UTILIZING UNSTRUCTURED FEEDBACK DATA FROM MRO REPORTS FOR THE CONTINUOUS IMPROVEMENT OF STANDARD PRODUCTS

Abramovici, Michael; Gebus, Philip; Göbel, Jens Christian; Savarino, Philipp
Ruhr University Bochum, Germany

Abstract

Industrial product-service systems (IPS²) enable new opportunities for the IPS² provider as the development department gets access to valuable feedback information from the use phase in order to improve follower product generations or currently active product instances. Yet, systematic analyses of this data during product improvement processes at most consider key figures from structured sources or subjective feedback information. The paper proposes a feedback design assistant that allows product developers to derive requirements from knowledge-based analyses of a product's unstructured and structured feedback information, whereas the knowledge extraction from unstructured data is emphasized.

Keywords: Design informatics, Knowledge management, Product-Service Systems (PSS), Product Lifecycle Management (PLM), Feedback management

Contact:
Philip Gebus
Ruhr University Bochum
Chair of IT in Mechanical Engineering
Germany
philip.gebus@itm.rub.de

Please cite this paper as:
Surnames, Initials: *Title of paper*. In: Proceedings of the 21st International Conference on Engineering Design (ICED17), Vol. 6: Design Information and Knowledge, Vancouver, Canada, 21.-25.08.2017.

1 INTRODUCTION

Industrial product manufacturers are increasingly aiming to change their sales strategies from just selling the ownership of their product to availability-driven and service-oriented business models, such as selling the usage of their product and assuring a contracted availability (Meier et al., 2010). On the one hand, this enables companies to increase the revenue during the use phase of their product's lifecycle. On the other hand, they get access to valuable feedback information in order to improve follower product generations (Roy et al., 2016). In such new business models in the context of so called industrial product-service systems (IPS²), the provider is obligated to guarantee his IPS²'s contracted availability. Therefore, providers must interact with several service partners, suppliers and customers (Meier, 2012). In particular, the provider relies on different service partners that execute maintenance, repair and overhaul (MRO) activities during the use phase of an IPS²'s lifecycle. In order to do so, service partners might require information from other parties in the IPS² network. Therefore, MRO reports need to be shared across the different parties in the network. This enables entirely novel opportunities for the IPS² provider's development department, such as exploiting feedback information about failure-prone components documented by customers and service partners in maintenance, repair and overhaul (MRO) reports to improve future product generations, which have not been accessible before.
Most of these industrial products are modular, mass-customizable solutions based on standard components (e.g. belts, spindles or pumps), which underlie periodical design improvements (Abramovici and Lindner, 2011). Design processes that seek to improve future product generations at most include the analysis of structured product usage data (feedback information), such as sensor data or metadata (Dienst et al., 2012). In reliability engineering (Design for Reliability) as a part of the Design for X paradigm for example, several types of usage data, such as e.g. time, cycles or on/off switches to failure (O'Connor and Kleyner, 2012) are considered to determine a component's or product's reliability. In addition to structured sources, if at all isolated subjective feedback information from revocations or customer complaints is considered in these approaches (Abramovici and Lindner, 2011). The subjective characteristics of these information make it difficult to derive universally necessary improvements. A recent expert study underlines the need for intensified usage data consideration during product development (Abramovici and Herzog, 2016).
Nevertheless, unstructured (textual) usage data is barely considered, although it contains valuable information, e.g. causes for a machine malfunctioning in MRO reports. Furthermore, over 95% of big data is unstructured (Gandomi and Haider, 2015). Regarding machine malfunctions, a lot of causes for these malfunctions cannot be captured by sensors but are hidden in repair reports that have been collected during the last decades, which might not even have a sufficient amount of sensors installed. Thus, making this information available to the product development in addition to structured data opens up a whole new potential to improve future product generations (Lindner, 2015).
This paper provides a *feedback management concept* to support product developers in improving follower product generations or current product instances of industrial mechatronic products by exploiting aggregated, objective feedback information from *unstructured usage data* combined with structured feedback information. The benefits of the concept in the context of product improvement processes are demonstrated by introducing and validating a product use information feedback design assistant.

2 REQUIREMENTS

The analysis of several industrial use cases and more than 100 real-life MRO reports in this context has led to the overriding requirements presented in this section, which the concept needs to fulfil to assist product developers during product improvement processes.
The use cases consider the rotary spindle unit of a Wire Electrical Discharge Grinding (WEDG) machine. The provider of the WEDG (IPS²) solely considers subjective feedback information from revocations, complaints or warranty cases to identify failure-prone components in the current generation as input for product improvement processes combined with objective feedback information from condition monitoring of active product instances. As mentioned before, subjective feedback information is only of limited suitability, while considering only strictly structured data may lead to inefficient or wrong product or component improvements. Exemplary, after analysing structured sensor data (e.g. spindle rotation speed, spindle running time, ambient temperature of the spindle unit and last

maintenance date of the unit), the drive belt has been identified as the most critical part of the WEDG as it cracks most frequently. Consequently, the product developer will derive new requirements that aim at strengthening the belt. However, the main cause for the crack of the belt is hidden in unstructured MRO reports locally created by service technicians that noticed steel cuttings that accrued during the milling process in the chassis of the driving belt. Thus, deriving requirements for improving the chassis would be more efficient as the cuttings cause micro cuts in the belt, which ultimately lead to the belt crack (see Figure 1). Derived requirements can either be used to improve follower product generations or to optimise currently active product instances. As in the example mentioned, unnecessary, inefficient and not beneficial product improvements might result from not including unstructured feedback information. In order to avoid these, the concept must meet the following superordinate groups of requirements:

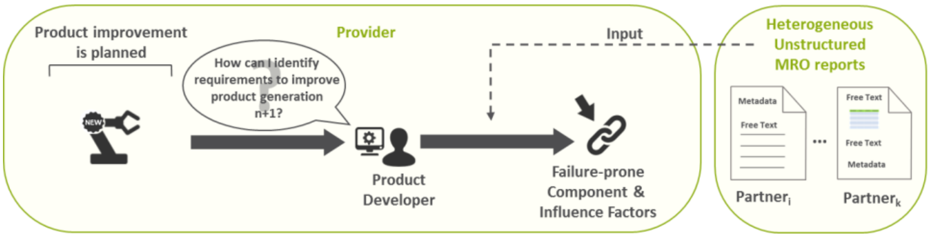

Figure 1. Problem Description

Objective Feedback Information Must Be Available to the Provider (R2.1)

To allow targeted and beneficial product or component improvements, not solely subjective feedback information must be available to the IPS² provider. Additionally, objective feedback information from structured sources (sensor data) combined with unstructured sources (failure descriptions, causes and solutions in MRO reports) must be accessible by product developers in order to enable in-depth analysis of failure-prone components. Furthermore, the management of product instance-related and product type-related data is mandatory as well as the mapping and relation between one another.

Unstructured and Heterogeneous Feedback Information Must Be Considered (R2.2)

According to estimates, not only 95% of corporate data is unstructured, but 80% of corporate knowledge is stored in textual form (Laudon et al., 2010) and so are MRO reports. To exploit this valuable knowledge during product development, heterogeneous unstructured data must be considered as each partner in IPS² networks records his MRO reports in an individual way. This unstructured feedback information has to be transformed into a structured form in order to enable automatic analyses.

Dynamic Changes Within the IPS² Partner Network Must Be Regarded (R2.3)

Partners, customers and products in IPS² networks change continuously. The latter for example might get reconfigured during use phase. These characteristic dynamics must be an integrative component of the concept in order to be applicable in IPS² scenarios. Therefore, the underlying data model of the concept needs to adapt to these dynamic changes and allow flexible extensions. As repair reports shall be focussed, malfunction causes, descriptions and respective solutions must be considered as well as product type and instance data including their relationships one to another.

Aggregated and Accurate Knowledge Must Be Presented to Product Developers (R2.4)

As manual analysis of the vast amount of feedback information is not reasonable, it must be pre-processed in such a way that only important information is displayed to the developer in an aggregated manner. Added value emerges if the accuracy of the knowledge extraction and mapping from textual unstructured documents reaches a certain level. As dealing with highly heterogeneous unstructured textual data accompanies a certain degree of inaccuracy, a threshold of 90% accuracy must be achieved in both: extraction and mapping.

3 RELATED WORK

Several research approaches aim at improving engineering processes and methods by considering feedback data from a product's use phase. The majority of these research activities consider only structured feedback data sources, while only few regard unstructured data. In this section, the most relevant works in this research area related to the scope of this paper are presented and distinguished from the presented concept:

Approaches Considering Structured Feedback Information

The framework by (Igba et al., 2015) feeds back in-service data to the product development phase. The authors propose ways of how to enhance the information capturing process from in-service reports in the future (basically by structuring the capture process through e.g. dropdown boxes). Nevertheless, the framework does not focus on how to analyse the large quantities of already existing unstructured textual in-service reports.

The product use data feedback design assistant by (Abramovici and Lindner, 2011) is the basis of the presented approach. The design assistant utilizes Bayesian networks to allow product developers to find critical product components. It enables what-if analyses on possible influence factors discovered in strictly structured (sensor) data in regard to a certain failure event.

The methodology presented in (Abramovici and Lindner, 2013) bases on object-oriented Bayesian networks and aims to support product developers in improving already existing mass-produced standard products. The methodology uses solely structured sensor data. The authors state that future works should imply different data sources as well as unstructured data.

The feedback assistance system by (Dienst, 2014) uses structured data from the use phase and stores them in a data warehouse to allow product developers to improve future product generations by using knowledge-based methods to analyse the usage data. The methodology does not regard unstructured data at all. Including unstructured data from the use phase requires adapted text mining methods and shall be the focus of future works.

The methodology presented in (Neubach, 2010) allows product developers to find weak spots in products by using what-if analyses on structured feedback data. The methodology also extends Product Lifecycle Management (PLM) approaches by providing a metadata model that allows product instance data management. Unstructured data is not considered in this approach.

Approaches Considering Unstructured Feedback Information

The framework for improving industrial goods by contextual knowledge provision (Dienst et al., 2012) considers customer and service information in product development. Engineers can access the information within their usual working environment, e.g. CAD tools. The approach bases on simple keyword detection, which is why it does not deliver a sufficient knowledge extraction precision. Additionally, the abstract nature of the approach compounds an industrial application.

The concept for sustainable IPS² presented in (Mamrot et al., 2016) bases on systems engineering and regards field data. The concept includes (unspecified) data and text mining methods to analyse field data and derive measures for weak spots of the IPS² within the product development phase. The approach considers service reports but does not mention concrete methods to be used or how exactly they should be combined to extract required information. Problems, such as heterogenic wordings or synonyms, are not addressed.

Resulting Need for a Methodology to Integrate Unstructured Objective Feedback Information

The analysis of the approaches in the scope of this paper shows that there is a need for a methodology to integrate unstructured objective feedback data from the product use phase into the product development phase as a concrete methodology remains still to be developed.

4 CONCEPT FOR CONSIDERING UNSTRUCTURED FEEDBACK DATA IN PRODUCT DEVELOPMENT

In order to fulfil the overriding requirements from Section 2, which have been identified in the early project phase, a novel concept for unstructured feedback data consideration in product development has been developed (see Figure 2).

In a first step, a product lifecycle management (PLM) approach has been extended to the effect that not only product type data but also product instance data, e.g. condition monitoring data, is considered and mapped to the respective product components (R2.1). PLM is a concept to manage engineering data and documents as well as processes across the whole product lifecycle. At the same time a Computerized Maintenance Management System (CMMS) has been defined as the basis for digital MRO task documentation (R2.2).

Secondly, in order to allow an automatic analysis of the unstructured feedback data, a methodology for knowledge discovery in unstructured MRO information has been developed (R2.2). The methodology is used to analyse, aggregate and evaluate heterogeneous unstructured MRO reports that are shared by service partners and customers (depending on the business model) in the IPS² network. It automatically extracts the malfunction description(s), cause(s) and solution(s) as well as metadata from the MRO

reports and transforms them into a structured form. To simplify the scenario, only repair reports are considered. Subsequently, this extracted aggregated information is mapped and conjunct to several core data from the IPS² provider's Product Lifecycle Management (PLM) approach.
In order to be usable in statistical knowledge-based analyses during product improvement processes, the conjunct information needs to be clustered and mapped to similar failure causes that refer to the same product components. Thus, a semantic knowledge base had to be developed that represents the clustered data and builds the foundation for the knowledge-based analysis (R2.3). Product developers can perform what-if analyses utilizing our product use information feedback design assistant from previous works (Abramovici and Lindner, 2011) to identify weak spots in the current product generation (R2.4). The resulting knowledge can be used to reconfigure current product instances to ensure their availability and to derive requirements for follower product generations improvements. While the overall concept addresses different scientific issues, this paper lays emphasis on the developed methodology to analyse, aggregate and evaluate unstructured data as well as the dynamic data model used to represent the extracted knowledge.

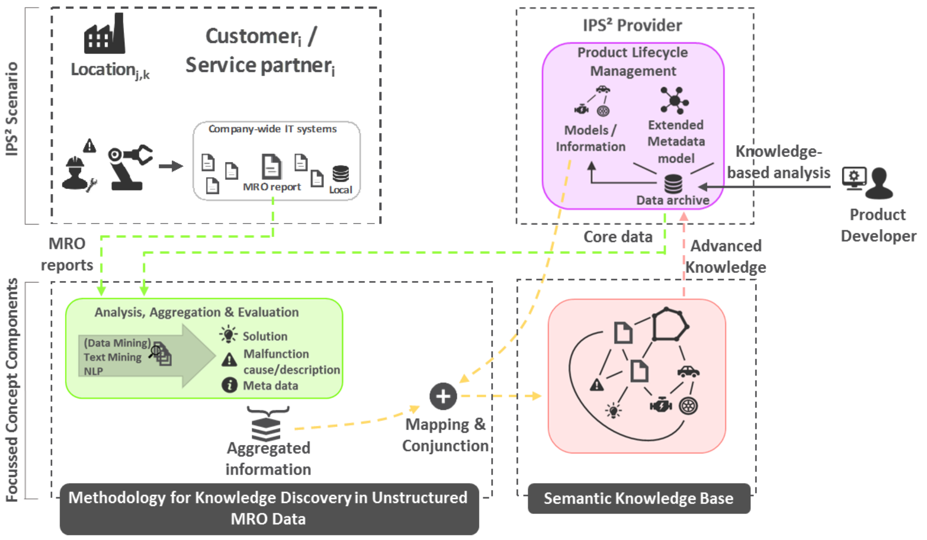

Figure 2. Simplified concept overview and focus of this paper

4.1 Methodology for Knowledge Discovery in Unstructured MRO Data

To make unstructured information from the use phase beneficially available in product improvement processes, only task-relevant aggregated information should be displayed to the product developer as manual analysis of the vast amount of data is not reasonable. Therefore, relevant information contained in unstructured MRO reports must be extracted and mapped one to another. The methodological approach presented in this section (see Figure 3) addresses these issues.
After documenting a technical repair activity, the new report about e.g. a plier that does not grab industrial steel wheels anymore is sent and registered to the processing component before it is stored into the knowledge base. In a first step, the available structured metadata of the report is extracted, while the unstructured textual part is divided into sentences as each sentence is processed separately.
Depending on the underlying data structure of the CMMS, structured metadata can be mined by the use of regular expressions (e.g. to identify repair and machine downtimes), specific keywords (e.g. to determine the failure category, such as mechanical, electronic, hydraulic or operator error) or directly by considering and exploiting the underlying CMMS data structure. Regarding the processing of the unstructured components, to address any spelling mistakes done by the service technician, each word is checked and corrected if necessary by calculating the edit distance between the word and similar words in a dictionary, which includes words from thesauri and previously processed textual documents. In a next step, algorithms analyse the sentence semantically and syntactically. Exploiting supervised machine learning models, the approach computes each words part-of-speech (POS) tag, e.g. "wheel" is a noun, and parses it syntactically by identifying "plier" as the subject and "wheel" as direct object for example. Afterwards, utilizing a trained machine learning model from the natural language processing research field, the approach computes each word's lemma (base form) based on its POS tag. Exemplary, the sentence "The pliers are not grabbing the wheels" is transformed to "the plier be not grab the wheel",

which will be important for mapping similar failure causes consecutively. At the same time named entities are identified by considering all nouns of the sentence and utilizing the thesaurus of the provider's Product Lifecycle Management approach, which supports the process by providing possible synonyms and designations of the physical components. As the IPS² that is affected by the malfunction is namely stated in the repair report and the provider takes track of product instances within his Product Lifecycle Management approach, the pliers can be identified as a component of the specific IPS². Thus, no additional manual work is required and the identification of the components is independent of the concrete product instance as the required information for this process is derived from the provider's available PLM approach. Additionally, the tense of the sentence is identified (past or present tense) by the analysis of the part-of-speech tags because it is required in a posterior step. The "Stuttgart-Tübingen Tagset" (STTS), which is the standard POS tag set for German, defines several different POS tags for the past tense of verbs, such as "VVPP", "VAPP" or "VMPP". As each word's POS tag is computed (as described above), the tense of a sentence can easily be identified by checking the occurrence of verbs with the previously mentioned tags.

Complete repair reports consist of three main unstructured components: failure description, failure cause and the solution. In order to allow a statistical analysis during product improvement, these components need to be identified and afterwards mapped to alike sentences in the knowledge base. Therefore, the sentence has to be transformed into its vector representation. The vector contains each word lemma only once (in the following named as term) represented by its normalized frequency ("TF") in the sentence multiplied by its inverse document frequency ("IDF"). For classification, an individual machine learning model has been developed. More than 100 real life repair reports have been analysed and used to train the (supervised) model. The model uses the word lemma vector and the identified tense as features (input values) to classify the sentence into failure description, failure cause, solution or irrelevant. Sentences that have been classified as the latter are not further regarded.

Many repair reports are incomplete and do not explicitly include the failure cause. In this case, the cause is implicitly mentioned in the solution (troubleshoot activity), which documents the actual repair task. For statistical use during product improvement, the cause must be explicitly documented. Therefore, it has to be derived from the solution. This can be achieved by using a thesaurus, which includes antonyms of often used verbs and adjectives in the MRO sector. The MRO report analyses yielded that the actual repair activities (cleanse, replace or repair) are mainly described by repetitive vocabulary, which is why the size of the developed thesaurus is kept within reasonable limits. Considering the exemplary documented solution "The reflector has been cleansed", one can derive the explicit cause "The reflector is dusty" by using the knowledge computed before (during the dependency parsing) as the reflector is semantically related to the verb "cleansed".

Mapping alike failure causes, descriptions and solutions one to another is crucial for beneficial use during product improvement as one and the same issue can be expressed in several different ways. But for statistical usage, similar cases must be treated as the same case. Therefore, a method had to be developed to achieve this task. The basic idea is to compute the cosine similarity between word vectors of the same category and referring to the same product type. Unfortunately, the TF-IDF representation mentioned before delivers too many false positives as the approach down weights essential keywords, such as the component name (e.g. the pliers) and verbs (e.g. cleansed, repaired), because they occur often since repair reports mainly briefly summarize the most important task-relevant information. Therefore, only the normalized term frequency in the sentence itself is used as weight representation. Additionally, unnecessary, often occurring terms are filtered out using a stop words list, such as "be", "have", "the" or "proper" for example. This leads to word vectors that mainly contain important information, e.g. the components, key verbs and adjectives. Additionally, a thesaurus adds synonyms of these terms to the vector. The thesaurus from the provider's PLM approach adds synonyms of the components (e.g. "mirror" as synonym of "reflector"), while an extended general thesaurus adds synonyms of often occurring, maintenance-related verbs and adjectives that describe the damaged state of the component (e.g. "broken", "defective", "knackered" as synonyms of "damaged"). As the failure causes describe a (standard) component's damaged state in a neutral way (independently of the IPS²), meaning e.g. in the simplest scenario some part "is broken" or "is dirty", the underlying thesaurus is also independent of the target machine and thus, can be transferred to arbitrary production IPS² containing standard components. Still, not all heterogeneous wording can be mapped one to another because a 100% accuracy is unrealistic due to the highly heterogeneous nature of unstructured textual data. Nevertheless, the data model presented in Section 4.2 helps to further improve the mapping accuracy.

Even if the cosine similarity between two failure causes does not surpass the predefined threshold of e.g. 90%, one can derive a similarity between the causes if the solutions are identical and refer to the same product components. The following example shall clarify the idea: Consider two malfunction causes "reflector is covered with dirt" and "laser measurement unit does not work properly". The cosine similarity between these causes does not surpass 90% due to different wording and detail used. However, both have similar solutions, e.g. stating somehow that the reflector has been cleansed. As the solution's cosine similarity surpasses the threshold of 90% and refers to the exact same product component (reflector), a similarity between the failure causes can be derived and persisted in the data model.

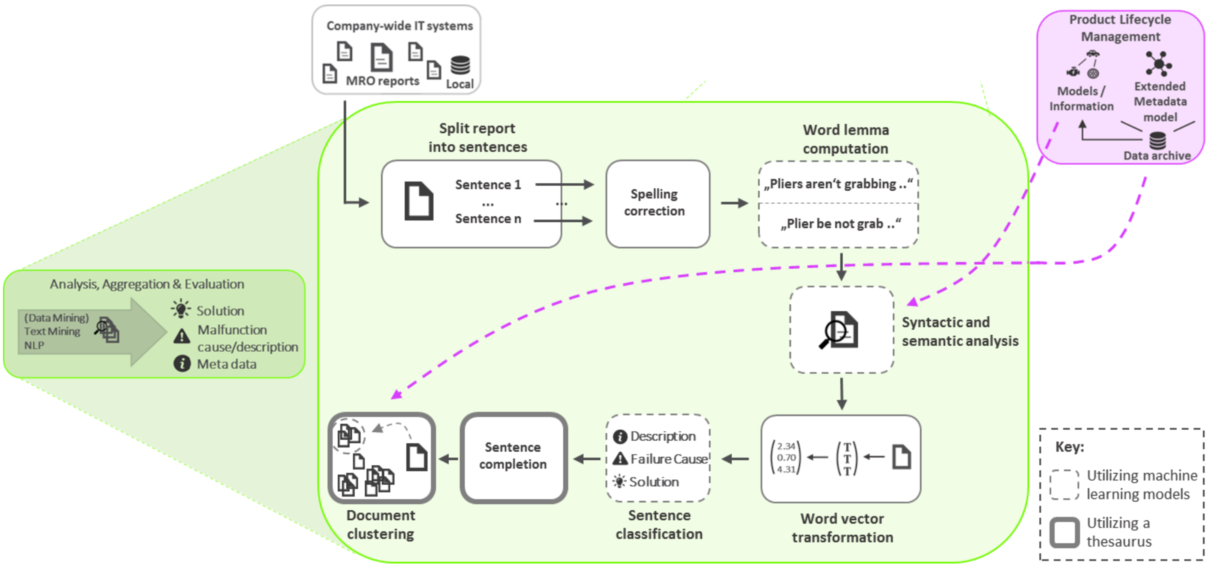

Figure 3. Methodology to process unstructured MRO data

4.2 Dynamic Data Model as Core of the Semantic Knowledge Base

To represent the knowledge discovered from unstructured MRO reports, a dynamic and semantically rich data model had to be developed that provides the basis for the knowledge-based analyses. An excerpt of the developed data model, its behaviour and the resulting aggregated failure cause clusters required for knowledge-based analyses are shown in Figure 4.

(a) The core part of the data model represents a subset of information from the IPS² providers Product Lifecycle management approach conjunct with extracted knowledge from unstructured repair reports. The data model maps required information from PLM, such as the product structure, CAD models, variants and component information into a semantic graph structure as representing and analysing this highly connected information with traditional data models becomes a complex, confusing and inefficient task. The data model bases on the property graph model that stores objects as vertices with attached properties. The relationships between two objects are modelled by directed edges following the naming convention of the Resource Description Framework (RDF), as in each relationship consists of a subject, a predicate and an object. As the model is individually adaptable and regards many different aspects, a simplified excerpt shows its main functionality. A new repair report has been processed stating that the reflector of the pliers is dirty (failure cause) and has been cleansed (solution) using the methodology described in Section 4.1. Extracted knowledge from the repair report is connected semantically by a directed ":REFERS_TO" edge to the identified product components (pliers and reflector). Failure description(s) are linked to cause(s) via directed ":CAUSE_FOR" edges, while the latter are connected to respective solutions or actual repair activities by ":SOLUTION_TO" relations. Of course, terms are stored in the model including their oath form (word lemma) and the term frequency value as an attribute of the relationship between the term and the actual sentence it occurs in. The solution for example would be linked to the terms "reflector" and "cleansed" with both having a normalized term frequency of 0.5. This is required for similarity computations as described in section 4.1. Furthermore, product type data and instance data is separately considered. Although concrete failure descriptions, causes and solutions are linked to product instance components, the latter are of course connected to the product component types via ":INSTANCE_OF" relationships and thus, can be considered during the product improvement process.

(b) Similarities between alike sentence types are represented by ":SIMILAR_TO" edges. If a similarity between two solutions has been identified, the data model automatically derives a similarity between the failure causes and links the causes with the respective edge.
(c) As a result, all failure descriptions, causes and solutions connected by a ":SIMILAR_TO" edge shape clusters. Regarding failure causes, each cluster represents exactly one cause of failure type. Therefore, these clusters can now be combined with the structured data from our previous works to perform what-if analyses utilizing the feedback design assistant.

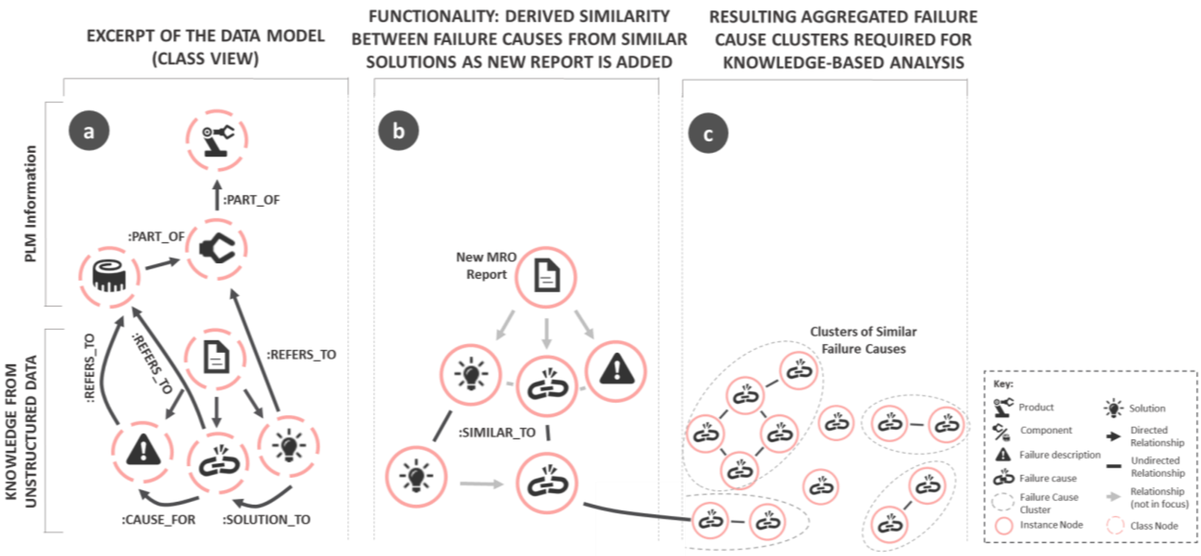

Figure 4. (a) Semantic knowledge base overview (b) functionality upon extension (c) cluster representation

5 IMPLEMENTATION AND VALIDATION

5.1 Implementation

The presented concept has been implemented as an extension of the prototypes presented in (Abramovici et al., 2016) and (Abramovici and Lindner, 2011). It is implement as a JAVA-based web service architecture. Each component of the methodology (see Section 4.1) is accessible by its own service allowing flexible extensibility and inclusion. The prototype pulls MRO reports from Computerized Maintenance Management Systems (CMMS). Product type and instance information required to build the data model is synchronised with PTC Windchill. The machine learning models utilize the Waikato Environment for Knowledge Analysis (WEKA). The data model (see Section 4.2) has been implemented as a graph in Neo4J. The thesaurus bases on GermaNet. As maintenance-related synonyms and antonyms are not covered well in GermaNet, an additional thesaurus has been developed to address these needs and to simulate an IPS² network-specific thesaurus.

5.2 Validation of the Methodology for Knowledge Discovery in Unstructured MRO Data

The two most critical parts of the presented methodology are the sentence classification model that is used to classify each sentence into failure description, cause, solution or irrelevant and the mapping of alike sentence types.
In order to validate the classification model, several different machine learning models have been tested using different sets of features and algorithms. More than 100 repair reports have been analysed and used to train a supervised model. A common measure for testing the accuracy of machine learning models is the F-measure. The model described in Section 4.1 utilizes support vector machines and achieves the highest tested accuracy with a F-measure of 96.4% in a 10 folds cross-validation test, which to put it simply means that it classifies 96.4% of the sentences correctly. Thus, the classification methodology surpasses the threshold from requirement 2.2 and therefore is suitable for the concept.
The mapping of failure causes has been manually tested by examining all clusters of failure causes. In total, 4 out of 114 have not been assigned to alike causes if applicable. A closer look revealed that all

four were not mapped correctly because the thesaurus has not considered the respective synonym used in these causes. Nevertheless, the predefined threshold of 90% has been surpassed. As both critical components have been successfully validated, the methodology meets the requirements and therefore is suitable to be used in the overall concept to improve standard products.

5.3 Validation of the Overall Concept to Improve Future Product Generations by Usage Data Consideration

The product use information feedback design assistant has been validated before considering solely structured feedback data (Abramovici and Lindner, 2011). It remains to validate the additional benefits of regarding the knowledge hidden in unstructured data. Therefore, in the early project phase, it has been determined to validate the approach by extending the previous use case by unstructured product usage data as described in section 2. As before, the rotary spindle unit of a Wire Electrical Discharge Grinding machine is considered. The drive belt has been identified as the most critical part. Product developers can identify the influence factors and their impact on the crack of the belt by utilizing knowledge discovery algorithms in product use information from the previous product generation provided by the feedback design assistant prototype. In addition to the identified influence factors determined from analysing structured data (spindle rotation speed, spindle running time, ambient temperature of the spindle unit and last maintenance date of the unit), novel information from unstructured data now can be considered that could not be captured by sensors. In the majority of cases, cuttings from the metal wheel milling process have been found in the housing of the driving belt (see Figure 5). Due to the rotation and vibration, these cuttings cause micro cuts in the belt, which lead the belt being more prone to cracking. The influence of the cuttings in the housing is now displayed to the product developer. Without this additional knowledge, the product developer consequently would have derived requirements necessary to enhance the belt by analysing the influence of the factors captured by sensors. However, in this case the improvement of the chassis would be more helpful as can be concluded from the additional knowledge extracted from unstructured feedback data. If further investigation is required, a view of the semantic data model can be opened that displays relevant causes and relations of the respective influence factor (see Figure 5).

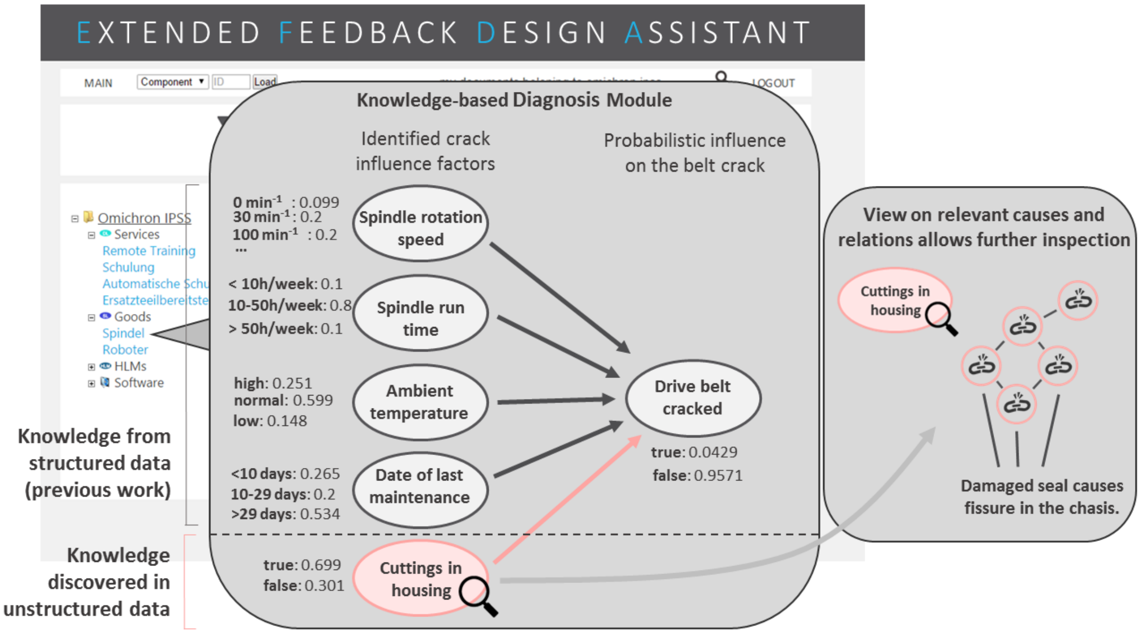

Figure 5. Additional value through knowledge from unstructured data during the product (component) improvement process

6 CONCLUSION AND FUTURE WORK

The presented concept aims at supporting product improvement processes by considering structured and unstructured product use information. The paper focusses on a methodology to process unstructured data in a way that it can be exploited during the product development phase and generate additional monetary and knowledge value. Product developers can utilize the developed feedback design assistant to derive requirements from analysing influences of a certain event. In order to do so, a dynamic and semantically rich data model has been developed and presented. The concept has been implemented and validated by extending our product use information feedback design assistant that has been continuously refined throughout the past years. The methodology achieved promising results. The additional knowledge extracted from unstructured data not only helps to identify additional improvement potentials of product components, but also shows that a sole consideration of structured data not always reveals the component's true vulnerabilities and thus in the worst case leads to unreasonable product improvements. Nevertheless, only the knowledge discovery in unstructured data has been validated using real life industrial data. The structured data base has been constructed using random values that have been derived from one test data set. Thus, a validation of both approaches using real life data remains to be done. Furthermore, the thesaurus used for the implementation and validation must be extended for an industrial application as the mapping process relies on the quality of the thesaurus. In future work, a visualization approach for the feedback design assistant's knowledge-based (what-if) analysis needs to be developed. Additionally, a demonstrator is in the works to simulate real life structured data.

REFERENCES

Abramovici, M., Gebus, P., Göbel, J.C. and Dang, H.B. (2016), "A Semantic Information Retrieval Framework within the Scope of IPS2-PLM", *Procedia CIRP*, Vol. 47, pp. 294–299.

Abramovici, M. and Herzog, O. (Eds.) (2016), *Engineering im Umfeld von Industrie 4.0: Einschätzungen und Handlungsbedarf, acatech Studie*, Herbert Utz Verlag, München, München.

Abramovici, M. and Lindner, A. (2011), "Providing product use knowledge for the design of improved product generations", *CIRP Annals - Manufacturing Technology*, Vol. 60 No. 1, pp. 211–214.

Abramovici, M. and Lindner, A. (2013), "Knowledge-based decision support for the improvement of standard products", *CIRP Annals - Manufacturing Technology*, Vol. 62 No. 1, pp. 159–162.

Dienst, S. (2014), *Analyse von Maschinendaten zur Entscheidungsunterstützung bei der Produktverbesserung durch die Anwendung eines Feedback Assistenz Systems*.

Dienst, S., Uhr, P., Fathi, M., Abramovici, M. and Lindner, A. (2012), "Concept for improving industrial goods via contextual knowledge provision", in Lindstaedt, S. (Ed.), *Proceedings of the 12th International Conference on Knowledge Management and Knowledge Technologies*, ACM, New York, NY.

Gandomi, A. and Haider, M. (2015), "Beyond the hype. Big data concepts, methods, and analytics", *International Journal of Information Management*, Vol. 35 No. 2, pp. 137–144.

Igba, J., Alemzadeh, K., Gibbons, P.M. and Henningsen, K. (2015), "A framework for optimising product performance through feedback and reuse of in-service experience", *Robotics and Computer-Integrated Manufacturing*, Vol. 36, pp. 2–12.

Laudon, K.C., Laudon, J.P. and Schoder, D. (2010), *Wirtschaftsinformatik: Eine Einführung, Wirtschaft*, 2., aktualisierte Aufl., Pearson Studium, München, Boston, Mass. [u.a.].

Lindner, A. (2015), *Methodik zur Verbesserung standardisierter Produkte auf der Basis der Auswertung von Nutzungsdaten deren Vorgängergenerationen, Berichte aus dem Maschinenbau*, 1. Aufl., Shaker, Herzogenrath.

Mamrot, M., Nicklas, J.-P., Schlüter, N., Winzer, P., Lindner, A. and Abramovici, M. (2016), "Concept for a Sustainable Industrial Product Service Systems based on Field Data", *Procedia CIRP*, Vol. 40, pp. 687–692.

Meier, H. (2012), *Integrierte Industrielle Sach- und Dienstleistungen: Vermarktung, Entwicklung und Erbringung hybrider Leistungsbündel*, Springer Berlin Heidelberg, Berlin, Heidelberg.

Meier, H., Roy, R. and Seliger, G. (2010), "Industrial Product-Service Systems—IPS2", *CIRP Annals - Manufacturing Technology*, Vol. 59 No. 2, pp. 607–627.

Neubach, M. (2010), *Wissensbasierte Rückführung von Produktnutzungsinformationen in die Produktentwicklung im Rahmen einer Product-Lifecycle-Management-(PLM)-Lösung, Maschinenbauinformatik*, Bd. 2010,2, Shaker, Aachen.

O'Connor, P.D.T. and Kleyner, A. (2012), *Practical reliability engineering*, 5th ed., Wiley, Hoboken, NJ.

Roy, R., Stark, R., Tracht, K., Takata, S. and Mori, M. (2016), "Continuous maintenance and the future – Foundations and technological challenges", *CIRP Annals - Manufacturing Technology*.