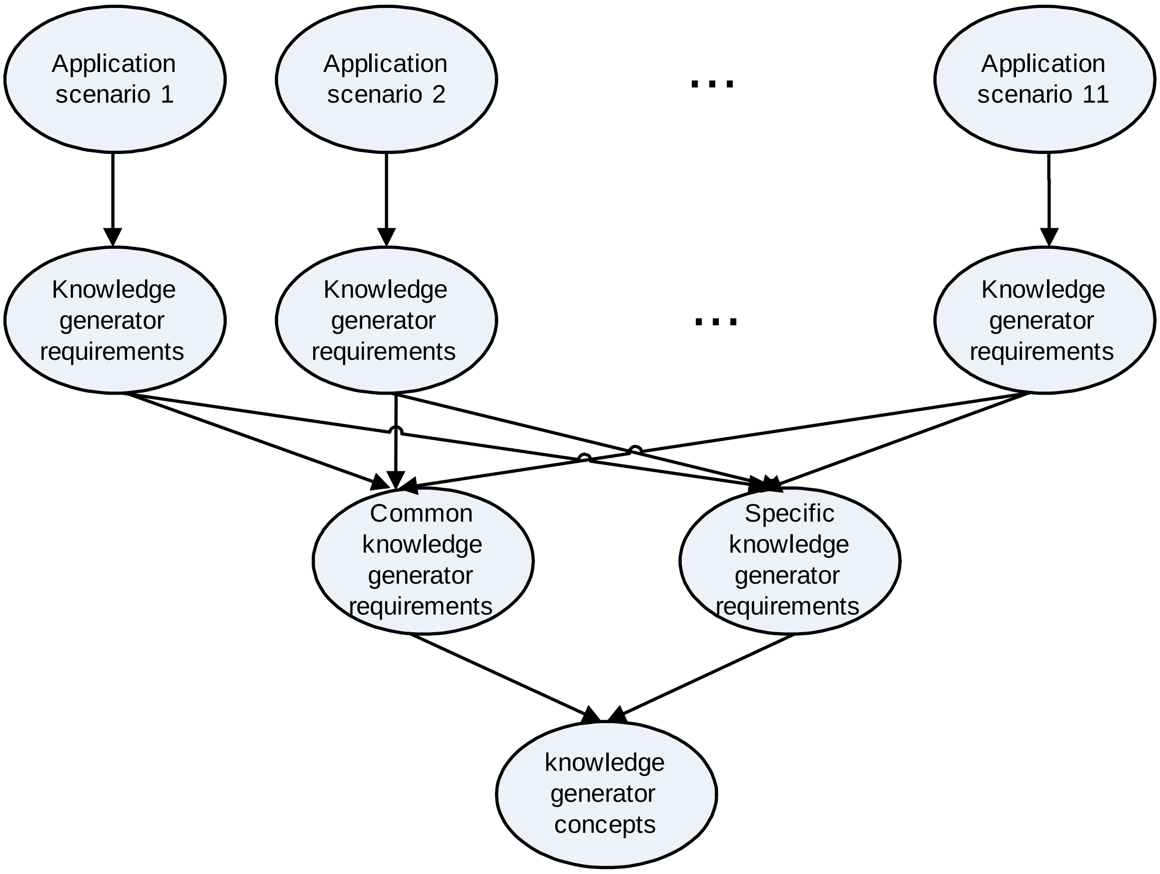


Figure 8: Approach to defining the knowledge generation concepts

5 Concepts for transforming field data to knowledge

5.1 Description of the main steps of the transformation of field data into knowledge

A general overview of the main steps through which data/information may go in order to generate the required knowledge is shown in Figure 9. It is worth noticing that not all the PROMISE application scenarios will consider all the steps shown in Figure 9 and also they may use different tools and methods at each step. The steps involved in the process are briefly described in the following sub-sections.

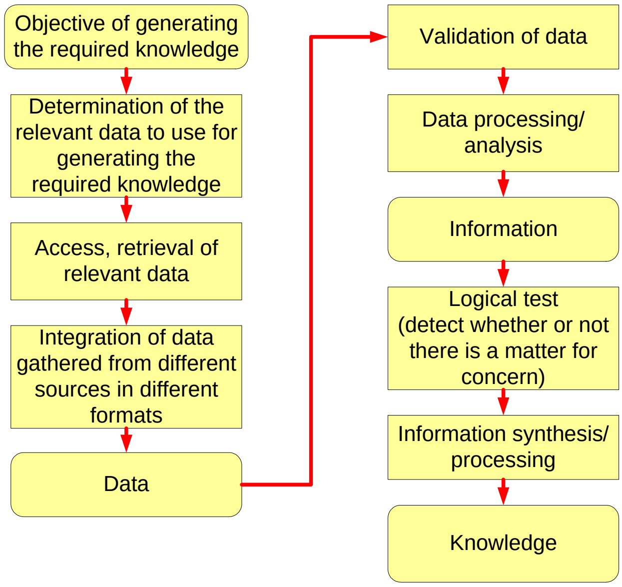


Figure 9: Main steps of transforming product field data into knowledge

5.1.1. Objective of generating the required knowledge

The objective of generating the required knowledge is the ultimate goal of the process of moving from a situation where we have data, to a situation of acquiring knowledge. In order to successfully produce knowledge from the existing data that we have obtained, we must know *what* we intend to do with the basic data: do we wish to solve a problem? make a decision? etc. When we have decided what our ultimate objective is in the knowledge process, then we can compare the results of the process at the end, to the original objective. In the PROMISE project, the ultimate objectives may be outlined briefly in the 11 application scenarios as follows (see Table 3).
For some application scenarios such as A10, we should further decompose the general objective provided in Table 3 into sub-objectives. For example in Table 3, the general objective of A10 is to generate DfX knowledge to use for the improvement of some design aspects however there are various aspects of DfX knowledge that can be considered such as deign for safety knowledge, design for lifecycle cost knowledge, etc. and the data related to these aspects should be transformed separately.

Table 3: Application scenarios—existing objectives[1]

App. Scenario	Relates to…	Main EOL step	Main Objective(s)
CRF (EOL)	A1	EOL	- Identification of components that are worth to reuse when deregistering the vehicle - Transfer on the component of "some" relevant information about its post – deregistering life
Caterpillar (EOL)	A2	EOL	- To prove the closure of information loop between the knowledge required to manage the field population as well as to make EOL decisions - Aggregation of available engineering

[1] The objectives are extracted from Appendix A "The Demonstrators" of Deliverable DR3.2

			data, filed data and ancillary information - Use of PEIDs to track components and automatically maintain related data linkages so it can be transformed into the required knowledge
INDYON	A3	EOL	- Increasing the recycling rate - Tracking and tracing the relevant product data together with the materials during all recycling processes and supply of these data for product design - Support of decision making systems with relevant data.
CRF (MOL)	A4	MOL	- Development of predictive maintenance algorithms to predict the engine oil wear out of a specific vehicle on the basis of ad-hoc mission profile indicators - Collection of prediction values on a ground station to remotely manage a fleet of vehicles - Collection of ad-hoc mission profile indicators on a ground station to remotely manage a fleet of vehicles with the objective of having information about the mission profile of each vehicle belonging to the fleet
Caterpillar (MOL)	A5	MOL	- To prove the closure of information loop between the information related to field application embedded in field data and knowledge required to manage the field population - Aggregation of available engineering data, field data and ancillary information - Use of PEIDs to track components and automatically maintain related data linkages so it can be transformed into the required knowledge
FIDIA	A6	MOL	- Traceability of components - Use of machine field data to support predictive maintenance - Generation of knowledge to support design improvement of components
MTS	A7	MOL	- Use of gas boilers field data to support predictive maintenance through the application of evolutionary diagnostic and prognostic algorithms - Testing of the capabilities of PEID regarding the collection of data from boiler serial protocol and from additional sensors, the storage and the transmission of data with long distance communication - Testing of back-end, PDKM, PLM and DSS as to their ability to handle the information on MOL, to analyse it and enable the users to access it for predictive maintenance
Wrap	A8	MOL	- To make faster End-of-line testing - To improve the compressor efficiency

			- To improve the cooling circuit pressure - To improve the control of Internal/external Temperature
INTRACOM	A9	MOL	- Efficient collection, integration and management of information about the product - Reception of data about product operation - Transformation of operational data into valuable knowledge to support decision support system - Decision making support to engineers and technicians about product improvements and problems solving/preventive maintenance
Bombardier	A10	BOL-Design	- Generation of DfX knowledge to support engineers in improving various aspects of the design such as reliability, availability, maintainability, life cycle cost, environment and safety. - Decision support to the transformation of product field data into DfX knowledge - Inclusion of the knowledge management functionality in the PDKM system to manage the DfX knowledge structured according to a predefined work breakdown structure
Polimi	A11	BOL-Production	- Production system reconfiguration to improve overall enterprise performance; prompted by feedback from MOL, EOL (kit of tools)

Several comments follow from Table3:

- In Table 3, only the main step of focus of the application scenario is mentioned. However some application scenarios are concerned with more than one step. For example, the main focus of A6 and A9 is MOL but they are also interested in the generation of knowledge to support design improvement.
- In Table 3 all objectives of the application scenarios described in Appendix A "The Demonstrators" of Deliverable DR3.2 are mentioned. However not all these objectives are related to the topic addressed in Deliverable DR7.1 that consists of transforming product field data into knowledge. For example, in A7, the objective "Testing of the capabilities of PEID regarding the collection of data from boiler serial protocol and from additional sensors, the storage and the transmission of data with long distance communication" is not related to the topic addressed in this deliverable.
- In Table 3, some objectives are well detailed to the point of specifying the method such as "Use of gas boilers field data to support predictive maintenance through the application of evolutionary diagnostic and prognostic algorithms" in A7 and some others are not sufficiently detailed such as "Transformation of operational data into valuable knowledge to support decision support system" in A9.
- The type of data transformation required is quite different from one application scenario to another. For example in A10, the field data to consider is historical data whereas in some other application scenarios such as A4, the field data to consider is real time (checked at "sufficiently small" interval times) data.

5.1.2. Determination of the relevant data to use

There is a requirement to determine the correct data to obtain in order that we may ultimately achieve the objective outlined in the sub-section above. In the normal course of events, the analyst has an over-abundance of data available to choose from in the context of any given objective; data may be coming from opposing systems, entities, externally, internally etc. The key to maintaining a coherent process that allows the movement from data to knowledge is to choose the correct data that is of relevance to the objective in question and to subject this to the knowledge process. This requires a selection procedure: selection is necessary in order to choose the right data at the right time, and to discard irrelevant data. In the application scenarios A1-A11, the determination of the correct data to achieve the objectives outlined has resulted in a wide disparity of data types and data requirements to be collected.

Almost all companies collect some lifecycle data through sensors or by some individuals such as maintenance personnel. However, the goal of the collection of these lifecycle data is not necessarily to enable the generation of the required knowledge expressed in some of the objectives described in Table 3. Two problems arise: (i) which data among the collected is suitable for the generation of the required knowledge and (ii) is all the necessary data needed to generate the required knowledge within the available data.

The appropriate way (Figure 10) to obtain the relevant data is to identify first the needed data and then to use methods and tools that allow the acquisition of this data. The issue of how to collect relevant data will not be addressed in the DfX demonstrator even if we believe that this issue has a great impact on the accuracy of the knowledge to be generated.

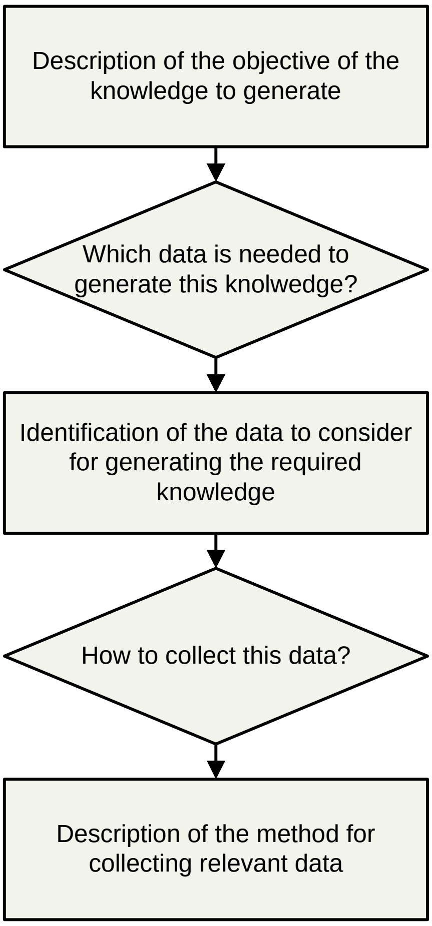


Figure 10: Approach to defining a procedure for collecting relevant data to generate specific knowledge

5.1.3. Access & Retrieval of the relevant data

Once the relevant data has been determined, it becomes important to know its whereabouts, its access capabilities, and whether it can be retrieved or not. Data that cannot be accessed cannot be processed into knowledge, and so it cannot be determined whether the data would successfully uphold the given objective required. This was, to a large extent, the situation faced by analysts until recently when faced with data access and retrieval from MOL and EOL scenarios. Now thanks to improved data storage and access techniques (e.g. RFID tags, PDA readers etc.), this—it is hoped—will become a much-reduced problem. The retrieval of the relevant data is an important systems question normally examined under the heading of the various types of middleware available to interrogate the relevant database, decide the correct data to access and read, and to return this data so that the analyst may use it. In the application scenarios A1-A11, the access and retrieval procedures to be adopted are expected to be widely different in BOL, MOL, and EOL.

5.1.4. Integration of data

When relevant data has been collected from a number of different sources, for example from a number of different databases or PEIDs—it may not be uniform; that is, it may require analytical attention to ensure that it is consistent and applicable. Common uniformity techniques include the use of algorithms and other techniques (AHP etc.) to ensure that the data to be used to inform the development of information and subsequently the production of knowledge is consistent, may be compared, and is free from subjectivity. Common forms of data that are *not* uniform include: aggregations of data that uses different critical subgroups; irrelevant data included with relevant data; varying interpretations of data resulting in differences; subjectivity on the part of the collector which may result in the loss of some relevant data, the retention of irrelevant data, or the mismeasurement of the data itself. Integration of data deals with the processes put in place to avoid these problems.

5.1.5. Validation of data

In a step connected with the integration of data, validation examines the results of the integration process in the sub-section above—i.e. the data that has been declared uniform. Validating the data at this stage is important in order to assess its further usage; the data that is used to develop information must be specific to the information-type in order to be placed in its proper context and given a structure. Validation ensures that the information to be produced is accurate and is useful for the purpose of achieving the objective set at the start; validation requires the data to be not only free from internal error (the function of integration above) but to be of relevance to the future steps in the knowledge process.

To ensure the validity of the data regarding the type of analysis it must undergo, some precautions should be observed. For example concerning the size of data sample, one has to look at: (i) the number of different components, parts or systems of the same family that are being monitored, (ii) the time span over which a given component, part or system was observed and related data concerning it collected, (iii) external conditions such climatic conditions, operational conditions, human implication, etc. The conditions of validity such as threshold assignments depends on the analysis methods used.

5.1.6. Data processing & analysis

This step provides the data with the relevant context so that it may be interpreted as information. This is performed by supplying an analysis framework to the data so that it may be processed with similar pieces of data to form information with structure and context. Data is often processed by comparison, or by the addition of further elements to its field in order to provide a context for its development. Typical examples of comparative data processing includes the development of graphs, which provide pictorial information from the data; while the addition of other forms of data combined may produce information also—an example being the analysis of the failures of a component (data type 1) over a period of time (data type 2), which reveals that its reliability is less than what was indicated at the design stage; in this example data type 1 and data type 2 are combined to produced the information required.

There exist a variety of mathematical and statistical methods and tools that can be used for the analysis of (historical or real-time) product lifecycle data. A short sample of existing tools is provided in table 4.

Table 4: some existing tool for data analysis

Tool	Description	Reference
OLAP (OnLine Analytical Processing)	Designates a category of applications and technologies that allow data analysis over large collections of historical data (data warehouses), supporting the decision-making process.	http://www.olapreport.com/fasmi.htm http://www.ondelette.com/OLAP/dwbib.htm
S-PLUS® 7	It builds on the S-PLUS platform to meet the rapidly growing demand for better and faster forecasts and estimations, to deliver the capability to handle extremely large data sets, and to integrate the power of predictive analytics into business processes across the enterprise. The goal: to enable organizations to achieve the benefits of quantitative decision making by integrating statistical analysis when, where, and how it is needed.	http://www.gras.de/fileadmin/user_upload/downloads/gras/s-plus/S-PLUS_business_white_paper_final_v3.pdf
Nlighten™	It enables engineers to reveal hidden design and value-improvement opportunities from in-service product field data, and analyze the effectiveness and financial impact of their product design decisions-- past, present and future. It provides the actual field data to product engineers in rapid time, allowing quick and easy access to historical product and parts failure metrics. Additionally, complex statistics are transformed into an easy-to-use and understand format that design engineers and decision makers can quickly act upon with full confidence.	http://www.ninatek.com/lantern_home.asp
CAfdE®	It provides a tool for the analysis of field or test failure data. The failure	http://www.bqr.com/BQR-2005-1.pdf

	data can be collected from various IT tools, such as ERP, CMMS, CRM and other special tools developed by the user. CAfdE® can also be used as the field failure data entry by the technician. It contains algorithms for calculating field MTBF & reliability growth, fixed time and sequential test planning and the results are saved in a core database.	
FRACAS+ Software	The Failure Reporting Analysis and Corrective Action System (FRACAS) can be used to collect record and analyse system failures. The failures are reviewed and corrective actions identified and verified. This process can be used to greatly improve the through-life reliability of the target system.	http://www.isograph-software.com/frcover.htm
Reliability Workbench	It applies many known reliability analysis methods from within a single, fully integrated, program. Isograph's Reliability Workbench now includes the fault and event tree analysis capabilities used by the international reliability community since 1986.	http://www.isographdirect.com/workbench.htm
ITEM ToolKit (Reliability Software)	It contains 5 modules for performing reliability prediction (MTBF) analysis where module is designed to analyse and calculate component, sub system and system failure rates in accordance with the appropriate standard. After the analysis is complete, ITEM ToolKit's (Reliability Software) integrated environment comes into its own with powerful conversion facilities for transferring data to other modules of the program.	http://www.itemuk.com/relpred.html
Weibull++ 6	It is software performing life data analyses utilizing multiple lifetime distributions, including all forms of the Weibull distribution, with an interface geared toward reliability engineering.	http://www.reliasoft.com/Weibull/
ActiveFactory™	It provides data trend analysis, sophisticated numerical data analysis using Microsoft Excel, comprehensive data reporting using Microsoft Word, and the capability to publish this valuable real-time and historical plant data to the Web or company intranet.	http://www.wonderware.com/products/activefactory/
	.	

Several comments follow from Table 4:

- Even if some analysis techniques provided by these tools are suitable for data analysis in some application scenarios, they still need to be integrated in the overall process of transforming product field data into required knowledge;

- As the requirements regarding data analysis in the 11 application scenarios are not yet definitely established it is difficult to determine which tool is suitable to which application scenario.
- The suitability of one tool or another to a given application scenario depends not only on the type of data analysis needed but also on the IT environment on which the data transformation process is to be implemented.
- A characterization of the 11 application scenarios regarding their needs about data analysis can help identify the appropriate data analysis methods and tools for each application scenario.

5.1.7. Logical test

The main objective of this step is to decide whether to continue the transformation process or not. This consists of determining whether there is a matter for concern or not. The detection of a matter of concern means that the process should continue in order to find a solution to the problem detected and if no matter of concern is detected the process may be stopped.

If the objective of data analysis is to detect whether there is problem or not regarding a specific issue, then the analysis should be followed by a logical test to see whether there is a matter of concern or not. However, if the objective of the analysis is to solve a given problem then the analysis will not be followed by any logical test.

In the example of generating knowledge for design for reliability, if the analysis of reliability data reveals that there is no deviation from the expected reliability performance; one can conclude that there is no matter of concern.

5.1.8. Information synthesis & processing

The synthesis and processing of Information ultimately leads to the generation of knowledge. Knowledge is information selected and combined, plus the influence of tacit experiences, ideas, insights, values and judgements of analysts involved in the process.

The information generated from product field data analysis may concern the status, operational condition, specific performance or other issue of a component, part or a system. Sometimes this information needs more processing to yield the necessary knowledge for making decisions and taking actions.

There are various methods that enable the transformation of information into knowledge. Among them we can quote the comparison technique that consists of comparing the obtained information about a given situation to other known situations. Another known technique is the connection technique that consists of connecting the obtained information to other pieces of data, information or knowledge. For example if the analysis of the reliability of a component reveals that the component's reliability is abnormally low, this information should be connected to the search for the main cause(s) of the this low reliability; and the identification of the cause(s) negatively impacting the reliability can be used as knowledge to improve the reliability during re(design) of similar components.

5.2 Collection of knowledge generation requirements of the 11 application scenarios

5.2.1 Introduction

PROMISE is concerned with 11 application scenarios in the Automotive, Railway, Heavy Load Vehicle, EEE and White goods sectors, where 2 are related to BOL (A10 & A11); 6 are related to MOL (A4-A9); and 3 are related to EOL (A1-A3). These 11 application scenarios are outlined in table 5.

Table 5: Application Scenarios

Application Scenario	Partner	Description	Product Lifecycle Phase where knowledge is to be applied
A1	CRF	PROMISE EOL information management for monitoring End of Life Vehicles	EOL
A2	Caterpillar	PROMISE EOL information management for heavy load vehicle decommissioning	EOL
A3	INDYON	PROMISE EOL information management for tracking and tracing of products for recycling	EOL
A4	CRF	PROMISE MOL information management for predictive maintenance for trucks	MOL
A5	Caterpillar	PROMISE MOL information management for heavy vehicle lifespan estimation	MOL
A6	FIDIA	PROMISE MOL information management for predictive maintenance for machine tools	MOL
A7	MTS	PROMISE MOL information management for EEE	MOL
A8	Wrap	PROMISE MOL information management for EEE	MOL
A9	INTRACOM	PROMISE MOL information management for Telecom equipment	MOL
A10	Bombardier	PROMISE BOL information management for Design for X	BOL
A11	Polimi	PROMISE BOL information management for Adaptive Production	BOL

In order to ensure the relevance of the report presented here to PROMISE objectives, an examination of the consequences of applying the knowledge process outlined in the section above for each application scenario is envisaged here. Thus, the knowledge process will be analysed for each scenario in BOL, MOL, and EOL to determine the similarity and distinctness of each application scenario in relation to each other; this will provide useful information so that an existing PDM system may be chosen in TR7.3, which is to use the output of this report as part of its requirement analysis.

Each application scenario has been analysed as to its knowledge content via a specially developed "knowledge application scenario" checksheet which is outlined in the next sub-sub-section (5.1.10). Sub-sub-section 5.1.11 presents the knowledge generation requirements of the 11 application scenarios and sub-sub-section 5.1.12 presents the results of the analysis of the knowledge generation requirements of the 11 application scenarios.

5.2.2 Knowledge Application Scenario Checksheet

The knowledge application scenario checksheet provided the basis for an examination of each application scenario on the basis of the knowledge process outlined in section 4 above. The knowledge application scenario descriptions were completed by the WP R7 participants based upon existing documentation, in particular, deliverable R3.2 which developed the original application concepts.

The knowledge application scenario template consisted of 8 areas, which the participant had to fill in for each application scenario in order to provide an overview of the knowledge issues involved. The format roughly followed that of the knowledge process outlined in section 5 above; that is, it consisted of the question areas outlined in table 6.

Table 6: Knowledge Application Scenario Template

Knowledge Application Scenario Question area	**Detailed Question(s)**
Objectives of obtaining the required knowledge	• What are the purposes for which the knowledge to be generated from the field data in the application scenario is used for? What are the end results that the application scenario is trying to achieve? • In which form the knowledge to be generated from the field data in the application scenario should be provided?
Determination of the relevant data	• What are the different types of the gathered field data (to be gathered) needed to obtain the required knowledge? What systems for data retrival are in place? Do these systems provide data in the correct context? Is the data/information to be used for generating the required knowledge located in only one source? • What are the data systems in which the field data is (should be) stored? • How the relevance of data is (should be) determined?
Access, retrieval of relevant data	• What systems are (should be) used to access the relevant data? • What systems are (should be) used for data retrieval? Do these systems provide data in the correct context?
Integration of data	• In which forms the field data are available? • How is the data integrated? Can uniformization be ensured?
Validation of data	• What external conditions can influence the validity of data? • What are the subjective factors in the gathered field data (for example, human judgements)? • What factors can influence the accuracy of data (incompleteness, incorrectness, censoring, etc.)?
Data processing / analysis	• What methods and tools are (should be) used for the analysis of data?
Logical test	• What cases correspond to the existence of a matter of concern?
Information synthesis	• What tools and methods are (should be) used to generate knowledge from the information obtained through data processing/analysis?

5.3 Knowledge generation requirements of the 11 application scenarios

The results of the knowledge application scenario template for each of the 11 application scenarios is given in detail in Appendix A; table 7 below outlines the main points of these results.

Table 7: Knowledge Application Scenarios—Results

	Knowledge generator type 1			Knowledge generator type 2						Knowledge generator type 1	
Application Scenario	**A1**	**A2**	**A3**	**A4**	**A5**	**A6**	**A7**	**A8**	**A9**	**A10**	**A11**
Objective	PEID assessment in EOL ELVs to effectively implement the component removal decision	Focus on information gained during EOL events and how rigorous management of the information can improve EOL responsiveness; also provide feedback to BOL and MOL functions	PEID assessment in processes of recycling plastic material	Predict optimal time for maintenance activity, in particular oil change, air filter change, brake change	Aggregation of available engineering data, filed data and ancillary information	Traceability	Predict remaining lifetime for components of a gas boiler to derive recommended actions for service engineers	Improve service for refrigerators by predicting failures, Maintain efficiency levels in terms of energy consumption.	Support the engineers and technicians to decision making about product improvements and problems solving/preventive maintenance.	Generation of DfX knowledge to support engineers in improving various aspects of the design such as reliability, availability, maintainability, life cycle cost, environment and safety.	How the PROMISE platform can be used to improve the overall enterprise performance by adapting the production system to the large number of product and process modifications prompted by the availability of feedback information concerning the whole product life cycle.
Determination	4 field data types determined	Performance data from from machine during use to BOL and/or MOL	Field data provided by raw material, field data generated	Three different types of data: environment, usage, and	Based on the objective	Generic and Maintenance data	Three different types of data: internal status of product, usage, and	Types of field data: usage, environment	Data from three different sources (EMS, SIS, Call Tracking System)	Field data to use for generating DfX knowledge is captured by CM (Condition Monitoring) /CBM (Condition Based	Five field data types are outlined.

			during the processes (bin descriptors) and expert knowledge (tacit and explicit)	maintenance.			environment.			Maintenance), FRACAS (Failure Reporting Analysis and Corrective Action System), Service and/or PEIDs (Product Embedded Information Devices).	
Access	PDA reader to PLKM system maintained by the dismantler	Data stored on PEID; PDA reader??	Mobile RFID readers to WMS maintained by the producer	Wireless transmission to maintenance engineer at garage and to ground station	Data stored on PEIDs	Computer	Internet based access from PDKM, wireless PEID reader could be used	DSS accesses field data by proxy device	Internet based access to PDKM and PDAs.	Most of the data to be used for generating the DfX knowledge is available within a specific database or file system which means that this data can be accessed electronically.	N/A
Integration	Assume no uniformization issues	Multiple types of data, but within this uniformization may be assumed?	N.A.	Can be assumed	Can be assumed.	Already done	Can be assumed.	N.A.	Can be assumed	Field info database should integrate all kinds of field data and provide them in form suitable for analysis activities.	Cannot be assumed
Validation	Subjective factors kept to a minimum; external impacts on field data may affect validity though	Not specified	N.A.	N.A.	Dependent on the data analysis method	Not necessary	N/A	N.A.	Failure of sensors, quality and accuracy of data registered by the technicians,	The validation of the data to be analyzed is mainly dependent on the data analysis method to be considered.	System reconfiguration
Data processing / analysis	Data processing via a threshold method: i.e. does component reach certain quality/cost etc.	Data processing via software methods	Data processing by rule based systems	Prediction algorithm	Predictive methods	Mathematical analysis	Predictive maintenance algorithm	Predictive maintenance algorithm	Knowledge management for translation and transformation of data to knowledge. DSS to support the decision making about problems and product improvements.	Statistical/mathematical methods.	Data processing dependent upon the tools by different sets of people involved.

Logical test	Test is whether the correct removal decision for each component is being implemented	Not specified	Development of product quality and costs	Remaining lifetime of a component below defined threshold	Not specified	Faults	Remaining lifetime of a component below defined threshold	Values exceeding thresholds for compressor efficiency, cooling circuit pressure, and internal temperature	Further analysis required	In this step, it has to be decided whether the data analysis results reveal an underlying design problem or not. For example, only if the reliability index of the observed component is abnormally low, then there is a matter of concern in which case investigations to know what causes it is necessary.	Optimality of a suggested new process layout.
Information synthesis	DSS allows comparability of 4 field data types: law , quality, cost, inventory to produce information synthesis and knowledge of removal decision	Software?	Data processing by rule based and expert systems	On the on-board diary, and possibly in the PROMISE Decision Support System.	Broadly described	Mathematical - Statistical	In the PDKM and in the DSS system.	Performed in the DSS.	In the PDKM and in the DSS system.	An appropriate processing and synthesis of all necessary information – under governance of the appropriate DfX specialist and supported by DSS – can generate the required knowledge.	In PROMISE Decision Support System

5.4 Results of the analysis of the requirements of application scenarios

The knowledge generation requirements of the 11 application scenarios are outlined in full in Appendix A. Where possible the checksheet was completed by participants that were directly working upon the application scenario itself; this improved the validity of the results returned. It is worth noticing that not all the required information is provided as the end-users have not yet a clear-cut view of their requirements regarding the transformation of product field data into knowledge.

Table 7 presents the generalised set of the results that have been deduced for the 11 application scenarios. Note that the table has been divided into two types of knowledge generator: type 1 and type 2. These knowledge generator types may be envisaged as in figure 11; the key difference between both is the reliance upon real time data (in reality "sufficiently small" interval times data) plus eventual past data and auxiliary data or past field data (i.e. data that has already been gathered and stored in data systems). Note that this is a generalisation of a predominant characteristic of the application scenarios: some application scenarios may have recourse to using the opposing knowledge generator in certain circumstances; however the knowledge generator type specified here is the one that is the principal type in use for each.

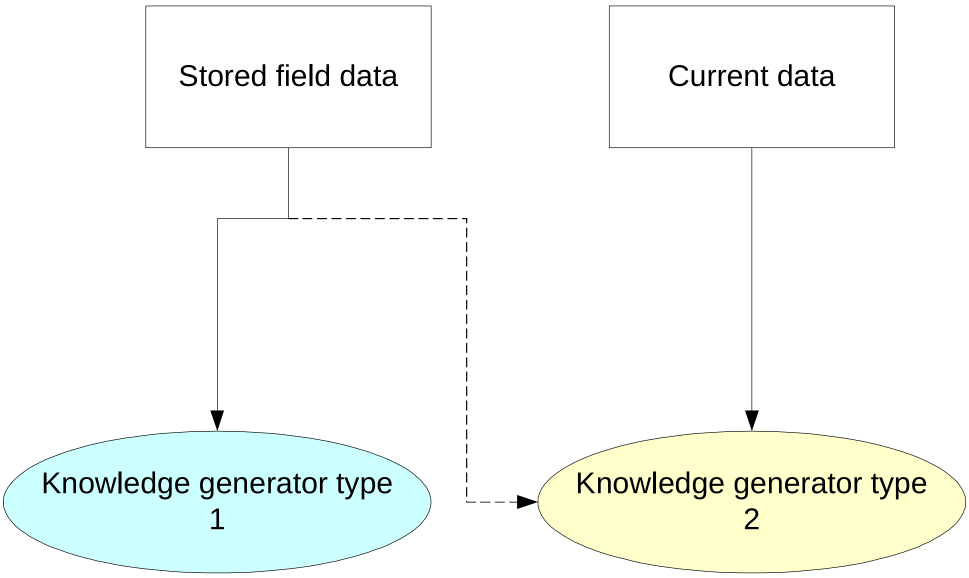


Figure 11: Knowledge generator type 1 & 2

The two different types of knowledge generator distinguished in table 7 are: type 1 (A1-A3 and A10-A11) and type 2 (A4-A9). The knowledge generator types recognise two main processes of generating knowledge:

- one that uses field data that is already gathered and stored in a database; and
- one that considers real time data (in reality "sufficiently small" interval times data) plus eventual past data and auxiliary data mainly for monitoring the operational condition of the system for use in predictive maintenance.

Further analysis reveals that the knowledge generator type 1 may be depicted as a flowchart as in figure 12. Here the majority of the process steps of the model for the knowledge process outlined previously in section 5 are repeated. Note the input of the analyst at each stage of the process from raw field data to data validation, and from data validation to information synthesis to knowledge generation. Note also the role of the personnel with the relevant knowledge of the objective towards which the field data is to be used: they have control over the data validation and choice of the relevant field data steps. This type of knowledge generator is based primarily upon field data and the transformation of this into knowledge with the assistance of the experience, ideas and knowhow of expert personnel in the field.

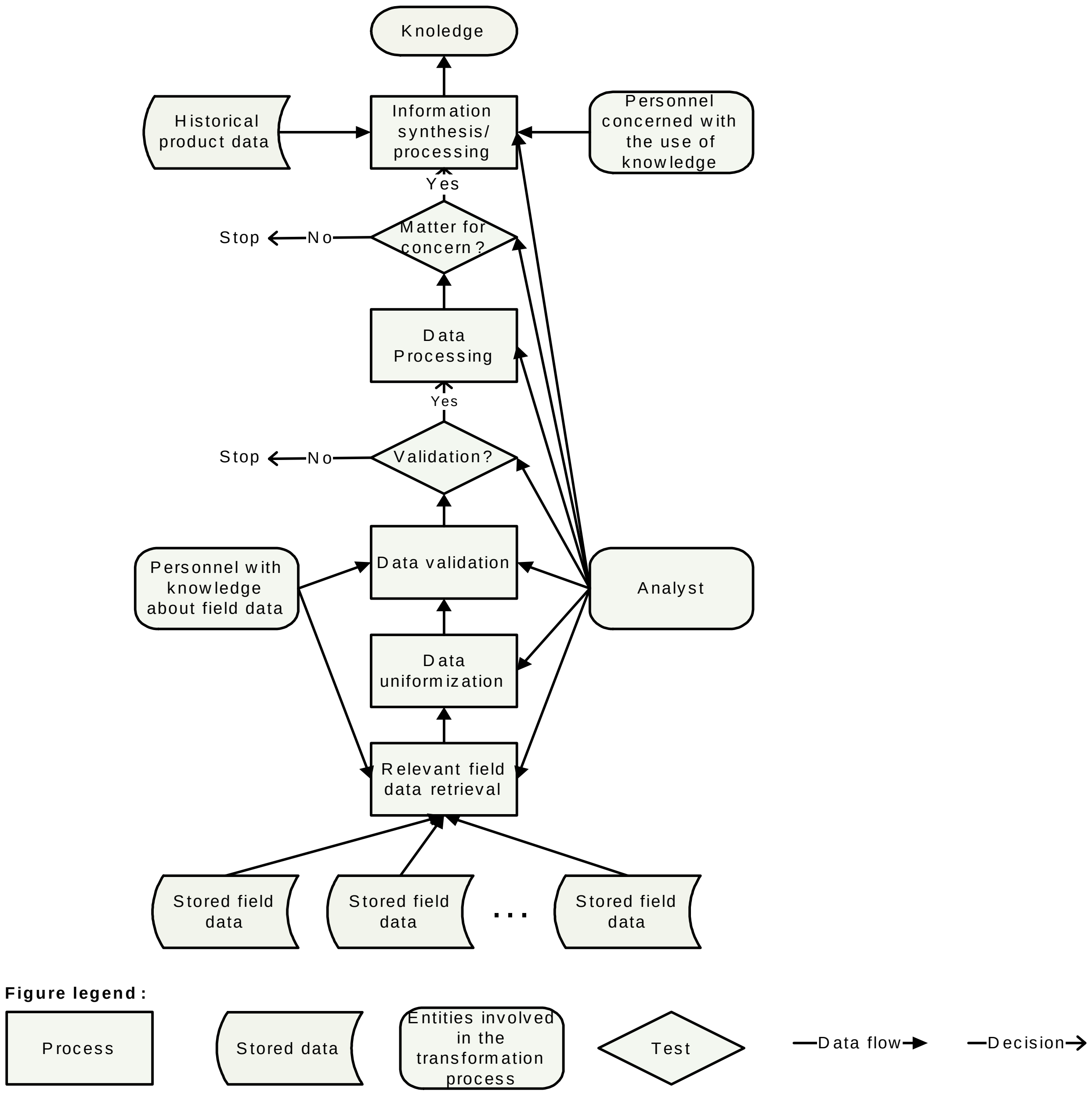


Figure 12: Knowledge generator type 1

This knowledge generator may be contrasted with knowledge generator type 2 (see figure 13), which uses real time field data and a host of other historical field data to achieve knowledge. Note that decision-support acts as a structuring tool from which knowledge may be generated as outlined previously by Hicks et al. (2002) in Figure 4 above. A diagnostic process is generally followed with the various types of data used in the process, before a prognostic answer—achieved through the aid of decision-support—is supplied; this becomes the produced “knowledge”.
The main purpose of the diagnostics module is to assess through on-line sensor measurements the current state and operational condition of the critical components in the system (Vachtsevanos and Wang, 1999). The measurement of the current state and operational condition of the critical components in the system can be achieved through the continuous measurement of the values of a set of appropriate parameters (Biagetti and Sciubba, 2004).
The main purpose of the prognostics module is to analyse the input from the diagnostics module and historical field data using appropriate models in order to draw a picture for the current situation and potential consequences for the future (Vachtsevanos and Wang, 1999).
In the case of predictive maintenance which is an important concern for many application scenarios considered in PROMISE, the prognostics module has the function of linking the diagnostic information and the maintenance scheduler (Vachtsevanos and Wang, 1999).

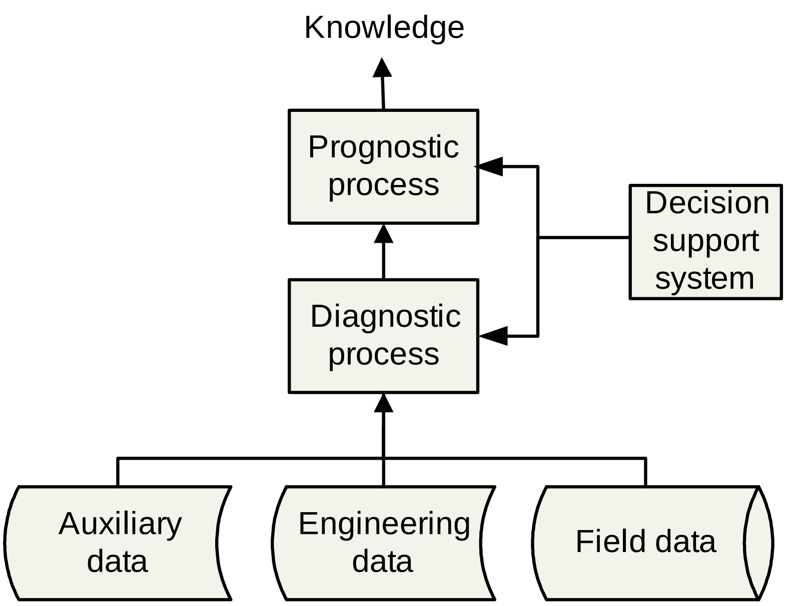


Figure 13: Knowledge generator type 2

Among the techniques used to deal with the diagnostics/prognostics problems, Vachtsevanos and Wang (1999) quote Stochastic Auto-Regressive Integrated Moving Average (ARIMA) models, fuzzy pattern recognition principles, knowledge-intensive expert systems, nonlinear stochastic models of fatigue crack dynamics and polynomial neural networks. Specific tools such as the watchdog Agent™ developed by the Center for Intelligent Maintenance Systems at the University of Wisconsin can be considered (Lee, 2003).

In Table 8, Schroer (2002) provides a comparison of the different approaches that can be used for diagnostics/prognostics problems where he emphasises the strengths and weaknesses of each approach.

Table 8: Comparison of approaches for diagnostics/prognostics problems (Schroer, 2002)

Approach	Strengths	Weaknesses
Rule-Based	• Easy to understand due to their intuitive simplicity • Well-proven, with many deployed applications. • Inference sequence can be easily traced. • Shells are widely available which makes the development easier	• Development and maintenance can be long and time-consuming • Knowledge acquisition bottleneck. • Generally only faults anticipated during the design phase can be diagnosed
Models based on structure and behaviour	• As the model is a "correct" model, theoretically all faults can be diagnosed, however in practice, this is difficult to achieve. • With the appropriate software it should be possible to generate models from CAD data.	• Computationally intractible on models with large numbers of components. • Generating adequate behavioural models for complex devices (e.g., a microprocessor) is a serious challenge. • Developing a complete and consistent model is difficult. For example, a "correct" model will not be able to diagnose a bridging fault. • Knowledge of fault types is often not included and this can lead to the diagnosis of nonsensical faults. • Development times can be long.
Diagnostic Inference Models	• Provides good diagnoses when good sources of diagnostic information are available • Fairly well proven with many deployed applications.	• Only usable where good sources of diagnostic information are available, therefore diagnostic issues will have to be considered at design time.
Case-Based	• A fairly intuitive and easy to understand process • The knowledge acquisition bottleneck can be overcome as learning is continuous and incremental.	• Can only diagnose once an adequate casebase has been built. • It is not always apparent how diagnostic inferences are arrived at. • Typically, collected cases are domain specific and cannot be generally applied.
Fuzzy logic and neural networks	• Good at dealing with incomplete and inaccurate information	• As a sole approach their ability to diagnose complex systems is questionable, however, combined with other approaches a good solution may be possible.

6. Further Analysis—Application Scenarios

The generalised results emanating from the analysis of the 11 application scenarios may be outlined under their two generic knowledge generator types: 1 and 2. This is performed in the next two sections (6.1 and 6.2), and the assumptions and hypothesis used to make these generalisations are outlined in section 6.3.

6.1 Type 1 knowledge generator

The Type 1 knowledge generator is depicted again in Figure 14. The following are notes and generalised results observable for this generator type.

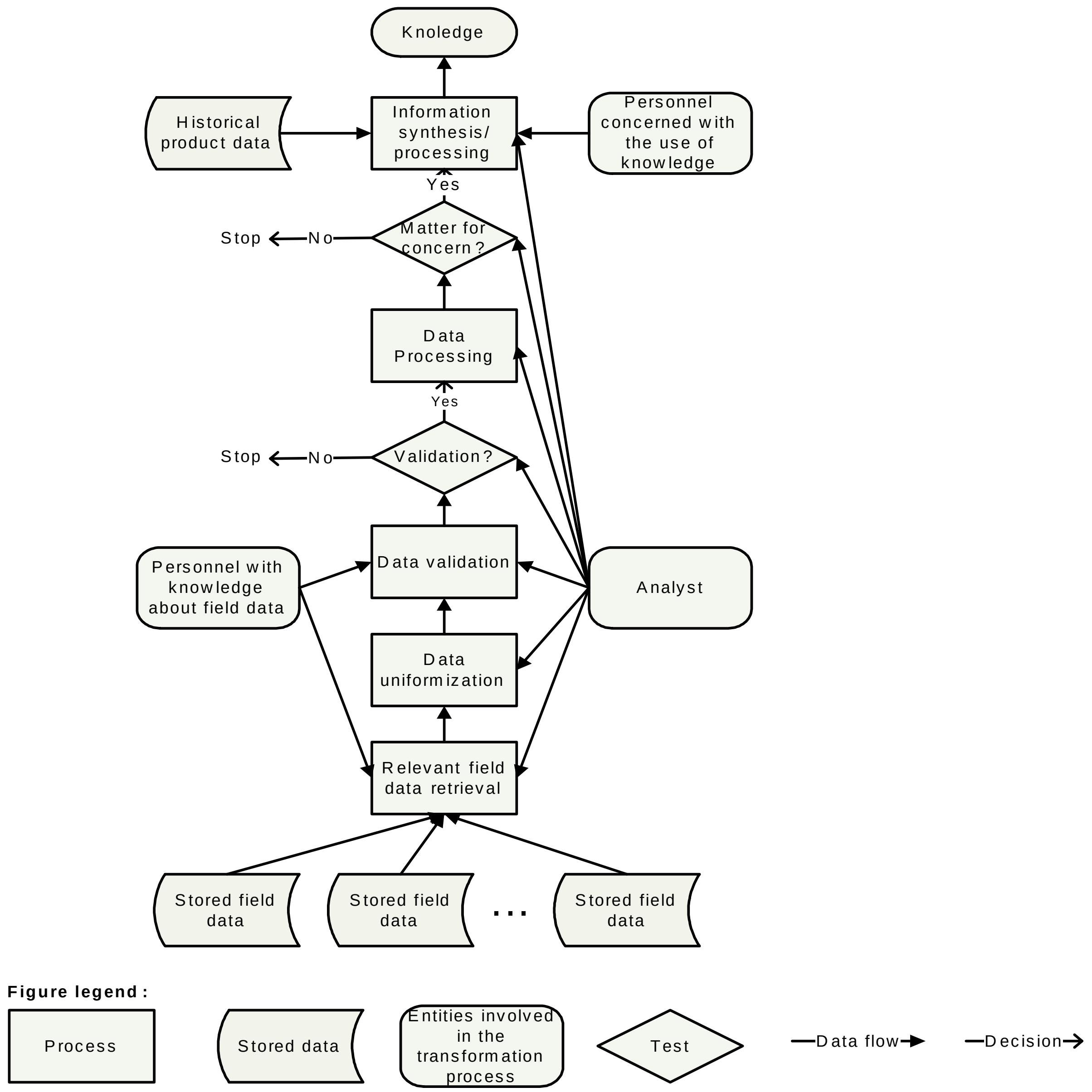


Figure 14: Knowledge generator type 1

Objective:

With respect to the application scenarios that are focussing pre-dominantly upon a usage of previously stored (past) field data (knowledge generator type 1), there are clearly defined objectives for both BOL and EOL phases of the product lifecycle. The objectives of the BOL phase are focussed upon the continual upgrading of existing knowledge bases to improve DfX techniques, and the improvement of process designs to support the implementation of the production process. The objectives of the EOL phase are chiefly concerned with an examination of existing resource recovery applications in place in order to support an efficient (from both economic and environmental points of view) treatment of EOL products, and the maintenance of these together with the additions of new techniques to improve the EOL environment to enable more efficient feedback to both MOL and BOL, and the use of PEID technology to improve resource recovery (see Figure 15).

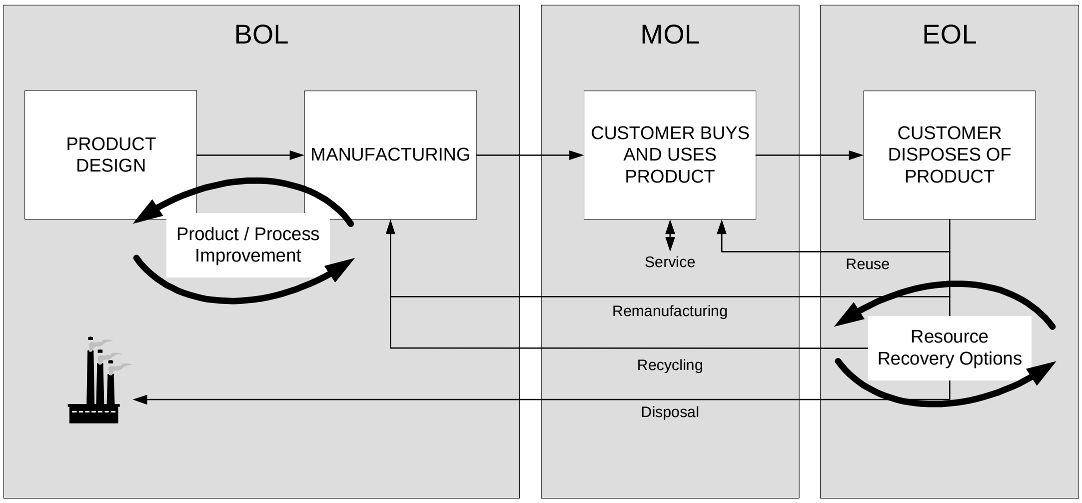


Figure 15: Type 1 knowledge generator – objectives

Determination:

Determination consists of identifying the relevant product field data for achieving the objectives of data transformation.
Within these objectives, a large variety of data types are determined by the knowledge process in the type 1 knowledge generators. The EOL application scenarios are predominantly concerned with data types associated with the PEIDs used in the EOL arena: data types include environmental and usage data upon vehicles and construction equipment, and usage data upon recyclables from the automotive industry. These data types are used as an aid to the pre-existing data stored in the EOL practitioner's databases to make resource recovery decisions concerning EOL issues, such as: inventory tracking & tracing, reuse of components, recycling materials, disposal etc. In the BOL arena, data determination centres upon the location of the requisite data to perform correct *product* and associated *process* procedures for the design and manufacturing of the product. Data determination here examines DfX issues—such as reliability, availability, maintainability, lifecycle cost, safety, environment, etc. of the product; whilst issues such as Life Cycle Analysis, Line Balancing, and the upgrading of the existing product processes—which is dependent upon what is being manufactured—results from these decisions.

Access:

Accessing the requisite data as a step of the knowledge process for type 1 knowledge generators requires a considerable outlay in auxiliary applications. For example in EOL, the chief process for the determination of the correct data to be accessed requires the use of PEIDs (such as RFIDs) on the product itself to store the requisite data, PDA (Product Data Acquisition) readers to allow for the transferral of the data from the RFID to the EOL practitioner's back-end system, and various software components located in the back-end system to process the data further—these are common requirements for all of the EOL application scenarios. Similarly for the BOL application scenarios: product design processes, utilising DfX procedures, face a range of auxiliary applications that support the access and retrieval of data from MOL (maintenance data from CM (Condition Monitoring) /CBM (Condition Based Maintenance), FRACAS (Failure Reporting Analysis and Corrective Action System), Service and/or PEIDs (Product Embedded Information Devices) and from other sources, such as PDM, databases, and standards and regulations (from various sources); and also *process* design, which is dependent upon both the product design and the manufacturing environment to specify the require access technologies to obtain the relevant data.

Integration:

In the type 1 knowledge generators, the issue of data integration is currently ambiguous. In the BOL environment, with the uncertainty of the development of *manufacturing processes* required to support production and the development of the product, it is impossible to be sure of issues concerning the integration of various types of data that *may* be required—the situation is entirely context-specific, making it impossible to determine, in advance, whether there will be a wide variation in the data types available or not; a similar situation exists for the development of DfX procedures for products: integration, of sorts, will take place in a dedicated database that will "integrate" the existing field data—this pre-supposes the existence of nicely-conformable data upon which integration rules may be applied. In the EOL arena, an assumption is being made that the data may be successfully integrated; the assumption is based upon the premise that the relevant data may be grouped into a number of "types" thus facilitating integration—for example, the data types for EOL include cost data, time data, quality data, maintenance data etc. Within these data types themselves, integration is not further assumed.

Validation:

The type 1 knowledge generators require validation of their data to ensure that the information to be generated from it is consistent and useful. For the BOL application scenarios it is clear that certain problems may cloud the validity of the data that has previously been regarded as integrated: DfX principles are entirely dependent upon the data analysis methods used to aggregate and align the data to its objectives, whereas system reconfigurations and changes in personnel, products, capacities etc. may have an adverse impact upon the validity of the data emanating from the manufacturing process itself. For those application scenarios that reside in the EOL arena, the imposition of subjective factors—such as subjectivity in data collection, incomplete databases etc.—may result in validation issues with the collected data, despite the fact that it is integrated. Some of the application scenarios for EOL are currently assuming that validation issues will not be a concern though.

Data processing/analysis:

The data processing/analysis methods that are to be applied by the type 1 knowledge generator application scenarios are similar in origin for EOL and BOL. For each component in EOL a threshold method combined with various "performance metrics" that may be evaluated is favoured: data processing into information is done via a system that sets various "thresholds" for key performance indicators that are successively measured and compared against existing database information; if the performance indicators pass a certain "threshold"—for example, a given threshold of quality, cost, labour time etc.—then the component passes to the next threshold; this process continues to select components that "pass" and are recovered, and those that don't and are recycled or disposed. For BOL with regard to DfX principles, mathematical and statistical methods are the favoured method for data processing and analysis; the input data is interrogated by various mathematical and statistical algorithms to determine their processing into information. Note that for A11, the context-specific nature of the application scenario makes it difficult to determine the exact data processing and analysis that the data will undergo; this will have to be decided in real-time.

Logical Test:

With regard to the type 1 knowledge generators, a logical test determines whether there is a matter for concern in the development of appropriate knowledge from the information derived. This is handled in broadly similar ways for the EOL spectrum, but differently by the BOL spectrum. In the EOL arena, the chief test is to determine whether the correct resource recovery decision is being implemented; this may involve an analysis of whether the correct component is being removed from a vehicle, or construction equipment, or whether sufficient product quality and cost resource recovery rules have been determined for the recycling of materials. In the BOL arena, the optimality of a new manufacturing process must be tested via simulation and other techniques to determine if the new design should be implemented; for the product designed from DfX principles, the logical test concerns the detection of whether the data analysis results reveal an underlying design problem or not. For example, only if the reliability index of the observed component is abnormally low, then there is a matter of concern in which case investigations to know what causes it is necessary.

Information processing / synthesis:

For the type 1 knowledge generators information processing and synthesis takes place in a similar manner to data processing and analysis. For each component in EOL a threshold method combined with various "performance metrics" that may be evaluated is favoured: information processing into knowledge is done via a system that sets various "thresholds" for key performance indicators that are successively measured and compared against existing database information; if the performance indicators pass a certain "threshold"—for example, a given threshold of quality, cost, labour time etc.—then the component passes to the next threshold; this process continues to select components that "pass" and are recovered, and those that don't and are recycled or disposed. For BOL with regard to DfX principles, if a matter of concern is resulted from the data analysis then this information should be worked out further by considering other data and information in order to determine the nature of problem, what caused it and its potential consequences. An appropriate processing and synthesis of all these information can generate the required knowledge. Note that for A11, the context-specific nature of the application scenario makes it difficult to determine the exact information synthesis that the information will undergo; this will have to be decided in real-time, after the various data analysis processes have been completed.

6.2 Type 2 knowledge generator

The Type 2 knowledge generator is depicted in Figure 16. The following are notes and generalised results observable for this generator type.

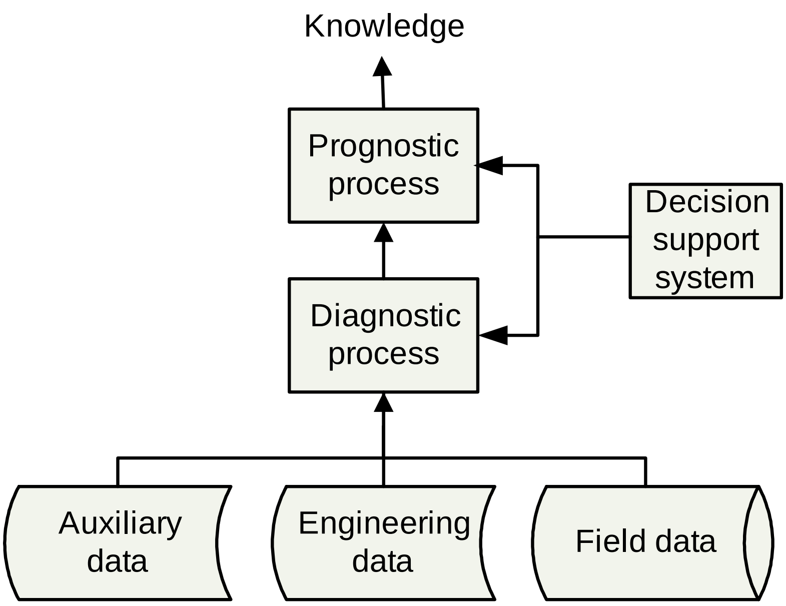


Figure 16: Knowledge generator type 2

<u>Objective:</u>

With respect to the application scenarios that are focussing pre-dominantly upon a usage of real-time (in practice, at "sufficiently small" interval times) field data together with past field data (knowledge generator type 2), there are clearly defined objectives for the MOL phase of the product lifecycle. The objectives of the MOL phase are focussed upon predictive maintenance, product life estimation, traceability and service issues including demand forecasting for spare parts that require continual monitoring in the MOL phase. The availability and usage of data in the MOL phase that has previously been lost owing to no fixed method for its recording, has now become available thanks to PEID techniques of data storage; the continuous monitoring of the state and operational condition of the product is now possible which enable the generation of knowledge for supporting predictive maintenance, demand forecasting for spare parts, product life estimation, etc. (see Figure 17).

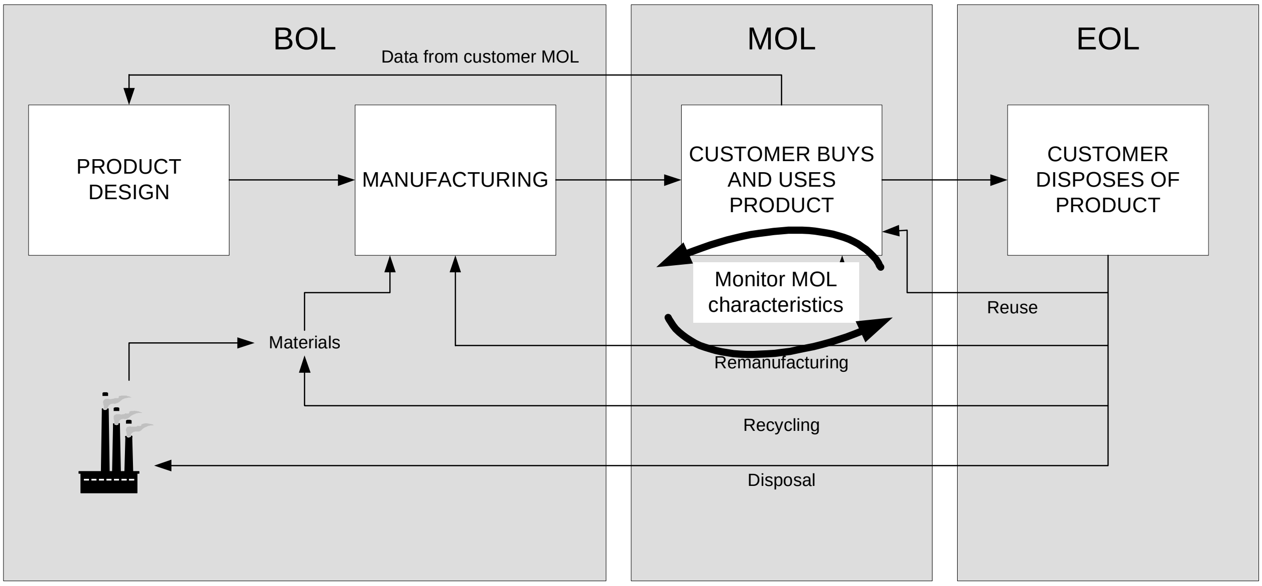


Figure 17: Type 2 knowledge generator – objectives

Determination:

Within these objectives, a large variety of data types are determined by the knowledge process in the type 2 knowledge generators. The MOL application scenarios are predominantly concerned with data types associated with the PEIDs used in the MOL arena: data types include environmental and usage data for predictive maintenance; data for traceability requirements; and service and conditions-based data. These data types are used as an aid to develop the databases of the MOL practitioners so that they can make decisions concerning MOL issues, such as: lifetime-of-components predictibility, tracability of components in MOL, decision support for maintenance personnel, and optimal maintenance activity predictions.

Access:

Accessing the requisite data as a step of the knowledge process for type 2 knowledge generators requires a considerable outlay in auxiliary applications, from field devices in the proximity of the PEIDs on the MOL product, to back-end systems to support decision making. The chief process for the determination of the correct data to be accessed requires the use of PEIDs (such as RFIDs) on the product itself to store the requisite data, PDA (Product Data Acquisition) readers to allow for the transferral of the data from the RFID to the MOL practitioner's back-end system, and various software components located in the back-end system to process the data further—these are common requirements for all of the MOL application scenarios. Technologies mentioned range from Internet access at the back-end system level, to wireless and mobile hand-held technologies that are useful for technicians taking data samples in the field.

Integration:

In the type 2 knowledge generators, the issue of data integration is currently assumed. In the MOL environment, with the pre-defined data types that facilitate integration—for example, the data types for MOL include maintenance data, reliability data, availability data etc.—allow for the assumption that the data to be collected is sufficiently integrated to make no difference to the measurement arena. However an assumption is being made that the data may be successfully integrated; the assumption is based upon the premise that the relevant data—grouped into a number of "types"—will not have internal integration issues also.

Validation:

The type 2 knowledge generators require validation of their data to ensure that the information to be generated from it is consistent and useful. Again in the MOL phase, validation issues are currently being assumed for many of the application scenarios; data validation is taken as given if personnel successfully monitor the PEIDs in the field correctly for the pre-defined data types stipulated at BOL *before* the product entered its MOL. Thus validation is assumed to be pre-built into the system by the design team at BOL; issues that may affect validity of data, however, include the defect operation of PEIDs, of PDA readers etc. or the inappropriate use of the back-end system software applications available to the MOL practitioner.

Data processing/analysis:

The data processing/analysis methods that are to be applied by the type 2 knowledge generator application scenarios are similar for all MOL application scenarios. The MOL environment comprises of data analysis techniques that involve both the use of *predictive algorithms* and *mathematical analysis* of the feedback from the PEIDs that are located upon the product in the field. The diagnostics/prognostics techniques are the most suitable for this kind of problems (more details about the diagnostics/prognostics are provided in sub-section 5.4). For each component in MOL, data is collected that is transformed into information via a process of data analysis; the data is interpreted by algorithms that allow the essence of the data to be extracted to enable the development of maintenance and service information that can impact upon the MOL product. This information can then be further developed to derive knowledge. This knowledge can materially affect the MOL product in the form of the maintenance activity, servicing, safety testing etc.

Logical Test:

With regard to the type 2 knowledge generators, a logical test determines whether there is a matter for concern in the development of appropriate knowledge from the information derived. This is currently an ambiguous issue for the MOL application scenarios. The most widely accepted interpretation of the logical test involves the setting of "thresholds" that, if passed by test data, will trigger a certain response from the system. The MOL data—maintenance data, reliability data, availability data etc.—may be supplied with benchmarks which act as thresholds; these thresholds may be held in the MOL practitioner's back-end system and data from the field may be analysed against these. If the data fails a certain threshold, this provokes a response from the system: i.e. make a recommendation to the user to replace a component with a lower-than-expected lifetime likelihood; remove components with faults; service component whose working-condition data deems this to be necessary; and maintenance of a product whose data fails threshold test etc.

Information processing / synthesis:

For the type 2 knowledge generators information processing and synthesis takes place in a similar manner to data processing and analysis. The MOL environment comprises of information analysis techniques that involve both the use of *predictive algorithms* and *mathematical analysis* of the feedback from the PEIDs that are located upon the product in the field. If a matter of concern is detected through the data analysis, then the search of causes of the problem, the potential consequences and their severity is crucial. The synthesis and processing of the information resulted from the data analysis with other information from past field data, engineering data etc. through comparison, combination or other techniques and the use of experience and ideas and values from expert personnel upon the information allow generating the required knowledge for supporting predictive maintenance or other purposes in the type 2 knowledge generator.

6.3 Hypothesis and Assumptions

The above research must be contextualised in order to be understandable to the reader; in this section we specify the main hypotheses and assumptions that are adopted in this deliverable in order to define the boundaries of our work.

- In this deliverable, we assume that the field data is gathered by personnel, sensors, RFIDs or other PEIDs as described in the different application scenarios. Consequently we are not concerned in this deliverable about how to gather field data;
- As the knowledge to be generated is intended to be used for various purposes and in different lifecycle phases and the field data to be used for the generation of knowledge is of different types and from different lifecycle phases, it is difficult to develop a detailed approach for generating knowledge from field data that can be suitable to all the application scenarios considered in PROMISE. Consequently the different steps of the overall approach to the transformation of data into knowledge will be detailed to the level of providing the different methods and tools that can be used in addition to guidelines about the type of problems for which each method/tool is suitable.
- The requirements for the definition of the main steps of the concepts of generating knowledge from field data will be based on the description of the PROMISE application scenarios.
- In this deliverable we do not deal with the issues of knowledge storage, knowledge management, knowledge usage, etc.
- In this deliverable we do not consider the issues related to where the different steps of the approach are taking place, which tools are used to support them and how the inputs and outputs are transferred between them.

7. Conclusions

This deliverable contributes to one of the main objectives of PROMISE that consists of closing the information loop between the different product lifecycle phases. To achieve this goal this deliverable considers the exploitation of product field data gathered at different product lifecycle phases in order to generate appropriate information and knowledge that can be used for different purposes and in different product lifecycle phases.

The main issue addressed in this deliverable is the one that consists of transforming product field data into useful knowledge where the main objective being the identification of models able to satisfy the requirements of the 11 application scenarios considered in PROMISE regarding the transformation of product field data into specific knowledge.

Regarding the 3 key concepts of data, information and knowledge that are intensively used in this deliverable we privileged a pragmatic understanding of these concepts that is closer to the PROMISE context far from the abundant philosophical debates about the differences and relationships between them.

To define the models that are needed for transforming product field data into suitable knowledge, we followed an approach composed of three main steps: (i) the collection of the requirements regarding the transformation of product field data into knowledge of the 11 application scenarios considered in PROMISE, (ii) classify these requirements into common requirements and specific requirements and (iii) define the models according to the common and specific requirements. We required that the models be "complete" in the sense that they respond to the requirements of all the application scenarios regarding the transformation of data into knowledge, "non-redundant" in the sense that they have quite distinct features that make them different from one another and "minimal" in the sense that their number should be kept to a minimum while preserving "completeness" and "non-redundancy".

The collection of the requirements regarding the transformation of product field data into knowledge of the 11 application scenarios considered in PROMISE is done using a checksheet addressing the main issues related to the transformation of product field data into knowledge. These issues are: purposes for which the knowledge to generate is to be used, determination of the relevant data needed for the generation of the required knowledge, access and retrieval of the relevant data, integration of data originated from different sources, validation of data before starting the analysis/processing, data analysis/processing, logical test to detect whether there is matter of concern or not and finally synthesis/processing of information to obtain the required knowledge.

From the data collected about the requirements regarding the transformation of product field data into knowledge of the 11 application scenarios using the checksheet described above, it appears that some of the issues are not described, some others are described but not with enough details and only few issues in some application scenarios are described with the required details. This fact has negatively affected the level of details in which the knowledge generator models are described.

After the analysis of the requirements regarding the transformation of product field data into knowledge of the 11 application scenarios considered in PROMISE, we come up with two different models; one considering real time field data and historical field data (knowledge generator type 2) and one considering only historical field data (knowledge generator type 1).

The knowledge generator type 1 is concerned with the generation of knowledge to support various decision making processes mainly in BOL ad EOL and the knowledge generator type 2 considers mainly the diagnostics/prognostics models to support predictive maintenance, product life estimation and demand forecasting for spare parts.

The lack of detailed requirements regarding the transformation of product field data into knowledge in the description of the 11 application scenarios prevents this deliverable from providing detailed descriptions of the steps through which the product should go in order to provide the required knowledge in the two models.

It was expected that WP R7 and other research WPs provide the application WPs with appropriate input regarding the methods and tools to be integrated in the 11 demonstrators considered in PROMISE however the lack in detailing the requirements regarding some issues such as the transformation of product field into knowledge yield that the contribution of the research WPs may be less than expected. Notwithstanding this fact, it can be stated that this deliverable has paved the way for the 11 demonstrators to efficiently describe their requirements regarding the transformation of product field data into useful knowledge which in their turn can provide valuable input to the continuation of task R 7.1 after the first 18 months.

References

Aamodt, A., Nygård, M., 1995, Different roles and mutual dependencies of data, information, and knowledge — An AI perspective on their integration, Data & Knowledge Engineering, 16, 3, pp. 191-222.

Alavi, M, Leidner , D., 1997, Knowledge Management Systems: Emerging Views and Practices from the Field, INSTEAD Working Paper, Fontainbleau, 1-31.

Badaracco, J.L. (1991) *The Knowledge Link - How Firms Compete Through Strategic Alliances,* Harvard Business Press, Boston.

Bassi, L. (1998) Harnessing the power of intellectual capital*, The Journal of Applied Manufacturing Systems,* Summer, pp 29-35.

Biagetti, T., Sciubba, E., Automatic diagnostics and prognostics of energy conversion processes via knowledge-based systems, Energy 29 (2004) 2553-2572.

Blake, P. (1998) The knowledge management expansion*, Information Today,* 15, 1.

Bohn, R. (1994) Measuring and managing technological knowledge*, Sloan Management Review,*Fall, pp 61-73.

Browne, J., Sackett, P. J., Wortmann, J. C., Future manufacturing systems—Towards the extended enterprise, Computers in Industry 25(1995) 235-254.

Carayannis, E.G. (1999) Fostering synergies between information technology and managerial and organisational cognition: The role of knowledge management, *Technovation,* 19, 4, pp 219-231.

Carlsen, A. and Skaret, M. (1998) Practicing knowledge management in Norway - Lessons from processes in small firms, *Working Paper*, Sintef, Norway.

Churchman, C.W. (1971) *The Design of Inquiring Systems: Basic Concepts of Systems and Organization*, Basic Books, New York.

Clark, D., http://www.nwlink.com/ ~donclark/performance/understanding.html, accessed on March 2005.

Compton P. and Jansen, R., A philosophical basis for knowledge acquisition. In: J. Boose, B. Gaines and J.G. Ganascia, Editors, *EKAW 89, Third European Workshop on Knowledge Acquisition for Knowledge-Based systems* (1989), pp. 75–89 (ECCAI/AFCET/ARC) .

Davenport, T.H. (1993) *Process Innovation: Reengineering through Information Technology*. Harvard Business School Press: Boston M.A.

Davenport, T.H. and Prusak, L. (1998) *Working Knowledge: How Organisations Manage What They Know,* Harvard Business Press, Boston.

Davenport, T.H., De Long, D.W. and Beers, M.C. (1998) Successful knowledge management projects*, Sloan Management Review,* Winter, pp 43-57.

Davenport, T.H., Jarvenpaa, S.K. and Beers, M.C. (1996) Improving Knowledge Work Processes. *Sloan Management Review*, Summer, pp 53-65.

Drucker, P. (1993) *Post Capitalist Society,* Harper Business, New York.

Glazer, R. (1991) Marketing in an information intensive environment: Strategies and implications of knowledge as an asset*, Journal of Marketing,* 55, pp 1-19.

Grant, R. (1996) Prospering in dynamically competitive environments: Organisational capability as knowledge integration*, Organization Science,* 7, 4, pp 375-387.

Gunasekaran, A. (1999) Agile manufacturing: A framework for research and development*, International Journal of Production Economics,* 62, 1-2, pp 87-105.

Harris, K. (1999) So, you're a knowledge worker: What Do You Do All Day? *Internal Report*, Gartner Group.

Hicks, B. J., Culley, S. J., Allen, R.D., Mullineux, G., A framework for the requirements of capturing, storing and reusing information and knowledge in engineering design, International Journal of Information Management, Vol. 22, Issue 4, 2002, Pages 263-280.

Jauw, J., Vassiliou, P., Filed data is reliability information: implementing an automated data acquisition and analysis system, Proceedings of the annual reliability and maintainability symposium, 2000, pp. 86-93.

Kiritsis, D., Bufardi, A., Xirouchakis, P., Research issues on product lifecycle management and information tracking using smart embedded systems, Advanced Engineering Informatics 17(2003) 189-202.

Knock, N., McQueen, R., and Corner, J. (1997) The Nature of data, information and knowledge exchanges in business processes: Implications for process improvement, The Learning Organization, 4, 2, pp 70-80.

Kotnour, T., Orr, C., Spauding, J. and Guidi, J. (1997) Determining the Benefits of Knowledge Management Activities. In *IEEE International Conference on Systems Man & Cybernetics*, Vol. 1 Orlando, Florida.

Laudon, K.C. and Laudon, J.P. (1999), *Management information systems: Organization and Technology in the Networked Enterprise*, Prentice Hall, New York.

Lee, J., E-manufacturing—fundamental, tools, and transformation, *Robotics and Computer Integrated Manufacturing* 19 (2003) 501–507.

Machlup, F., Mansfield, U. (Eds.), The study of information: interdisciplinary messages, (New York: John Wiley and Sons), 1980, 646-647.

Malhorta, Y., 1998, TOOLS @ WORK: Deciphering the Knowledge Management Hype, Journal for Quality & Participation, 21, 4, 58-60.

Martin, J. (1995) The Great Transition; Using the seven disciplines of enterprise engineering to align people, technology and strategy, Amacom, New York.

Meadow, C.T., Yuan, W., Measuring the impact of information: Defining the concepts**,** *Information Processing & Management*, Volume 33, Issue 6, November 1997, Pages 697-714.

Neef, D. (1997) Making the case for knowledge management: The bigger picture, *Ernst & Young LLP Working Paper* Center for Business Innovation, Boston.

Nonaka, I. (1991) The knowledge creating company, *Harvard Business Review*, Nov - Dec, pp 96-104.

Nonaka, I. and Takeuchi, H. (1995) *The Knowledge Creating Company*, Harvard Business Press, Boston.

Oh Y.S., D. S. Bai, D.S., Field data analyses with additional after-warranty failure data, liability Engineering & System Safety, Volume 72, Issue 1, April 2001, Pages 1-8.

Parlby, D. (1997) The power of knowledge: A business guide to knowledge management *Internal Report*, KPMG.

Polanyi, M. (1966) *The Tacit Dimension,* Routledge & Kegan Paul, London.

Quinn, J., Anderson, P. and Finklestein (1996) Managing professional intellect: Making the most of the best, *Harvard Business Review,* 74, pp 71-80.

Ruggles, R. (1998) The state of the notion: Knowledge management in practice, *California Management Review,* 40, pp 80-89.

Ruggles, R.L. (Ed.) (1997) *Knowledge Management Tools,* Butterworth Heinenmann, Boston.

Schroer, R., Revolutionary ideas in test, Aerospace and Electronic Systems Magazine, IEEE Volume 17, Issue 6, June 2002 Page(s):36 - 41

Stewart, T. A. (1997) *Intellectual Capital: The New Wealth of Organizations,* Doubleday.

Sveiby, K.E. (1997) *The New Organizational Wealth: Managing and Measuring Knowledge Based Assets,* Berrett-Koehler, San Fransisco.

Takeuchi, H. (1998) *Beyond Knowledge Management: Lessons from Japan*. www. Sveiby.com, pp 1-11.

Thierry, M., Salomon, M., van Nunen, J. and van Wassenhove, L. (1995) *Strategic Issues in Product Recovery Management*, California Management Review, Volume 1, No. 2, pp. 114-135.

Tuomi, I., Data is more than knowledge: implications of the reversed knowledge hierarchy for knowledge management and organizational memory, Proceedings of the 32nd Hawaii International Conference on Systems Sciences, 5-8 January 1999, Volume: Track 1, Pages: 1-12.

Vachtsevanos, G., Wang, P., An intelligent approach to fault diagnosis and prognosis, The 53rd Meeting of the Society of Machinery Failure Prevention Technology, MFPT Forum, Virginia Beach, April 19-22, 1999.

Wiig, K.M. (1995) Knowledge management: Where did it come from and where will it go?, *Expert Systems with Applications,* 13, pp 1-14.

Wikstrom, S., Normann, R., Anell, B., and Ekvall. G. (1994) *Knowledge and Value: A New Perspective on Corporate Transformation,* Routledge, London.

Wilmering, T.J., Yuan, J., VanRossum, D. (2003) A metadata architecture for mediated integration of product usage data, Proceedings of the of the IEEE Systems Readiness Technology Conference, AUTOTESTCON 2003, 22-25 September, Pages: 564-575.

Wilson, D.A. (1996) *Managing knowledge.* Butterworth Heinnmann- Oxford.

Appendix A: Knowledge Application Scenarios

The following sections outline the 11 knowledge application scenario descriptions derived from the application scenarios.

WP A1: PROMISE EOL information management for monitoring End of Life Vehicles

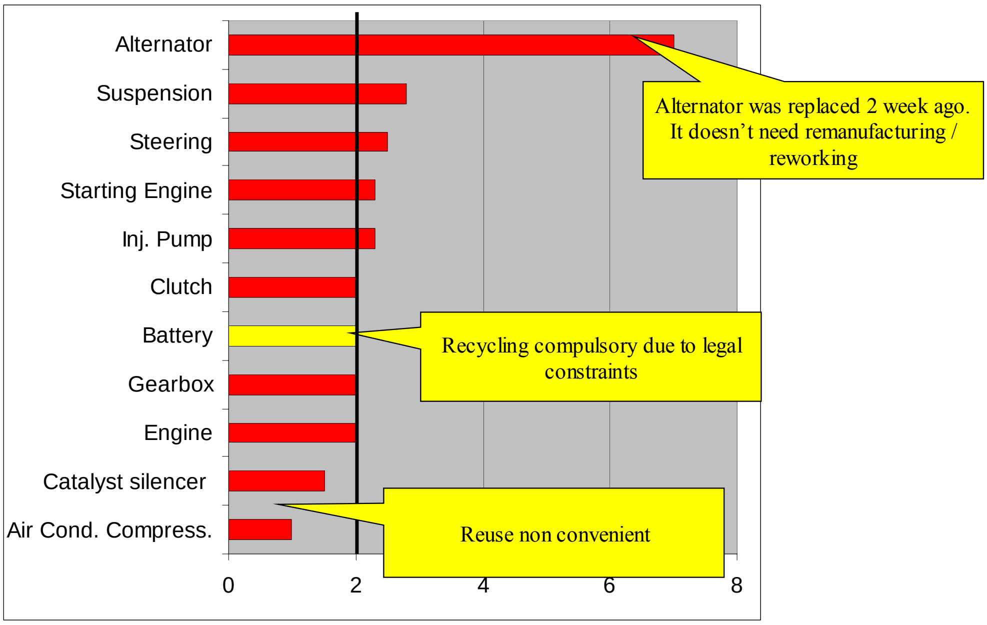
Objectives of obtaining the required knowledge

What are the purposes for which the knowledge to be generated from the field data in the application scenario is used for? What are the end results that the application scenario is trying to achieve?
Assess the use of PEIDs for improved decision-making in End of Life process management, and materials tracking. The output of the decision-making system is a list of components to be removed from the vehicle for reuse, remanufacturing, and recycling.

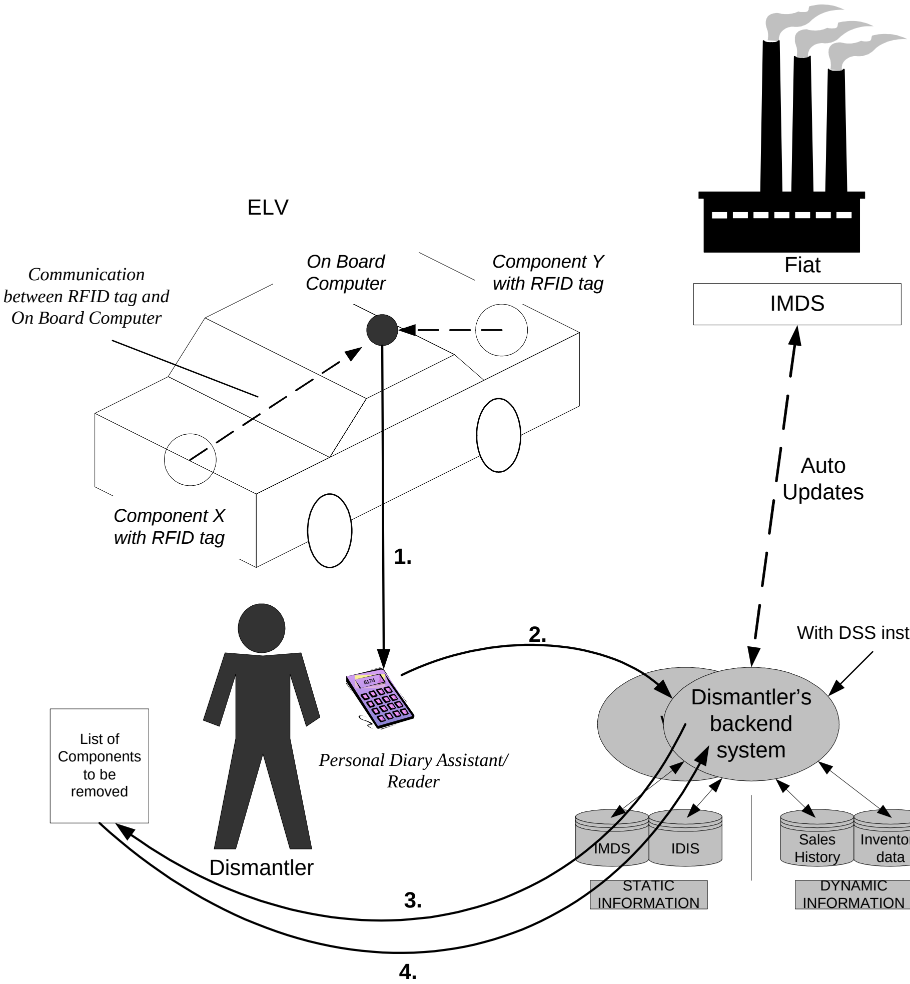
In which form the knowledge to be generated from the field data in the application scenario should be provided?
The output from the field data is a list of components to be removed (see figure). The component list is compiled from assessing ELVs that are input into the EOL process, and providing a "score" for each ELV component based upon its quality, cost to remove and current stock levels being retained.

LIST OF COMPONENT TO BE REMOVED

(IN DESCENDING ORDER OF WORTH REUSING SCORE)



Determination of the relevant data

What are the different types of the gathered field data (to be gathered) needed to obtain the required knowledge? What systems for data retrival are in place? Do these systems provide data in the correct context? Is the data/information to be used for generating the required knowledge located in only one source?
Raw data is collected from the on-board computer, which is connected to individual PEIDs (probably RFIDs) on individual components in the ELV (see figure below).  Inside the ELV, there is communication between individual RFID tags on components, and the on board computer. • Using a reader, the dismantler identifies the car via the on board computer, which provides vehicle details and information on the quality (field data type no. 1) of components (1.); • The dismantler interacts with their back-end system (2.), where the DSS resides—the DSS assess 3 further criteria other than quality: does the component HAVE to be removed by law (field data type no. 2—a list of components (such as batteries)—either YES or NO); does the component cost (field data type no. 3—dependent on current labour costs, market demand etc.) less to remove than will be fetched for it in the market; and are there many of these components in stock (field data type no. 4—

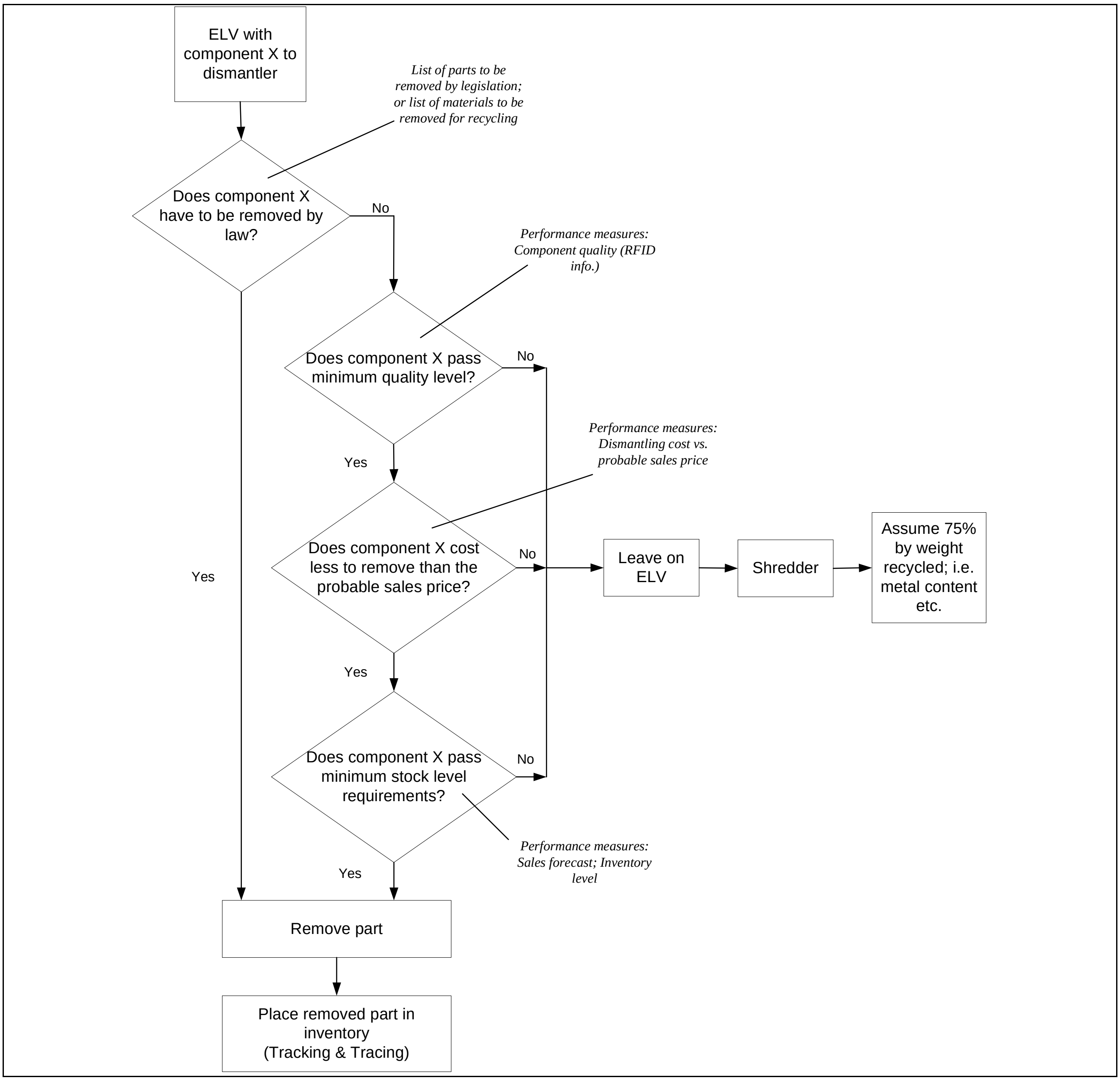
inventory levels);

- An ideal list of components to be removed is generated (3.) from the DSS;
- When the dismantler completes the job of dismantling the ELV components, any changes that had to be made to the printed list are transmitted back to the dismantler's backend system (4.);
- Automatic Updates from Fiat (A.) including inputs from Fiat via the Internet (for example IMDS), are part of a continuous stream of supporting information available to the dismantler;
- Automatic Updates from dismantler to Fiat (A.) on the dismantling / reuse rates on a periodic basis to allow compliance to the ELV directive.

Thus there are 4 field data types:
Quality—on the individual component in ELV;
Law—static list from EU;
Cost—in EOL dismantler's yard and current market trends;
Inventory—in EOL dismantler's yard.

ELV is source of quality data; Inventory and cost data are derived from the dismantler; Law comes from EU via Fiat.

What are the data systems in which the field data is (should be) stored?
Quality data is stored on the ELV's on board computer, based upon feedback from individual PEIDs on components. When the on board computer is accessed is accessed, the quality information is moved to the dismantler's back end system (PLKM) which consists of static and dynamic components. Static components include dismantling instructions and IMDS; while dynamic components include sales history and inventory data, as well as automatic updates from Fiat concerning legislation govering the removal of components by law, changes in dismantling instructions, new component information, etc. The various types of field data are stored in databases and accessed by the decision support system (DSS)—see figure, —which assesses each field data type successively to allow the removal decision to be made (i.e. whether a part should be removed for reuse / remanufacturing or left on the vehicle for material recycling).



How the relevance of data is (should be) determined?

The relevance of the data that is being returned is assessed via set of performance measures (see sample table below):

Information Required	Source of data / information
Dismantling costs	Inventory data; sales data
Sales forecast	Sales history over a specified period of time
Component quality	RFID tag
Inventory level for particular component	Inventory data
Probable sales price	List price (Fiat); Sales history (dismantler); quality of component

These performance measures enable the DSS to determine the best removal choice for each component, and determine whether the field data being returned is quality-oriented, cost-oriented, inventory-oriented, etc.

Access, retrieval of relevant data

What systems are (should be) used to access the relevant data?
Quality data from the on board computer on the ELV, it is envisaged, will be accessed using a PDA reader. Once in this format, the quality data can be passed on to the dismantler's backend system, where it will be stored in a database that will hold information on a number of entities (dependent on the component), such as: Mission profile statistics Maintenance history Environmental conditions etc.

What systems are (should be) used for data retrieval? Do these systems provide data in the correct context?
The on board computer retrieves data from the individual PEIDs (most likely RFID tags) that are placed upon individual components. Data retrival will be looking for those types of raw data as previously defined. A relatively consistent context is assumed.

Integration of data

In which forms the field data are available?
As has been stated previously, there are four types of field data impacting upon the DSS: Law Quality Cost Inventory. Each of these data forms are analysed separately in the decision module, thus uniformization and integration issues are avoided.

How is the data integrated? Can uniformization be ensured?
In the DSS there are four types of field data considered: Law—either component X (X being any component in the ELV) must be removed by law or not; a simple YES or NO here; If YES, then the quality of the component is considered. Quality will be based upon RFID tag data to determine whether the component falls within a pre-determined quality threshold;. The cost module examines the dismantling cost vs. probable sales price for the component and the decision to remove will continue to inventory if it is feasible. The inventory module takes into account the current state of the dismantler's stock levels against possible sales forecasts. As can be seen, the data is strictly segregated into law-type, quality-type, cost-type, and

inventory-type, with no cross-integration considered. Uniformization of the data at the performance measure level may be assumed for each individual component.

Validation of data

What external conditions can influence the validity of data?
The validity of the data can be impacted by changes to each of the four data forms. For example, with Law, a new component added to the list that HAVE to be removed, influences the validity of the list as it was used previously in a negative way. Similarly with the other field data forms: Quality of components as reported by RFID tags and actual quality by a visual inspection on the part of the dismantler, may give rise to discrepences between the list of components to be removed generated by the DSS, and what is actually taken out of the ELV. Similarly, significant changes to any of the cost measures used, or the inventory structures in place, may require retooling of individual field data forms.

What are the subjective factors in the gathered field data (for example, human judgements)?
Subjective factors are kept to minimum. As has been stated above, a visual inspection of parts by experienced dismantlers may be used to confirm the component's quality.

What factors can influence the accuracy of data (incompleteness, incorrectness, censoring, etc.)?
Incomplete databases of data means that adequate DSS modules for each of the four field data forms may be difficult to develop. This would result in guesses as to the adequacy of the component under investigation. For example, if a previous history of the quality of component X has not been built up, it may take some time and wrong decision-making to determine the pre-defined level of quality (for example) that should be set for the component. Obviously affects newly-introduced components most.

Data processing / analysis

What methods and tools are (should be) used for the analysis of data?
The methods that are to be used to analyse the field data are previously described in the decision flowchart above. There are four main decision modules that use a set of pre-developed performance measures in each module as a toolkit to measure the effectiveness of that module. The method is based upon previous results derived for the same performance measures being used to forecast the proper removal decision for the present component. Decision rules are to be developed based upon these performance measures whereby (for example) the quality of a component may be termed "excellent", "good", "moderate", or "poor" based upon a pre-determined level of quality being reached by the part in question. Similarly for the cost module: the component must reach a certain threshold whereby it can be determined whether it will be profitable to harvest the component for reuse / remanufacturing or to leave it on the ELV for recycling. Whether a component should be introduced into inventory will be determined by market demand, and the number of existing components already in inventory.

Logical test

What cases correspond to the existence of a matter of concern?
The DSS is tested by determining whether the individual field data modules are running correctly

–i.e. are the correct parts being chosen according to their quality? Are the proper costs taken into consideration in the cost module? Matter for concern: if the DSS is unable to stop the recycling or disposal of components that could have been reused or remanufactured earlier in the proceedings.

Information synthesis

What tools and methods are (should be) used to generate knowledge from the information obtained through data processing/analysis?
The overall knowledge objective is to determine the "status" of any particular component in an ELV—i.e. whether it is to be reused, remanufactured, recycled or disposed of. The DSS is the key information synthesis tool to be developed so that various types of comparability may be performed upon each component against previously-held knowledge to allow for the assignment of the component under examination to its proper "material flow". The tracking & tracing of this material flow is not included in A1.

WP A2 : PROMISE EOL information management for heavy load vehicle decommissioning

Objectives of obtaining the required knowledge

What are the purposes for which the knowledge to be generated from the field data in the application scenario is used for? What are the end results that the application scenario is trying to achieve?
The purpose of this scenario is to identify the basic framework for implementing the PROMISE End of Life (EOL) methodology on construction and mining equipment. The application scenario focuses on information that is gained during EOL events and how rigorous management of the information can improve EOL responsive to the event as well as provide feedback to BOL and MOL functions and tracking of total life cycle information (see appendix B, DR3.2, page 19). The primary objective of the proposed scenario is to manage the waste stream for MOL activities. In addition this information can provide feedback to the design and manufacturing sources as well as management to make the PLM processes more robust (see appendix B, DR3.2, page 19). The scenario will be used to identify, test, and document the information requirements, component requirements, information flow, and business case relative to life cycle management. The focus of the scenario is EOL responsiveness to customers needs and commercial drivers. However, there is a requirement for a systematic approach for identifying the opportunities to convert the data that is gathered during the defined MOL process into useful knowledge to better manage the design, production, and waste management processes. In this context the waste management processes includes recycling, remanufacturing, and disposal. Standard systems must be developed where possible to facilitate data flow, material flow and data management with a end goal to maximize reuse and minimize disposal within a viable economic model (see appendix B, DR3.2, page 19). A total number of four objectives is sought achieved by this demonstrator (see appendix B,

DR3.2, page 20):

PROMISE main objective #1: *To develop new closed-**loop life cycle information flow models** for BOL, MOL and EOL.* This scenario will use information relative to component life and failure modes gained during MOL to enhance the design process in BOL. It will use field population data and implied demand to enhance the logistics information for the component providers in the remanufacturing phase of BOL. It will also provide for the study of waste stream data to optimise EOL processes.

PROMISE main objective #2: **To develop new PLM system and IT infrastructure exploiting the capabilities of smart product embedded information devices. Embedded devices will form the bases of the data and information tracking during the MOL event that triggers the process generating the waste stream for EOL management. These devices will continue to be used during EOL to track and document data relative to the logistics and validation through the supply chain.**

PROMISE main objective #3: To develop new standards to allow the technologies and associated tools to be developed by the PROMISE project to be accepted by the market and allow it to expand quickly by creating an appropriate environment for the development of new innovative applications. **Standards will be required to convert the event data into a actionable information package. In addition the scenario will support the need for standards in device and information protocols. New standards must address the need for recycle and reuse parameters within an acceptable economic model.**

PROMISE main objective #4: To develop new working and business models appropriate for the use and exploitation of the new technologies and tools to be developed by all actors involved in a product lifecycle. **The scenario will be used to identify, test, and document the information requirements, component requirements, information flow, and business case relative to life cycle management. This will include processes and information management that will facilitate EOL activities including quantification of recyclable content and processes to validate proper levels of recyclable content as well as socially acceptable disposal processes.**

In which form the knowledge to be generated from the field data in the application scenario should be provided?
The A2 demonstrator will aggregate all kind of field data, product specific information, and additional information and transform it – supported by decision support systems – into knowledge that can be used to take productive and effective actions. The data must be accessible by the supply chain in a PDKM system. This demonstrator covers the closure of the information loop between product operation (MOL) and product recycle, reuse, remanufacture or disposal (EOL). However, to effectively impact these focus areas the demonstrator will also have to close the loop with manufacturing and design (BOL) (see D3.2, page 12).

Determination of the relevant data

What are the different types of the gathered field data (to be gathered) needed to obtain the required knowledge? What systems for data retrival are in place? Do these systems provide data in the correct context? Is the data/information to be used for generating the required knowledge located in only one source?
This scenario is primarily related to EOL activities. However, it impacts both BOL and MOL process as previously stated. Information will be collected from the machine during use. When onboard data processing determines that there is an "event", this event data will be transmitted to the appropriate source. For example, a major failure should transmit information directly to the

service people (MOL). Logistics information would also be sent out so the replacement part(s) (BOL) could be put in route to the destination of the failure. Manufactures would also be contacted in the case that no parts are available or if the supply of the needed part(s) falls below a designated quantity (MOL). Other importance performance data could be transmitted to the service people and/or designers to help understand how to determine the source of the problem or improve the design. (BOL). As components , assemblies, and machines are replaced in the MOL phase a waste stream is generated that will transfer the focus to EOL processes (see appendix B, DR3.2, page 20).

What are the data systems in which the field data is (should be) stored?
Information into EOL will be tagged by a RFID linked data base. This information will establish the bases for EOL decisions relative to reuse, recycle, and disposal. The first objective would be to reuse as much of the components , assemblies, or machines as possible. Often this will require some remanufacturing so information flow back to BOL will be critical to optimizing the process (see appendix B, DR3.2, page 20).

How the relevance of data is (should be) determined?

Figure 1 describes the decision process for the EOL sector:

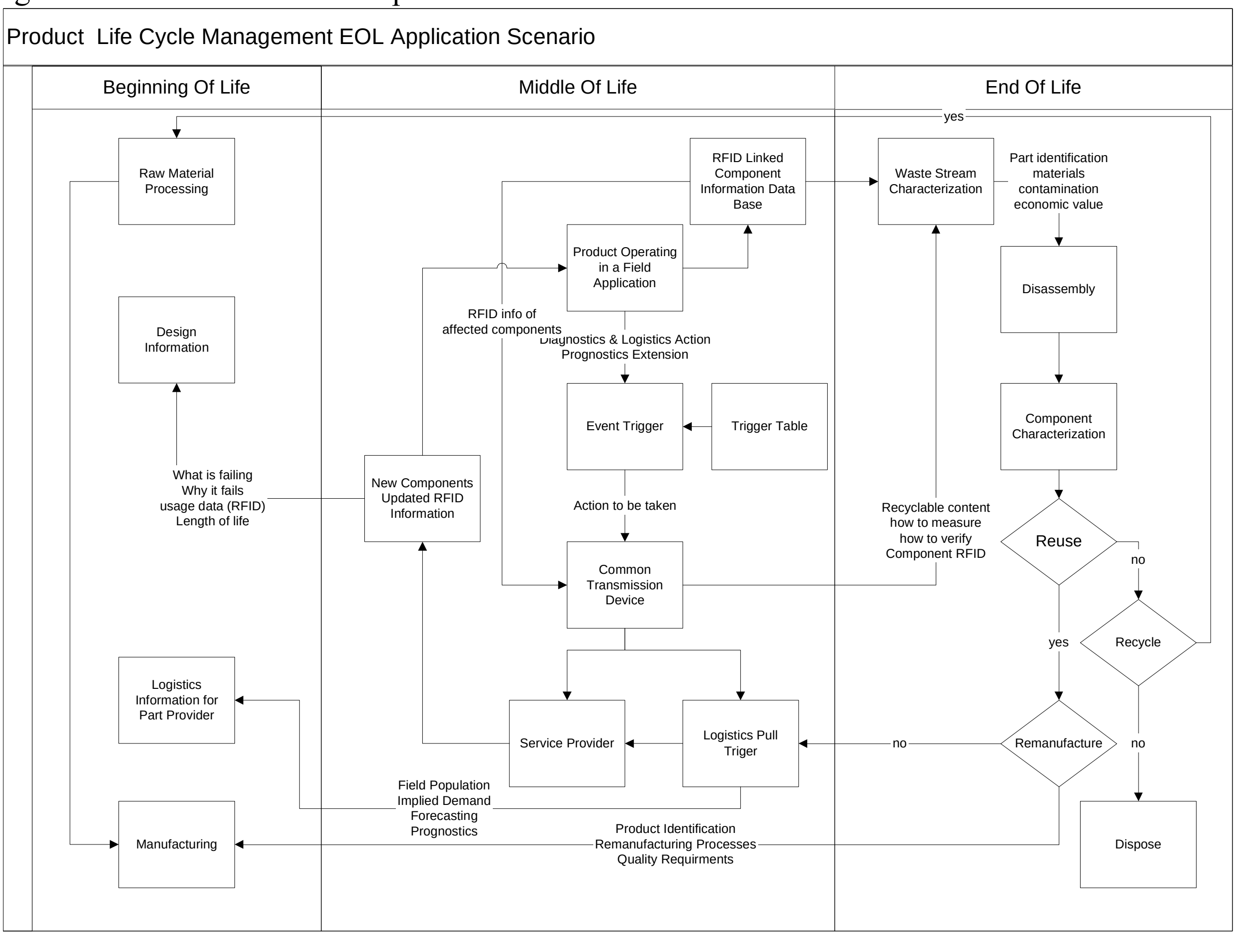


Figure 1: Flow chart of the EOL Application Scenario (see appendix B, DR3.2, page 23)

There are several requirements for this scenario that must be fulfilled for it to be successful. The real challenge is to develop standard methods and protocols that can be used for a number of different applications. The process start point is defined as a TTL or TTT operating in the field as illustrated in the top of the MOL box in Figure 4. In fact this can be characterised as any machine with some diagnostic and prognostic capability operating in its designed application (see appendix B, DR3.2, page 24).

There are a number of decision point in the EOL processes of reuse, recycle, and dispose depicted

in Figure 4. The information requirements and processes definitions related to these decision points have to be identified. However, each application will be different and must be identified by the end user of the system. The final challenge will be to fully define the information flow between the MOL process and the BOL and EOL processes. Some high level concepts are included in Figure 4. However, these will have to be further defined and specified by the PROMISE team (see appendix B, DR3.2, page 24).

Access, retrieval of relevant data

What systems are (should be) used to access the relevant data?
Data will be collected from PEID sensors through out the life of the product, component, and piece part. Most of the data will be stored in data bases that are associatively connected to the PEID via an identification number. In some cases the data may be stored on the PEID itself. This will have to be determined on each individual product. This data or information will be used in EOL to characterize the components and/or piece parts. The data will be used to retrieve the most up to date processes and procedures for disassembly, remanufacturing, recycling, and disposal. The database will contain information such as part identification, material, contamination, duration of life, and service conditions (see appendix B, DR3.2, page 25).

What systems are (should be) used for data retrieval? Do these systems provide data in the correct context?
At this point, it is not clear where or what data will be stored. The application scenario needs a bit further refinement to come to that stage (see appendix B, DR3.2, page 26).

Integration of data

In which forms the field data are available?
• Time/Date stamp • Event type • Relevant Event data • Action Required • Product serial number (TTT or TTL) • Component serial number (Specific component in question) • Machine hours • Component hours • Relevant sensor information (condensed or raw data) • Maintenance information • Misc. user input - . . If there are large amounts of sensor data that cannot be transmitted, an event could be triggered which informs a service man to come and manually collected the needed data from the machine and clear the storage device (see appendix B, DR3.2, page 25).

How is the data integrated? Can uniformization be ensured?
In a final product, all decisions would be made by proprietary Caterpillar software, however we will need a solution for the PROMISE demonstration. Possibly a portable PC could be placed on board the product to perform data analysis and storage for the demo. The need for storage would most likely be needed only in the case that the volume is too great to transfer with the chosen communication device (see appendix B, DR3.2, page 27).

Validation of data

What external conditions can influence the validity of data?
• Caterpillar is a large global organisation with products in use on every continent. This would require systems to be produced in many languages and the system would be required to handle very large amounts of data. • Caterpillar machines are serviced by independently owned dealers. Implementation and training of such a system will/would require a large investment with these dealers. • Some customers may perceive that Caterpillar is spying on them in order to avoid paying warranty claims (see appendix B, DR3.2, page 28).

What are the subjective factors in the gathered field data (for example, human judgements)?
For EOL consideration the business relationships must be established primarily in the recycle and disposal arena. Any component or part of a component that has potential recycle value must be identified as such. The recycle process must be fully developed and documented. A business to take the part and convert it to a somewhat original form must be identified and the process has to be established to transport theses parts to the appropriate location. The same relationship may be required for reuse process. However, many organizations have internal remanufacturing facilities (see appendix B, DR3.2, page 30).

What factors can influence the accuracy of data (incompleteness, incorrectness, censoring, etc.)?
Caterpillar machines work in very rugged conditions. Both extreme heat (55 °C) and extreme cold (-30 °C) conditions are encountered. Vibration, impact, large amounts of dust, oil, rain, mud, etc are also part of the normal operating conditions. These machines work in all weather conditions (see appendix B, DR3.2, page 26).

Data processing / analysis

What methods and tools are (should be) used for the analysis of data?
Software should be user configurable and relatively open so that the user (Caterpillar or a dealer) can customise it to fit a specific customer's needs. It could then also be customised monitor multiple components on a machine (engine, critical structures, etc) (see appendix B, DR3.2, page 26).

Logical test

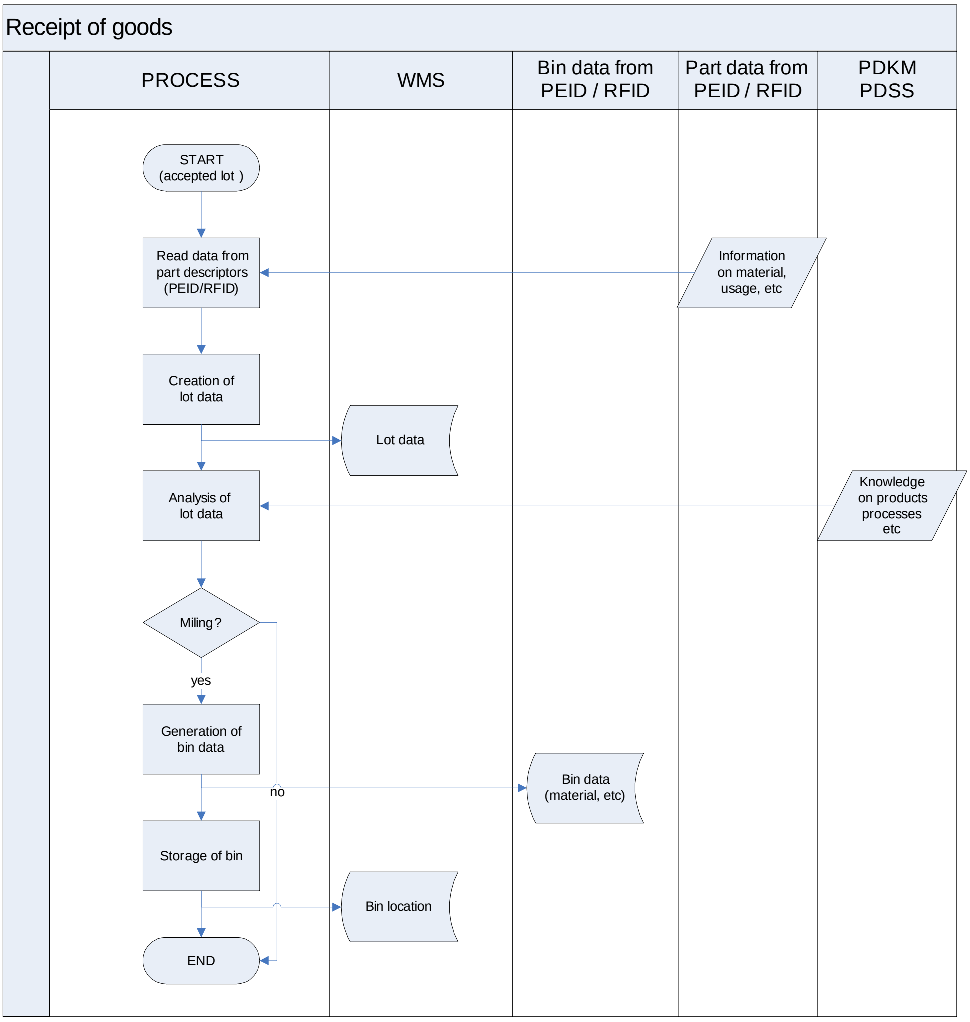
What cases correspond to the existence of a matter of concern?
?

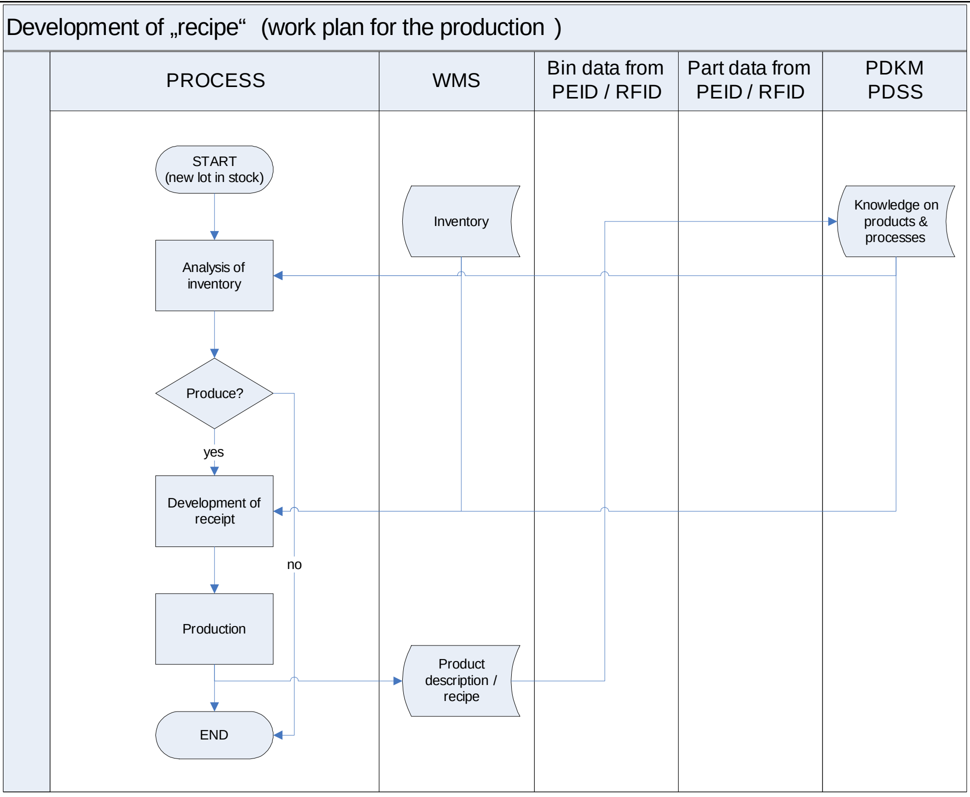
Information synthesis

What tools and methods are (should be) used to generate knowledge from the information obtained through data processing/analysis?
Since Caterpillar machines live for decades (50+ years) in the field, the life should be quite long. This should, at a minimum, match the time to the first major overhaul of the machine where devices could possibly be replaced (see appendix B, DR3.2, page 26).

WP A3 : PROMISE EOL information management for tracking and tracing of products for recycling

Objectives of obtaining the required knowledge

What are the purposes for which the knowledge to be generated from the field data in the application scenario is used for? What are the end results that the application scenario is trying to achieve?
Knowledge generated shall be used for supporting the decision making in plastic recycling processes. The overall process is to mill “old” plastic parts (e.g. car bumpers) and to use the milling product for the production of new raw material. Scenario mainly addresses the decision making during the A) receipt and preparation of raw material and B) recipe development for the production of new raw material. Both processes are depicted in the following: 



In which form the knowledge to be generated from the field data in the application scenario should be provided?
A) 1) Selection and composition of raw material (bumpers) for the building of milling lots 2) Generation of lot bin description (composition and properties of material) B) 3) Proposal for production plan (depending on inventory, market situation, costs) 4) “Recipe” for production lots (products) to be produced

Determination of the relevant data

What are the different types of the gathered field data (to be gathered) needed to obtain the required knowledge? What systems for data retrival are in place? Do these systems provide data in the correct context? Is the data/information to be used for generating the required knowledge located in only one source?
1) PEID / RFID provides data about raw material (e.g. car bumpers) such as material description (kind of plastic), age etc 2) PEID / RFID provide data about the content of a container (e.g. milled car bumpers) such as composition, humidity and about the quality of data. 3) warehouse management system (WMS) provides data about the position of containers and inventory of particular material

What are the data systems in which the field data is (should be) stored?
1) WMS 2) Container PEID

How the relevance of data is (should be) determined?
N.A.

Access, retrieval of relevant data

What systems are (should be) used to access the relevant data?
1) (mobile) RFID readers 2) system for tracking the containers (integrated in the WMS)

What systems are (should be) used for data retrieval? Do these systems provide data in the correct context?
Data retrieval will be done by WMS which are also ensuring the correct context of the data.

Integration of data

In which forms the field data are available?
Bumper: Type of material, colour, surface (paint), etc ***Bins:*** position, composition weight, quality, quality of data ***WMS:*** inventory, position of bins

How is the data integrated? Can uniformization be ensured?
Bumpers will be provided in charges. First decision is on how to proceed with the processing of the material. If either the quality of the material or the quality of the data is not adequate material will not be further handled. If both qualities are sufficient the whole charge will be milled and packed to one or more bins. Based on the data of the single items a description of the bin content needs to be generated. Content of the bins is the raw material for the production. Depending on the inventory of the warehouse decisions shall be generated on if and what to produce. Data required for this decision are the attached to the bins, stored in the WMS or are added to the decision making process by an expert.

Validation of data

What external conditions can influence the validity of data?
Data are partly of dynamic nature such as the description of the bin content which depends on the environmental conditions.

What are the subjective factors in the gathered field data (for example, human judgements)?
Expert knowledge (tacit and explicit) is required for each process step.

What factors can influence the accuracy of data (incompleteness, incorrectness, censoring, etc.)?
Quality of the description of the car bumpers is essential for the whole process. One task of the application scenario is to deal with a certain degree of incompleteness, incorrectness, etc. which means incompleteness is acceptable up to a defined threshold.

Data processing / analysis

What methods and tools are (should be) used for the analysis of data?
Expert knowledge coded in rule based systems and threshold methods shall support the decision on the milling process and the generation of the bin descriptors. Independent from the methods applied it is required to integrate the tacit expert knowledge to the overall processes.

Logical test

What cases correspond to the existence of a matter of concern?
Degree of product quality in correlation with related costs.

Information synthesis

What tools and methods are (should be) used to generate knowledge from the information obtained through data processing/analysis?
Any kind of "Expert" and / or rule based system, means a system which guides a user through a process, provide decision support etc based on explicit knowledge.

WP A4 : PROMISE MOL information management for predictive maintenance for trucks

Objectives of obtaining the required knowledge

What are the purposes for which the knowledge to be generated from the field data in the application scenario is used for? What are the end results that the application scenario is trying to achieve?
Support predictive maintenance by creating flexible maintenance plans. Data from PEIDs have to be analysed to determine the remaining lifetime of particular components in order to predict failures before they occur. Thus, a replacement can be performed at an economically optimal point of time. At the same time, the availability of trucks will be improved. A long term goal is the improvement of product design. There is no further details given on this, therefore the focus will be on the reduction of downtime using predictive maintenance.

In which form the knowledge to be generated from the field data in the application scenario should be provided?
The required knowledge is the predicted time of failure. However, this is rather uncertain and cannot be calculated exactly. Therefore it seems to be useful to classify the predicted remaining lifetime, e.g. more than 6 months - less than 6 months - less than 1 month - less than 1 week (or similar categories).

Determination of the relevant data

What are the different types of the gathered field data (to be gathered) needed to obtain the required knowledge? What systems for data retrival are in place? Do these systems provide data

in the correct context? Is the data/information to be used for generating the required knowledge located in only only one source?
The following data has to be gathered: - usage data (mileage, RPM, number of start-ups etc) - environment data (temperature, pressure) - maintenance data (number of component replacements) All the data will be provided from a single source: The on-board diary (on-board computer) which is integrated in the vehicle. Supposedly the context for data is provided (static product model could be stored on the on-board diary).

What are the data systems in which the field data is (should be) stored?
On the vehicle: the on-board diary. Summary statistics are to be sent to a ground station periodically and will most probably be stored in a backend database as well.

How the relevance of data is (should be) determined?
A prediction model will be needed to integrate the field data and estimate the remaining lifetime. This model will be based on physical properties of the material, e.g. percentage of degradation after a certain mileage etc. In general, the input data needed for the prediction model is the relevant data.

Access, retrieval of relevant data

What systems are (should be) used to access the relevant data?
• Backend systems at the ground station • Systems at an authorized garage

What systems are (should be) used for data retrieval? Do these systems provide data in the correct context?
Data retrieval will be performed by the on-board diary. It will then transmit the data to other systems using wireless connections. It should be possible to provide the correct context by the on-board diary.

Integration of data

In which forms the field data are available?
Data will be sampled by sensors and recorded by the on-board computer. This list shows the data that is planned for collection: 1. For predicting oil change interval, the following mission profile data should be gathered and used. – RPM – Fuel Cons – Trip duration – Number of engine start-up – Oil temperature – Engine load – Water temperature

- Engine working hours
  - Engine age
  - Boost Pressure

2. For predicting air filter change interval, the following mission profile data should be gathered and used.
   - Difference of air pressure before/after change
   - RPM
   - Temperature
3. For predicting brake change interval, the following mission profile data should be gathered and used.
   - Number of brake change
   - Energy consumption during brake action

How is the data integrated? Can uniformization be ensured?
The on-board diary collects and stores the information. As it also creates summary statistics from it, uniform data can be assumed.

Validation of data

What external conditions can influence the validity of data?
Failure of sensors, insufficient reliability of data transmission (rather unlikely though, as everything is wired and distances are short)

What are the subjective factors in the gathered field data (for example, human judgements)?
None

What factors can influence the accuracy of data (incompleteness, incorrectness, censoring, etc.)?
N/A

Data processing / analysis

What methods and tools are (should be) used for the analysis of data?
Prediction algorithm which in turn might be based time-series analysis, trend detection, and extrapolation.

Logical test

What cases correspond to the existence of a matter of concern?
There will be matter of concern if the predicted remaining lifetime of a component is shorter than the defined threshold (say, two weeks).

Information synthesis

What tools and methods are (should be) used to generate knowledge from the information obtained through data processing/analysis?
The information synthesis will be performed on the truck's on-board diary to display alerts on the dashboard in case of an upcoming failure. There is also information synthesis at the ground station to enable analysis on the whole truck fleet. It is not sure whether this analysis will be done within the PROMISE DSS.

WP A5 : PROMISE MOL information management for heavy vehicle estimation

Objectives of obtaining the required knowledge

What are the purposes for which the knowledge to be generated from the field data in the application scenario is used for? What are the end results that the application scenario is trying to achieve?
The knowledge to be generated is to be used at: • MOL: to responsiveness to the events (customer requirements in term of service and maintenance), • design (BOL): to improve some aspects of the design taking into account for example failure causes, • production (BOL): to improve logistics concerning supply of spare parts/components • EOL: to improve EOL processes taking into account for example recyclability, reusability, remanufacturability, disposability, etc.

In which form the knowledge to be generated from the field data in the application scenario should be provided?
The exact format in which the knowledge should be provided is not specified. However it's clearly indicated that the generated knowledge should support: • Fatigue life prediction, • predictive maintenance, • Design for X, • demand forecasting (for spare parts/components), • EOL issues.

Determination of the relevant data

What are the different types of the gathered field data (to be gathered) needed to obtain the required knowledge? What systems for data retrieval are in place? Do these systems provide data in the correct context? Is the data/information to be used for generating the required knowledge located in only one source?
The data to be gathered concern: • Component life, • Repair • Maintenance • Application severity • Etc.

What are the data systems in which the field data is (should be) stored?
The data to be captured from sensors should be stored in the "data acquisition system for sensors". From the description of the A5 PROMISE demonstrator, it follows that this software is not yet available.

How the relevance of data is (should be) determined?
Not specified in the description of the A5 PROMISE demonstrator but logically the relevance of

the data should be determined according to the purpose of the knowledge to be generated.

Access, retrieval of relevant data

What systems are (should be) used to access the relevant data?
Sensor software reader (wireless system).

What systems are (should be) used for data retrieval? Do these systems provide data in the correct context?
Sensor software reader (wireless system).

Integration of data

In which forms the field data are available?
There are various kinds of data to be considered: field data, engineering data and auxiliary information/data. The forms in which field data will be available are not specified in the description of the A5 PROMISE demonstrator.

How is the data integrated? Can uniformization be ensured?
The integration of data before its transformation into information is inevitable. However the way this integration will be realized is not specified in the description of the A5 PROMISE demonstrator.

Validation of data

What external conditions can influence the validity of data?
Not specified in description of the A5 PROMISE demonstrator.

What are the subjective factors in the gathered field data (for example, human judgements)?
Not specified in description of the A5 PROMISE demonstrator.

What factors can influence the accuracy of data (incompleteness, incorrectness, censoring, etc.)?
Not specified in description of the A5 PROMISE demonstrator.

Data processing / analysis

What methods and tools are (should be) used for the analysis of data?
The methods and tools to use in transforming data into information are not specified in description of the A5 PROMISE demonstrator. However, the output of this analysis “stress range” is indicated.

Logical test

What cases correspond to the existence of a matter of concern?
Not specified in description of the A5 PROMISE demonstrator.

Information synthesis

What tools and methods are (should be) used to generate knowledge from the information obtained through data processing/analysis?

The tools and methods to be used to generate knowledge from the information obtained through data processing/analysis are not specified in description of the A5 PROMISE demonstrator. However, it is indicated that the output should comprise “fatigue life calculation” for use in making decision about preventive maintenance.

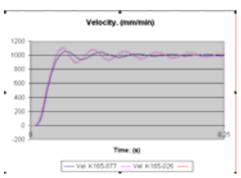
WP A6 : PROMISE MOL information management for predictive management for machine tools

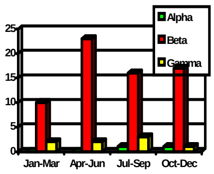
Objectives of obtaining the required knowledge

What are the purposes for which the knowledge to be generated from the field data in the application scenario is used for? What are the end results that the application scenario is trying to achieve?
The main Application Scenario purpose is traceability. It is important to know the history of the machine components: repairs, fault, substitutions, maintenance information, etc… Some eletrical or mechanical components once repaired could be installed on other machines, but the machine builder could ignore on which system that component was installed. At the next fault or reparation, it could be desiderable to know the ‘history’ of that element, it could be useful to perform statistical analysis on batch of components, etc…All this knowledge of working conditions is a useful feedback to the design department and technical assistence service. The final result is to obtain a overall knowledge of the parts (components) constituting a machine.

In which form the knowledge to be generated from the field data in the application scenario should be provided?

Knowledge should be provided in numerical format, strings and charts.
Should be visualized data like “Changed oil 23rd May 2005” or “ Batch Productionn°1000” and digital parameters (even in form of bar chart or trend lines) indicating the behaviour of components during the time.





E.G. It should be chosen a suitable output to outline this relation:
The last year at periodical check overshoot has increased its value of 10% (Field data) => Adjustment of Kv parameter (operation to be executed)

Determination of the relevant data

What are the different types of the gathered field data (to be gathered) needed to obtain the required knowledge? What systems for data retrival are in place? Do these systems provide data in the correct context? Is the data/information to be used for generating the required knowledge located in only one source?
Generic data: batch production, warranty time, fixings and substitutions, etc… Maintenance data: Test parameters like rising time, overshoot, error max, root mean square error, etc…

System in place for data retrival are the Computerized Numerical Control (CNC) and RFIDs. Data are located in these two sources.

What are the data systems in which the field data is (should be) stored?
It should be developed a backend system managing the data coming from the field.

How the relevance of data is (should be) determined?
Generic data are validated by technical personnel who is responsible of its input. Maintenance data are validated by computer algorithms that generate them.

Access, retrieval of relevant data

What systems are (should be) used to access the relevant data?
A backend computer should access (e.g. by a modem connection) to data stored on the local CNC.

What systems are (should be) used for data retrieval? Do these systems provide data in the correct context?
Relevant data stored on CNC are retrieved from RFIDs (installed on the machine and connected to the CNC) and are locally produced by the CNC (periodical tests).

Integration of data

In which forms the field data are available?
Field data is available as strings (eg. manually inputted by technical assistants) and numbers (eg. read from the sensors).

How is the data integrated? Can uniformization be ensured?
Data is integrated by using a coding with fixed length for each field. So whatever is the kind of data stored (string, numbers, etc) they are translated in a form that can be easily and univocally understood.

Validation of data

What external conditions can influence the validity of data?
Human error in input procedure of generic data. Malfunctions of CNC test procedures for maintenance data.

What are the subjective factors in the gathered field data (for example, human judgements)?
No subjective factor.

What factors can influence the accuracy of data (incompleteness, incorrectness, censoring, etc.)?
Incorrectness or incompleteness while manually entering the fields.

Data processing / analysis

What methods and tools are (should be) used for the analysis of data?
Maintenance data: Mathematical methods are used for the analysis of data. Current, position, velocity data are elaborated and transformed into information (relevant parameters).

Logical test

What cases correspond to the existence of a matter of concern?
Abnormal values for sensor data (current, velocity) usually correspond to degenerating components condition and can be used as indexes for the prediction of faults and better comprehension of technical problems when they occur on a machine

Information synthesis

What tools and methods are (should be) used to generate knowledge from the information obtained through data processing/analysis?
Should be used statistical analysis (e.g. fault percentage), analytical evaluation (e.g. out of thresholds) and advanced logical analysis (fuzzy logic, neural networks).

WP A7 : PROMISE MOL information management for EEE(1)

Objectives of obtaining the required knowledge

What are the purposes for which the knowledge to be generated from the field data in the application scenario is used for? What are the end results that the application scenario is trying to achieve?
Support maintenance and repair operations of gas boilers which are performed by the after sales service. The knowledge to be created is used for prediction of product failure. Thus, the availability of gas boilers can be improved.

In which form the knowledge to be generated from the field data in the application scenario should be provided?
The required knowledge is the predicted time of failure. However, this is rather uncertain and cannot be calculated exactly. Therefore it seems to be useful to classify the predicted remaining lifetime, e.g. more than 2 months - less than 2 months - less than 1 week (or similar categories). This knowledge will be combined with error codes and device information, and be sent to the decision support system. The DSS informs engineers about the current status of the boiler and makes recommendations for action.

Determination of the relevant data

What are the different types of the gathered field data (to be gathered) needed to obtain the required knowledge? What systems for data retrival are in place? Do these systems provide data in the correct context? Is the data/information to be used for generating the required knowledge located in only one source?
Data is gathered by embedded sensors and from the internal status of the gas boiler control board. - Internal status of product (switches, actuators etc) - Physical measures (various temperatures) - Usage data (RPM of fan, hot water flow rate, etc) For knowledge creation is necessary to combine the gathered data with product information from the PLM system.

What are the data systems in which the field data is (should be) stored?

Data is stored on the PEID. In case of events (failure, timer, other diagnostic events), data will be sent to a PLM system.

How the relevance of data is (should be) determined?
A prediction model will be needed to integrate the field data and estimate the remaining lifetime. This model will be based on physical properties of the components, e.g. percentage of degradation after a certain usage etc. In general, the input data needed for the prediction model is the relevant data.

Access, retrieval of relevant data

What systems are (should be) used to access the relevant data?
The data is only accessed by the PDKM. *Note: The application scenario suggests that the DSS does not access the data on the PEID directly. Instead, it requests a remaining lifetime prediction from the PDKM. It remains to be seen, whether this is consistent with the overall PROMISE architecture.*

What systems are (should be) used for data retrieval? Do these systems provide data in the correct context?
Data is being retrieved by the PEID. It can be assumed that it can provide the correct context as it is intended to analyze / pre-process the data locally.

Integration of data

In which forms the field data are available?
The data will be gathered from sensors in the gas boiler as well as from the gas boiler control board. It will then be stored on the PEID. PEID should gather the following data. − Temperatures − Switches status − Actuator status (both digital/analog) − Historical data − Parameters value − Boiler status (Central heating/Domestic hot water) − Command to modify parameters (e.g. Central heating temperature)

How is the data integrated? Can uniformization be ensured?
Uniform field data can be assumed due to local processing (see above). There might still be the need to integrate the field data with PDKM product information to create knowledge.

Validation of data

What external conditions can influence the validity of data?
N/A

What are the subjective factors in the gathered field data (for example, human judgements)?
None

What factors can influence the accuracy of data (incompleteness, incorrectness, censoring, etc.)?
N/A

Data processing / analysis

What methods and tools are (should be) used for the analysis of data?
Predictive maintenance algorithm based on the following techniques. - DOE (Design of Experience) technique - Statistics technique - Mathematical models - Test/Assess the reliability.

Logical test

What cases correspond to the existence of a matter of concern?
There will be matter of concern if the predicted remaining lifetime of a component is shorter than the defined threshold. The threshold will be set so that the engineers have enough time left to replace the component in questions before it fails.

Information synthesis

What tools and methods are (should be) used to generate knowledge from the information obtained through data processing/analysis?
Firstly, the field data coming from the PEID will be combined with PDKM product information to create knowledge, i.e. to determine the remaining lifetime of a component. This knowledge is used by the DSS to derive action recommendations for the service engineers.

WP A8 : PROMISE MOL information management for EEE(2)

Objectives of obtaining the required knowledge

What are the purposes for which the knowledge to be generated from the field data in the application scenario is used for? What are the end results that the application scenario is trying to achieve?
Identify technical problems with refrigerators in a timely manner to improve efficiency and reduce product failures. In particular, the following situations have to be diagnosed: - Compressor fails to start - Refrigerator Unplugged - Compressor On for too long - Defrost not Starting - Door left open

In which form the knowledge to be generated from the field data in the application scenario should be provided?
Diagnosis results on abnormal product behaviour (usage/consumption), efficiency level

Determination of the relevant data

What are the different types of the gathered field data (to be gathered) needed to obtain the required knowledge? What systems for data retrieval are in place? Do these systems provide data

in the correct context? Is the data/information to be used for generating the required knowledge located in only one source?
- Usage data (compressor running time, etc.) - environment (temperatures)

What are the data systems in which the field data is (should be) stored?
Internal flash memory of the appliance. However, the data will not be accessed there. Instead, the access is to be done using a proxy device which is connected to the refrigerator by Ultra Lowcost Powerline.

How the relevance of data is (should be) determined?
Prediction algorithms are required to analyse and interpret the field data in order to diagnose the situations mentioned above. The relevant data will depend on the required inputs of the algorithm.

Access, retrieval of relevant data

What systems are (should be) used to access the relevant data?
Only the proxy device enables access to collected field data.

What systems are (should be) used for data retrieval? Do these systems provide data in the correct context?
The Predictive maintenance software retrieves the data.

Integration of data

In which forms the field data are available?
N/A

How is the data integrated? Can uniformization be ensured?
There is no information given on that in the application scenarios. However, as there is a single point of data access (the proxy device), uniform data can be achieved easily.

Validation of data

What external conditions can influence the validity of data?
The proxy device communicates with the actual refrigerator using a new technology called Ultra Lowcost Powerline for data transmission over powerlines. It can not be said, how reliable this technology is, so there might be a chance of invalid data caused by losses of packets.

What are the subjective factors in the gathered field data (for example, human judgements)?
None

What factors can influence the accuracy of data (incompleteness, incorrectness, censoring, etc.)?
Malfunction of sensors on the appliance.

Data processing / analysis

What methods and tools are (should be) used for the analysis of data?
Predictive maintenance algorithm (see other predictive maintenance scenarios for details, it is very similar).

Logical test

What cases correspond to the existence of a matter of concern?
Low compressor efficiency. Cooling pressure outside of allowed range. Internal temperature deviation to large.

Information synthesis

What tools and methods are (should be) used to generate knowledge from the information obtained through data processing/analysis?
Predictive maintenance software accesses the data and applies prediction algorithms to the data set. It is unclear, whether this data interpretation is supported by a knowledge base (case history) as it is indicated in the application scenario description. The demonstrator description does not mention any knowledge base.

WP A9 : PROMISE MOL information management for Telecom equipment

Objectives of obtaining the required knowledge

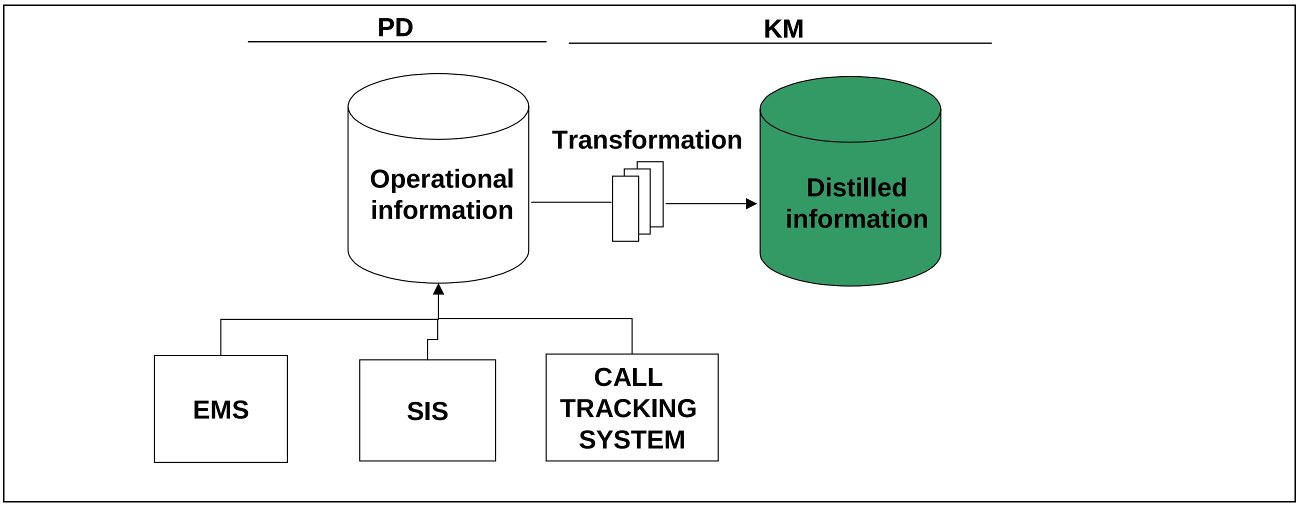
What are the purposes for which the knowledge to be generated from the field data in the application scenario is used for? What are the end results that the application scenario is trying to achieve?
The purpose of the scenario is to support the engineers and technicians to decision making about product improvements and problems solving/preventive maintenance. The collected information will be appropriately managed providing to the evolved actors the tools to correspond efficiently and effectively when a problem occurs and support decisions about product improvements. In more detail, the application scenario is trying to: • support the technicians to correspond effectively and in shorter time when a failure occurs • minimize design changes during product life and reduce design effort • improve the services provided to the customers

In which form the knowledge to be generated from the field data in the application scenario should be provided?
The data must be collected and integrated in order to be used for supporting the decision making for maintenance and product improvements, by providing problem identification, best solutions, repetitive problems. In more detail, a list of alarms, which will be processed, will give valuable input for preventive maintenance in combination with the data coming from the technicians. However, the format in which the knowledge should be provided is not specified, but the generated knowledge for example could be: • Indications/suggestions of similar problems • Indications of repetitive problems • Suggestions for solutions

Determination of the relevant data

What are the different types of the gathered field data (to be gathered) needed to obtain the required knowledge? What systems for data retrival are in place? Do these systems provide data in the correct context? Is the data/information to be used for generating the required knowledge located in only one source?
Data are gathered through automatic way, sensors, systems registrations and manually by the technicians. The different data that are needed to obtain the required knowledge are: • Alarms, IBAS keeps in the form of log files alarms to report on its performance, malfunction, and throughput degradation. The alarms are classified into Real, Active and Historical. Periodically, the IBAS alarms are reported to and processed by the Element Management System (EMS) that resides at the Network Operation Centre (NOC). • Info from the technicians about the product problems and maintenance (not always but in some cases) • Info coming from SIS (card's type, card's history etc.) • Info coming from Call Tracking System (type of problem, date, time and location when/why failure occurred, technician that handled the problem, solution). The systems that are used for data retrieval are EMS, Call tracking system and SIS. The above systems are independent which means that are not linked and information installed is not available to all involved actors in the whole maintenance and engineering team. The purpose is through the PROMISE solution these systems to be integrated.

What are the data systems in which the field data is (should be) stored?
Element Management system - EMS is used for the efficient operation management and supports the areas of: - Fault Management, - Configuration Management - Performance Management - Security Management Service Inter-registration System - SIS It is used by the service lab personnel to register problems, solutions and card history. It only contains information about cards that were restored in the lab. Call tracking system Tracks the calls made by customers, the problem described, the technician allocated and the solution. Product Data Knowledge Management – PDKM Will be used for gathering the operational data coming from the three previous systems by distilling and transforming these data to Knowledge supporting the decision making. Also, the info related to RFID tag (ID-it was decided that an ID will be used) will be stored to the PDKM (figure below). In addition, this PDKM should manage information about the instantiation of the product at each customer site and its history (Technical, Operational), other relevant data concerning the product like: - Documents (Design, Specification, Testing results…) - General Descriptions (Brochures, White Papers) - Engineering Change Orders and location data, as also data about customers and users.



How the relevance of data is (should be) determined?
The relevance of data could be determined according the symptoms of a problem, the frequency of the problem, the technician that made the maintenance, the location, the type of the component/material, the quality of data, the SLA-service level agreement etc. Using these indicators enable the DSS to determine the problem, the solution and best practices. To be clarified

Access, retrieval of relevant data

What systems are (should be) used to access the relevant data?
The data are accessed though the three current systems. EMS: The data concerning the operation of the IBAS can be accessed through the EMS and then can be transferred to the PDKM. These data can be used for the identification of the repetitive problems. Regarding the communication with the EMS, EMS implements a CORBA based North Bound Interface (NBI), so a NBI should be used to communicate with the EMS. For the communication between INTRACOM and the related NOC a VPN (Virtual Private Network) should be used. SIS: The data concerning the information about cards that were restored in the lab (card's history, problems and solutions) are accessed though the SIS. SIS uses statistic methods to process these data and reports are produced. Call Tracking System: The Call Tracking System includes info concerning the support that are provided to the customers, when customers experience problems that require software or hardware support, and call INTRACOM's hotline or send emails, fax. The system that INTRACOM developed and use to support and manage customer's requests is ADARES (ADvanced Action REquest System). ADARES is based on REMEDY platform. This system handles the requests in order to support customer providing solutions. In the future the actors will have the possibility to have access to all info regarding the product, solutions, best practices, repetitive problems etc. through the PDKM, where the info from the above systems will be integrated. Also, the customers will have access through the web to simple troubleshooting.

What systems are (should be) used for data retrieval? Do these systems provide data in the correct

context?
To be clarified

Integration of data

<table><tr><th>In which forms the field data are available?</th></tr><tr><td>EMS: alarms in the form of log files, info about configuration.
Call tracking: there is a system (ADARES) that is used to support and manage customer's requests. The system produces reports in the form of excels which information is related to:
ttype of problem,
date, time and location when/why failure occurred,
technician that handled the problem,
reports regarding the solutions given
materials that are consumed on the field
info about customers
SIS: info about cards that were restored in the lab. The registered info concerns problems, solutions and card's history.
Documents, brochures related to product
Manuals
Forms (Technical change order- ETA)
Descriptions (special characteristics, location etc)</td></tr></table>

How is the data integrated? Can uniformization be ensured?
The data are not integrated. An issue here is to define what uniformization means?

Validation of data

What external conditions can influence the validity of data?
Reasons that could be influence the validity of data could be failure of sensors, quality and accuracy of data registered by the technicians, as also the appropriate information that must be provided from the customer (card's replacements), which in most of the cases is not available to INTRACOM.

What are the subjective factors in the gathered field data (for example, human judgements)?
As only the technicians are responsible for the maintenance of the failures that occurs, the info concerning the problem and the solution must be recorded in such a way in order not to miss any important data. So, if the technician decides that some info are to detailed or forget to register some of the info then the gathered field data can be influenced.

What factors can influence the accuracy of data (incompleteness, incorrectness, censoring, etc.)?
Factors that can influence the data were described above.

Data processing / analysis

What methods and tools are (should be) used for the analysis of data?
EMS by an automatic way process the gathered data related to the product (IBAS). SIS uses statistic methods for data processing. Call tracking system uses the ADARES system to operate and process the gathered data, and according to the report that is needed to be produced from the database the appropriate fields are used.

In the future these systems will be integrated and appropriate algorithms, methods concerning knowledge management will be used for the translation and transformation of data to knowledge. DSS could support the decision making about problems and product improvements.

Logical test

What cases correspond to the existence of a matter of concern?
- Repetitive problems - alarms

Information synthesis

What tools and methods are (should be) used to generate knowledge from the information obtained through data processing/analysis?
The data coming from the existing systems (EMS, SIS, Call Tracking System) will be combined with PDKM product information to create knowledge. This knowledge is used by the DSS to derive action recommendations for the technicians and engineers.

WP A10 : PROMISE BOL information management for Design for X

Objectives of obtaining the required knowledge

What are the purposes for which the knowledge to be generated from the field data in the application scenario is used for? What are the end results that the application scenario is trying to achieve?
The DfX knowledge to be generated in this application scenario aims at supporting engineers in improving various aspects of the design such as reliability, availability, maintainability, lifecycle cost, environment and safety.

In which form the knowledge to be generated from the field data in the application scenario should be provided?
Concerning the representation of the DfX knowledge to be generated; it should be structured according to a predefined work breakdown structure (WBS). However, the form in which the knowledge should be provided is not specified.

Determination of the relevant data

What are the different types of the gathered field data (to be gathered) needed to obtain the required knowledge? What systems for data retrieval are in place? Do these systems provide data in the correct context? Is the data/information to be used for generating the required knowledge located in only one source?
1. The type of field data related to functions, systems and / or components to be considered are: - date, time and location when / where failure occurred - operating circumstances

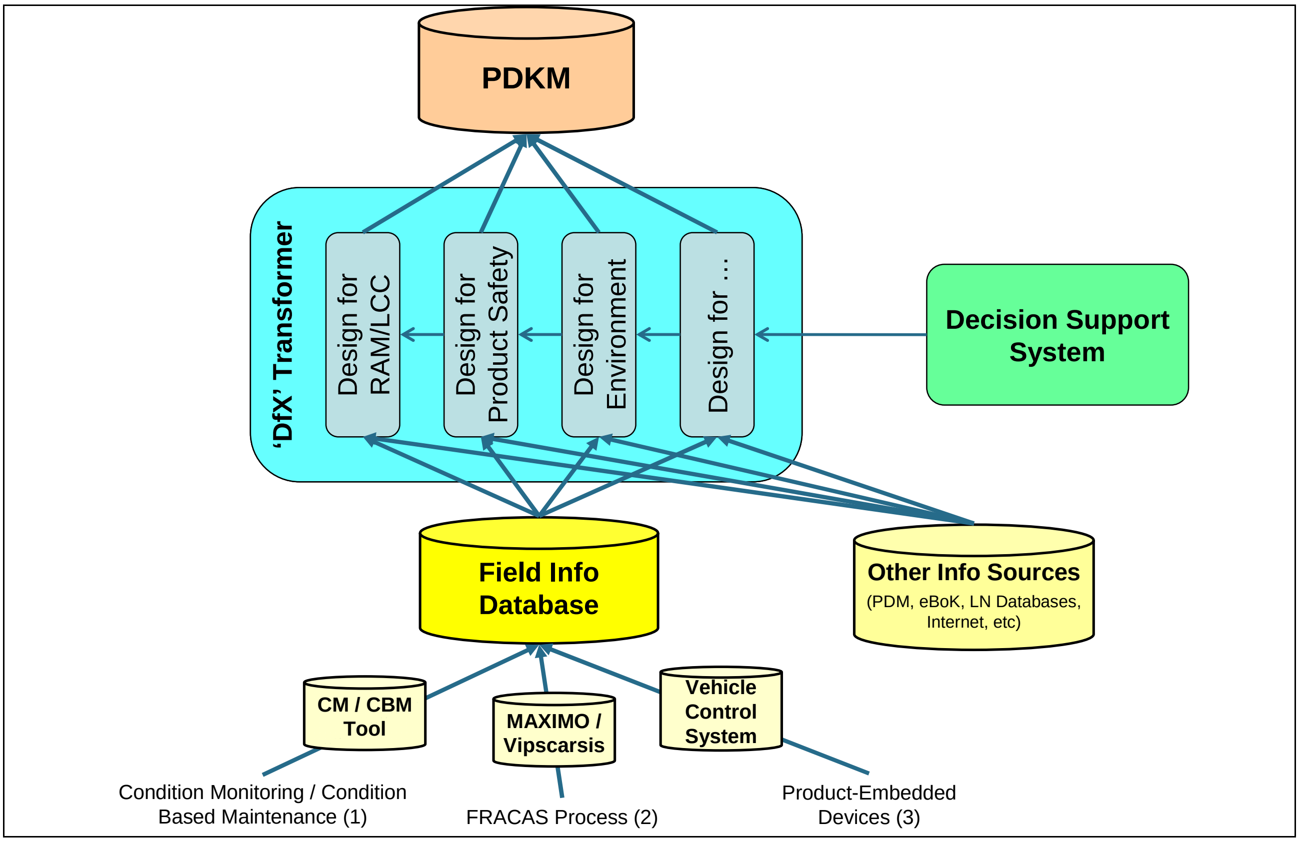
- environmental conditions - symptoms - effect on train service - operator's actions after failure - primary / secondary fault assignment - circumstances under which fault was first become apparent - operating distance, time, cycles, since it was put into service. 2. The program BTRAM (based on MAXIMO resp. VIPSCARSIS) is used to described the process of tracking the maintenance and repair history of vehicles and their registered component parts throughout the lifecycle. 3. Other relevant field data is gathered in event records [only for safety purposes] and inspection / maintenance reports 4. The data and information to be used for generating the DfX knowledge are located in different sources.

What are the data systems in which the field data is (should be) stored?
The field data to be used to generate DfX knowledge are available by CM (condition monitoring) – captured in the diagnosis system of the vehicle – data of component failures from FRACAS (Failure Reporting Analysis and Corrective Action System) – captured by VIPSCARSIS, MAXIMO) – by event recording and by inspections. The required field data shall be available in Field info database. All other data and information are available in other sources such as PDM system, Lotus Notes databases, EBoK's (Intranet), Internet and other similar data and information sources (DfX basic data, standards, etc.). A single system for all kinds of field data is not available currently at BT.

How the relevance of data is (should be) determined?
Not specified in the description of the A10 PROMISE demonstrator but logically the relevance of the data should be determined according to the purpose of the knowledge to be generated. The relevance of the data is determined by the responsible system engineer.

Access, retrieval of relevant data

What systems are (should be) used to access the relevant data?
The field info database shall gather all relevant field data and provide it then to the DfX transformer. All other data / information will be accessed directly by the DfX transformer in the corresponding system (see figure below).



What systems are (should be) used for data retrieval? Do these systems provide data in the correct context?
The vehicle control system TCMS (resp. diagnosis system) is used to retrieve and aggregate data from individual sensors that are placed upon individual components.

Integration of data

In which forms the field data are available?
Field data are mainly available in electronic format within the previously described systems. Few data is also available only on paper.

How is the data integrated? Can uniformization be ensured?
Currently there is no a single system able to mange all kinds of field data captured by CM (Condition Monitoring), FRACAS (Failure Reporting Analysis and Corrective Action System), event recording and inspection. The field info database should integrate all these kinds of field data provide them in form suitable for analysis activities.

Validation of data

What external conditions can influence the validity of data?
The validation of the data to be analyzed is mainly dependent on the data analysis method to be considered and depends mainly on technology, operating conditions, corrective and preventive maintenance.

What are the subjective factors in the gathered field data (for example, human judgements)?
Subjective factors – mainly definition and interpretation of contractual conditions – are considered as minimal.

What factors can influence the accuracy of data (incompleteness, incorrectness, censoring, etc.)?
A topic to be considered is mainly the incompleteness and incorrectness of failure reports which have not been filled out correctly.

Data processing / analysis

What methods and tools are (should be) used for the analysis of data?
The objective of data analysis/processing is to obtain accurate information about a specific aspect of the system under consideration. There exist various data analysis methods and tools and the selection of the appropriate one depends both on the objective of the analysis and the characteristics of the data to be analysed. For example, in the case of reliability domain it is widely recognized that the field reliability is inevitably different from the one predicted by the engineers at the design stage on the basis of simulations and laboratory tests where the real operational conditions cannot be perfectly reproduced. Therefore, the analysis of field reliability data to determine to what extent the field reliability is different from expected is crucial. Currently reliability is analysed by using FRACAS MRT – customized ACCESS based set of tools – creating reports using field data gathered in MAXIMO and Excel.

Logical test

What cases correspond to the existence of a matter of concern?
The objective is to detect whether there is a matter of concern or not. In this step, it has to be decided whether the data analysis results reveal an underlying design problem or not. For example, only if the reliability index of the observed component is abnormally low, then there is a matter of concern in which case investigations to know what causes it is necessary. This step can be followed by making decision about whether to go further with the process, to stop it or to pursue other objectives. If no matter of concern is detected after the processing/analysing of data, it may be preferable to redefine the objectives of the analysis, to consider additional data or simply to stop the process.

Information synthesis

What tools and methods are (should be) used to generate knowledge from the information obtained through data processing/analysis?
If a matter of concern is detected then this information should be worked out further by considering other data, information and previous knowledge in order to determine the nature and dimension of problem, what caused it and its potential consequences. An appropriate processing and synthesis of all these information can generate the required knowledge governed by the appropriate specialist engineer.

WP A11 : PROMISE BOL information management for Adaptive Production

Objectives of obtaining the required knowledge

What are the purposes for which the knowledge to be generated from the field data in the application scenario is used for? What are the end results that the application scenario is trying to achieve?
The purpose of this application scenario is to demonstrate how the PROMISE platform can be

used to improve the overall enterprise performance by adapting the production system to the large number of product and process modifications prompted by the availability of feedback information concerning the whole product life cycle. The demonstrator is a software for the support of the decision on the selection of the best change action to introduce in the production system related to the *cylinder head* of the FIAT multi jet diesel engine.
Thus the knowledge to be produced is generic in nature: it is not focused upon one specific area.
Two knowledge objectives:

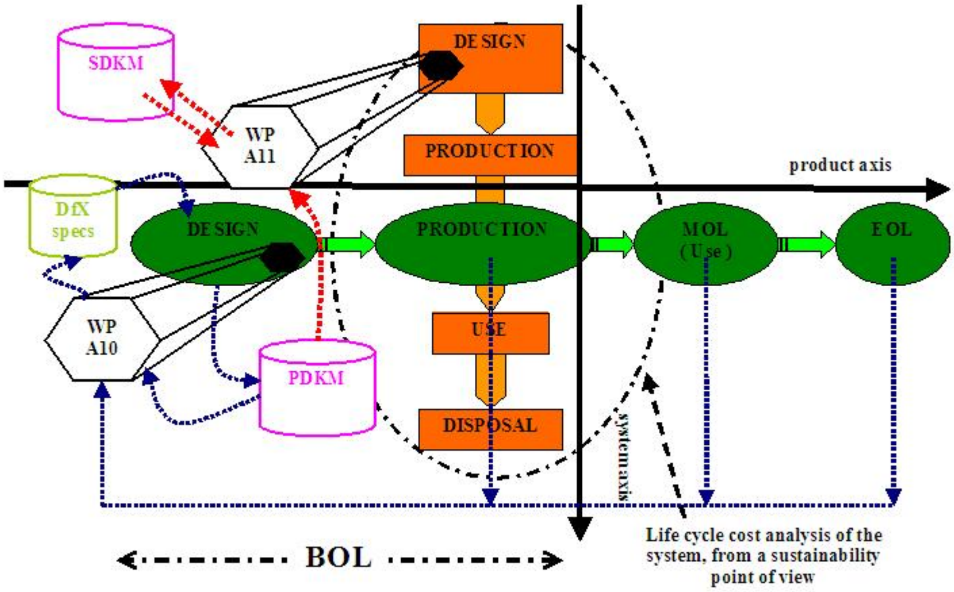
1. Production system reconfiguration— The optimal adaptation of the production system is important because it allows the continuous improvement of the product.
2. What if analysis— The What If analysis is important to quantify the impact of a potential modification of the product and/or the process. Indeed the product/process designer often considers a large variety of alternative modifications that are difficult to assess in terms of performances obtained at the factory shop floor level. For example it is not possible, without such an analysis, to properly assess the impact of all these alternatives in terms of system throughput, production cost, etc.

In which form the knowledge to be generated from the field data in the application scenario should be provided?
Adaptive Production mainly involves the BOL phase of a product, so that an integrated approach to the product & process system design can be carried out. The PROMISE platform gathers a great amount of data from the whole product lifecycle, data which are transformed into knowledge concerning the product. This can be used by the different product lifecycle stakeholders to improve one or more of the lifecycle phases and sub-phases. For instance they can be used in the BOL to modify features of the product as required by e.g. Predictive Maintenance or EOL processing. The great amount of collected data will increase product/process modifications relative to the current situation. Once the modifications have been decided, it is essential to make the system work according to the new "rules" in the most efficient way possible. To achieve this it is important to have a kit of tools which will help the decision maker to reconfigure the production system (e.g. the production line) or even to design/configure a new one. Knowledge must be delivered in the form of recommendations for Adaptive Production to take effect. For example, In more detail,. the possible adaptation actions include: • Introduction of new machines in the production line. • Introduction of new part transporters in the production line. • Introduction of additional WIP buffers in the production line. • Introduction of new work operators in the production line. • Modification of process parameters. • Modification of the number of fixtures flowing in the production line.

Determination of the relevant data

What are the different types of the gathered field data (to be gathered) needed to obtain the required knowledge? What systems for data retrival are in place? Do these systems provide data in the correct context? Is the data/information to be used for generating the required knowledge located in only one source?
The figure below outlines the application scenario. The aim is to reconfigure the system, given the modifications to the product/process. In the green blocks you can see the life cycle phases from the PROMISE point of view; Design and Production constitute together the BOL phase. In the orange blocks you can see the life cycle of the production system. An emphasis will be given to the sustainability of the system reconfiguration. With the word "system" is intended here the set of hardware and software resources whose aim is to realize the whole production process, (e.g.

production lines, FMSs, job shops). With PDKM (Product Data Knowledge Management) is intended the storage and management system of product data and knowledge, one of the essential elements of the PROMISE platform. With SDKM is intended the set of data (with the relative knowledge) concerning the system.
Whereas A10 is involved with analysing the product for DfX procedures, A11 is focused upon the process for similar procedures. Thus, the relevant field data comes from the four stages outlined in the production process: design of the process, production of the process, use of the process in product production, and disposal of the process to make way for a new process for product production.



What are the data systems in which the field data is (should be) stored?
The data will be held in the SDKM (set of data knowledge management—see above) system; as well as interaction with the to-be designed PDKM system.

How the relevance of data is (should be) determined?

There are two scenarios envisaged:

Scenario 1/A

The process/product modifications, or possibly the creation of a new product to be added to the production mix must be adapted to an already existing production system. The Adaptive Production paradigm forces the decision maker to decide how the system layout should be modified (e.g. "Should the number of machine tools be increased, decreased, or remain the same?" And, if it the case, "How many new machine tools should be added and what kind of machine tools?") or how the inter-operational buffers should be modified in order to maximize a certain type of objective function (e.g. the system throughput).

Scenario 1/B

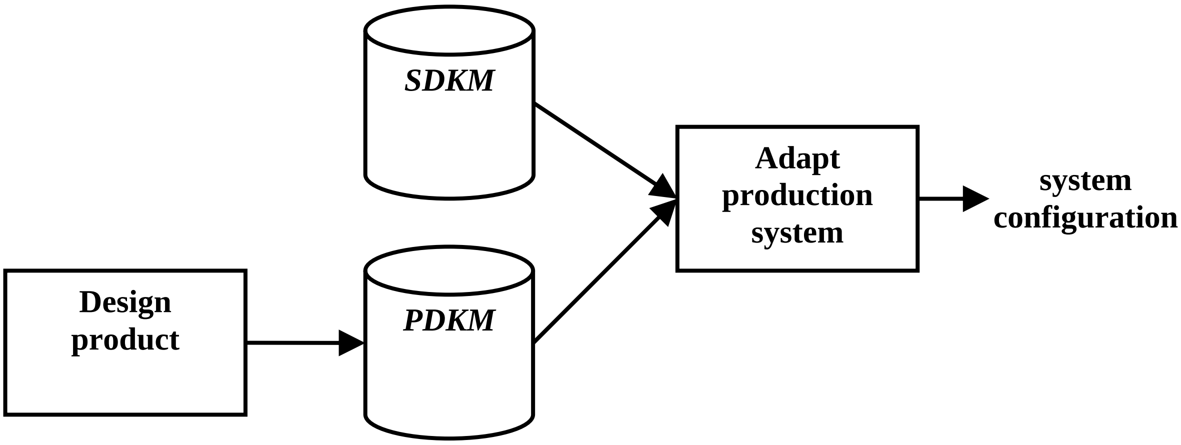
The process/product modifications affecting a family of products are identified and properly modelled, e.g. a set of scenarios for these type of modifications is given with the relative set of probability distributions. In this case the reconfiguration activities should take into consideration the whole set of modifications scenarios to make the same decisions as in Scenario 1/A, but with the possibility to make the decision process take into consideration all the implications which a certain feasible configuration can cause to the lifecycle of the system.

Five field data types are outlined:

1. To receive data concerning the product, with or without modification with respect to the present product.
2. To receive data concerning the process, with or without modification with respect to the present process.
3. To receive data concerning the production system, with or without modification with respect to the present system.
4. To optimise the configuration of the production system, adapting it to product and process modifications.
5. To evaluate the performance of the production system.

Access, retrieval of relevant data

What systems are (should be) used to access the relevant data?



This application scenario focuses on the production system, the product being any one of the products involved in the rest of PROMISE application scenarios. No physical component is needed for A11 Demonstrator, which is simply a software. This software will be able to simulate different configurations for the production system of anyone of the products considered in PROMISE. So we do not need any PEID to be attached to the machine tool because we do not need to follow the system lifecycle. Anyway here a simple model of the input and output data (see figure below) that can be found, in order to state from the very beginning all data and information involved in the scenario.

It is suggested that the relevant data in scenario 1/B may be accessed by the analysis and modelling of the future most probable scenarios for the product/process modifications. If such an

activity is properly carried out, the Scenario 1/B should be tackled using appropriate mathematical models (e.g. stochastic dynamic programming tools) and the results obtained should be of greater importance for the enterprise.

What systems are (should be) used for data retrieval? Do these systems provide data in the correct context?
"Not applicable" See answer to question above

Integration of data

In which forms the field data are available?
The ownership of the appropriate data about the future evolution of process plans, bills of materials and demand will enable the decision maker to face the Scenario 1/B. Otherwise only Scenario 1/A can really take place.

How is the data integrated? Can uniformization be ensured?
The backend software needed for the present application scenario relies on the presence of, at least, two software tools for data management, which correspond to the two big databases contained in the figure in 2.2. The PDKM software is the one defined and used inside the PROMISE platform, as been designed and implemented by the activities performed in WP R7. The SDKM is a simple list providing the needed data about the system. The data is integrated in the PDKM. Uniformization cannot be assumed for such a generic application scenario.

Validation of data

What external conditions can influence the validity of data?
Reconfiguration of production systems is nowadays carried out without taking into consideration the different sort of product/process modifications. This is because of the unavailability of the proper data needed for the analysis. The implementation of the Adaptive Production paradigm inside an already existing firm could have some negative aspects on business management, like: • The justification of costs due to either the first acquisition of the needed software or to the updating activities of the same software. • The difficulties inside the enterprise to implement the new system configuration due to the cost of the reconfiguration activities, though these system modifications are strongly based on some economical motivations, and more generally the hostilities to accept the output of the decision support system.

What are the subjective factors in the gathered field data (for example, human judgements)?
Not considered, but presumably large with such a generic application scenario.

What factors can influence the accuracy of data (incompleteness, incorrectness, censoring, etc.)?
Not considered; generic application scenario.

Data processing / analysis

What methods and tools are (should be) used for the analysis of data?
The people directly involved in the application scenario include: • *Production System designers*. Proofs the feasibility of new systems configurations by

using the demonstrator, and chooses the adaptations to be implemented into the present production system, i.e. the new system configuration; directly uses the DSS software. • *Product designers*. Evaluates the impact of the design modifications on the production system performance, both in terms of technical and economic performance, directly before the implementation of such modifications. • *Process designers*. Designs the production process, e.g. modifies the production process following what has been gathered from the field. Then they can make the same evaluations as the product designer, but concerning process modifications instead of product ones. • *Production planners.* Manages the production system, e.g. plans how to produce the products. Is involved in the adaptation of the production system to the new requirements because they are the one who can directly measure the implications of different potential system configurations on the production capacity offered by each configuration to the firm, thus allowing the production planner to evaluate different plans and modes of use directly comparing the system performance to the market demand. This is also enabled by the possibility of having this demonstrator and of using it in everyday practice.

Logical test

What cases correspond to the existence of a matter of concern?
The Logical test will be the analysis of a (new) configuration for the production system, in particular the one that best fits the objective chosen before the analysis. The output data will contain for example the number/type of machine tools, the number/type of inter-operational buffers (in particular their capacity), and so on; this must be tested for optimality.

Information synthesis

What tools and methods are (should be) used to generate knowledge from the information obtained through data processing/analysis?
Information synthesis will occur in the PROMISE Decision Support System, which is still to be defined in task TR9.1 from WPR9. In A11 the demonstrator provides robust support to the decision maker in deciding how to modify the production system layout, technology and equipment to satisfy the new product and/or process requirements.