Reference Architecture framework for enhanced social media data analytics for Predictive Maintenance models

Jens Grambau
Faculty of Print and Media
Hochschule der Medien
Stuttgart, Germany
grambau@hdm-stuttgart.de

Arno Hitzges
Faculty of Print and Media
Hochschule der Medien
Stuttgart, Germany
hitzges@hdm-stuttgart.de

Boris Otto
Faculty of Mechanical Engineering
Technische Universität Dortmund
Dortmund, Germany
boris.otto@tu-dortmund.de

Abstract— **Social Media data contains a lot of hidden information which is currently rarely used in the manner of service topics on product level. However, a deep analysis of existing predictive maintenance models shows, that the combined use of social media data with already existing data from products or internal service data can improve existing and new analytical models for an enhancing predictive maintenance. Therefore, this framework paper describes an approach how to gather, process and analyze Social Media data related to products of a power tool producer. The defined processes are executed with the Azure Machine Learning Studio and are visualized with Power Bi. The main result of this paper is the Reference Architecture which combines several processes combine heterogenous data sources and enable the "First time to Incident" algorithm which helps companies to to increase the precision of Predictive Maintenance.**

Keywords— Reference Architecture, Reference Model, IoT, Predictive Maintenance, Predictive Analytics, Predictive Model, Social Media, Data Mining, Big Data, Service, CRM, Framework, Visualization, Products and Services

I. INTRODUCTION

In a world of digital transformation analytical solutions are becoming more and more important for companies. Grambau illustrates in the State-of-the-Art paper the latest approaches of Predictive Models in different industries with a focus on service-oriented topics. It can be seen that there are only a few approaches that link purely machine-generated data with human-made data from social networks, although this fusion could considerably increase the accuracy of forecasts [1]. However, this fusion cannot generally be applied to every product and service, because often there is no data available in the social networks or no connection of the data can be created. Grambau's analysis of existing Predictive Maintenance (PM) models clearly reveals the gap in research on a common approach how to use Social Media data in the context of service for a PM method. Furthermore, it becomes clear, that for most of the Predictive Models in the B2B sector, the enrichment with Social Media data is difficult. In contrast, the B2C sector can profit a lot of Social Media data, as there is much more information available which is highly user specific and can be linked to specific products and services. Nevertheless, for B2B businesses it is recommended to search and use existing data in Social Media networks as the available information contains a lot of tacit knowledge about products or services. Moreover, the information is predominantly written by heavy users or subject matter experts.

The new technical possibilities for the innovative and improved processing and use of data offer companies and their products as well as their services great potentials and synergies. The insights gained from the analysis of existing Predictive methods revealed the gap in existing research and justifies the need of the development of a new enhanced PM Model that additionally includes Social Media data as a new source of information. It is assumed that with the use of Social Media data PM Models can be more accurate and precise. This new data can provide a new view on products and services which then will lead to a better customer service. Not only the service maintenance of products can be enhanced but customer ideas also lead to even more innovative product development. Additionally, patterns can be derived from the analysis of the Social Media data. Additionally, this framework provides an implementation approach which can help existing PM Models to expand their analytical possibilities and features. Furthermore, companies can react earlier on customer needs and can improve their R&D for existing and new products. This framework helps companies to improve their overall data analytics approach as Social Media data is taken into consideration. This paper proposes a technical architecture based on a new developed reference architecture as well as an approach how to use and enhance existing data bases with Social Media data.

The structure of this this paper is organized as follows: In section II the objective, the research question and the research method are outlined. In the subsequent parts the reference model framework and structure are defined, and the PM framework is introduced. In section five the application of the reference architecture and the specific use case is described. In section six first results as well as hypotheses for the application of the use case are explained. In the conclusion main results are summarized and an outlook for future work is provided.

II. OBJECTIVE, RESEARCH QUESTION AND METHOD

A. Objective and definition of research questions

As outlined above, preceding literature review of predictive methods revealed the lack of a general valid concept that describes the use of Social Media data in the manner of PM. Thus, this work aims in providing a description of a generalized PM Framework which shows the end-to-end process from gathering Social Media data, preprocessing and analyzing the data with Machine Learning methods, combining the data with machine-generated data and further visualizing and transferring the analyzed data. Such a generalized Framework could offer companies advantages in using Social Media data for their customer service as the additional access to Social Media data provides a new social view on services and products which in the end leads to a stronger relation and a better customer service.

In consequence, this work aims in answering the following research question (RQ):

- RQ 1: How should a PM Framework look like to use and include Social Media data and its hidden information for a PM application?

This main research question can be specified into the following research sub questions (RSQ):

- RSQ 1: How should the reference architecture of the enhanced PM Model be designed and how need the different layers to be adapted?
- RSQ 2: How should the enhanced PM Framework process look like for being able to gather, preprocess and analyze Social Media data?
- RSQ 3: How can the analyzed data be visualized and transferred to other systems that companies are able to use and benefit from the insights?

This paper refers to the second part of a dissertation, as shown in Fig. 1, the Concept Phase – Framework development and aims in developing of the PM framework based on the Reference Model for using Social Media data in this context.

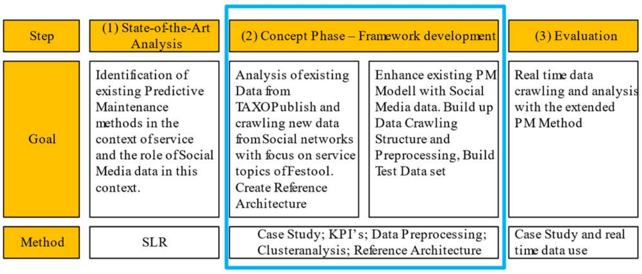

Fig. 1: Dissertation structured design, 2019.

The outcome of the research is a new framework which can be applied to different industries and sectors for the enhanced use of Social Media data for predictable services. Thus, a transformative worldview can be enabled as there is no longer only a technical view, but also a social view on the data.

The challenge of this research lies one the one hand in the nature of the data, as there are different sources with different heterogeneous formats available and on the other hand in the handling of the data to get a predictable outcome with the focus on the use cases.

The following sections of this paper outline the PM framework and its reference architecture, the PM Cockpit, and the PM Framework process. The main application of the PM Framework describes the case study and the generic process steps, the data collection, the data preprocessing, the data analyzation and the visualization of the results. In the result section the outcomes as well as hypotheses will be described which will be evaluated the continuative dissertation.

B. *Research Method*

A common research strategy in the discipline of business informatics comprises the construction of a reference architecture. Schütte defines the construction process of a new reference model through a structured process divided into five phases [2].

Within this work, a new reference model is created through following the five processes, that are described in detail below [3–5].

The first phase defines the basis and the common definition of the research question or the problem statement. It includes also a terminology which clearly identifies the attributes and objects which are used in the reference model. The research question defines the aim and the target of the new designed reference architecture.

A reference model framework defines a standardized model which is used to describe terms and elements of the framework on one basic language and reflects phase two. Schütte explains that the use of a so called "master-reference model" is useful for the universal structure and standardization.

Modelling languages are used in the third phase to formalize the framework which was defined in the second phase and create a so-called reference model structure. Therefore, detailed process and data modelling is done with usage of different languages like UML, EPK or ER-model. As the single languages are structured and consistent a syntactic and semantic correctness can be guaranteed. With a concrete case the processes and data models can be designed and be used for deducing of the new reference architecture.

In phase four the case described and modelled in phase three will be generalized. The idea is to create from a case-based structure a deduction to a generalized reference model with all its processes. The generalized reference model can then be used and applied to different companies or cases. This enables to use benchmarking based on the reference model. This means basically that from an internal and external company point of view the new reference model is applicable.

The application of the new generalized reference model reflects the fifth phase. It can be defined as the integration of the reference model in an existing environment of a company. The integration can vary from the implementation of single parts of processes or the aggregation of many reference models' parts to a lager reference model integration.

III. Reference Model Framework and Structure

A. *Construction of reference model framework*

The construction of a reference model is based on the consolidated view of knowledge. Schütte outlines the importance of using a defined language when constructing a reference model and additionally, recommends at a certain point of time a benchmark for the specific designed information system. Most of the designed reference models reflect process knowledge and thus, represent a good basic documentation that can be used for constructing the reference model framework. Moreover, the reference model can be developed out of a practical use case or out of research reflections [4].

This research describes the way of using existing or new data sources for a new analysis of products and services in the context of customer service. As in many companies dedicated data bases are used for different systems like CRM, ERP, MD, etc. the ANSI-SPARC architecture is set as the base reference model framework. The basic three-level architecture provides the advantage to differentiate between the user view, the internal concept level and the internal database level. Furthermore, this three-level architecture can be simply interpreted as a schema shown in Fig. 2, where all layers can communicate with each other to process data in the best way and to extract the maximum achievable benefit from data and the conceptual processes. In the internal database level, the physical data structure is defined which enables an efficient access to data. The main part of the conceptual level shows how data is processed, what the interfaces are and where the data is saved.

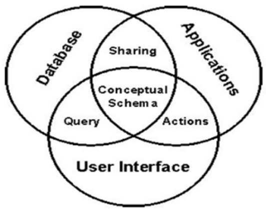

Fig. 2: ANSI-SPARC schema, [6].

The Standard ASCI-SPARC architecture is transformed for this work and shows already a few new elements which will be developed for the specific construction of the new reference model structure.
Fig. 3 provides an architectural template of the reference architecture for the Social Media data enhanced PM Model. It is aligned to the standard ANSI-SPARC Architecture which was designed to combine and conclude all necessary parts of the framework for a common understanding [6] [7].

The reference architecture recommends the new view on the developed framework. There are several works available which mention key structures of architectures with different level types. Ong, et al. describes their data driven architecture in a five layer perspective as the Data source layer and Metadata layer can be aligned to the Data Layer and the External Data sources in this paper [8]. As enterprise systems or services are mandatory to use and work with the data also the connection between the different layers and processes are important [9]. The connection illustrated by Paz et al. are implemented in this architecture in a more streamlined way. The internal layers are anyhow connected in-between and not only in this three-layer view. A lot of architectures are basically defined by an internal part, where most of the data is processed, stored transferred and visualized, and by an external part, that mainly consists of external services to gather, transfer or store the data [10]. The reference architecture was enhanced in having an external data source as well as the Data Layer which contains all available data within a company. Wang describes the different architecture layers in three main layers: The consumer, the platform and the providers which is aligned to End User Layer, Logical and Data Layer and External Data sources Layer [11].

As ANSI-SPARC was proposed in the mid-seventies it can be assumed that this architecture is outdated. However, in many modern RDBMS this architecture is still the basis for the construction and explanation of the database system landscape. The main benefit of the ANSI-SPARC architecture is the general point of view, that is used to describe an already existing system from a company in an easy way. Moreover, new data bases and its structures, new applications and its connection as well as new user interfaces can be designed and connected with the existing and new components [6].

By applying the ANSI-SPARC to this work, the basic reference model shown in Fig. 3 can be created in order to represent the structure of the cooperating company in the context of PM and the social web.

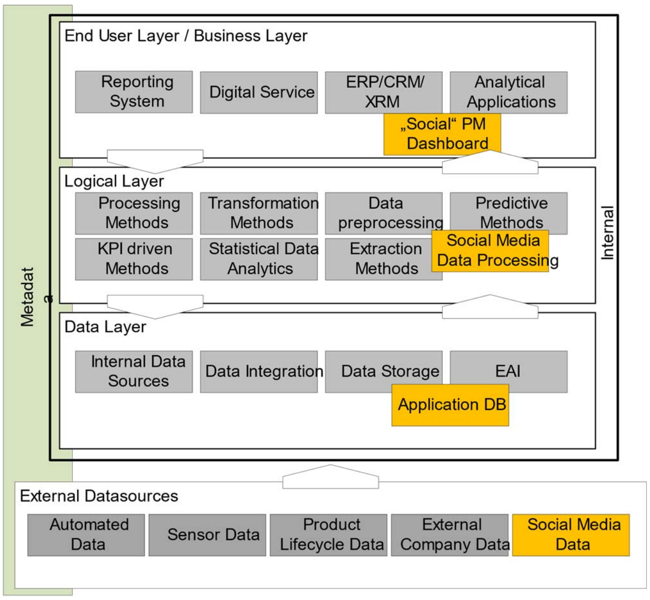

Fig. 3: Basic Reference Model derived from ANSI-SPARC, 2019.

B. *Construction of the rereference model structure*

The reference model framework paired with the reference model structure is the basis for the reference architecture for the enhanced PM analytical approach. A variety of industry sector can be supported by using this new model and improve their data analytics. To design and describe the new structure one base language, the Unified modeling language (UML), is used. UML is one of the most common notations to design and describe software systems. Its base is a common notation and semantic to visualize, construct and document models of business processes which should be available for object-oriented software development. There are 13 different types of diagrams to describe processes. For this work the type of activity diagrams is used, because procedures, processes and algorithms can be visualized and constructed [12–14].

The use case is based on a power tool producer located in Germany. This company is chosen because of the huge amount of data which can be gathered in the internet or social web. Although the business model is highly B2B driven, end users communicate a lot in the internet. As the social web data related to this company is not yet fully gathered, analyzed and used in the purpose of PM, this represents a suitable case for this work. The reference model structure outlined in the following is based on available social data about this power tool producer. In this use case there are no existing internal processes for data analytics. Thus, the established reference model is designed from scratch and can be applied to any IT environment of other company.

To build a new PM model different main processes are necessary which need to be implemented in the existing business processes of the company. The new designed processes can be divided into four main steps. In the beginning there is the data capturing process which is visualized in Fig. 4 and contains the description of the additional sources from the social web where specific product data will be gathered. A Social Listening tool (Microsoft Social Engagement) is used to gather specific data from open accessible social networks, blogs and newsfeeds. Furthermore, a web scrapper (WebHarvey) is used to gather product reviews from different platforms such as amazon, google shopping, CPO and further. The data can then be located at different places and must be transferred to one centralized local storage where the data can be processed. For this processing different interfaces are needed which are dependent on the used technologies. In this use case Azure Hub services are used to transfer the data from the Social Listening tool into an Azure SQL database and transferred to a local SQL database. The data from the web scrapper is saved directly in a local database.

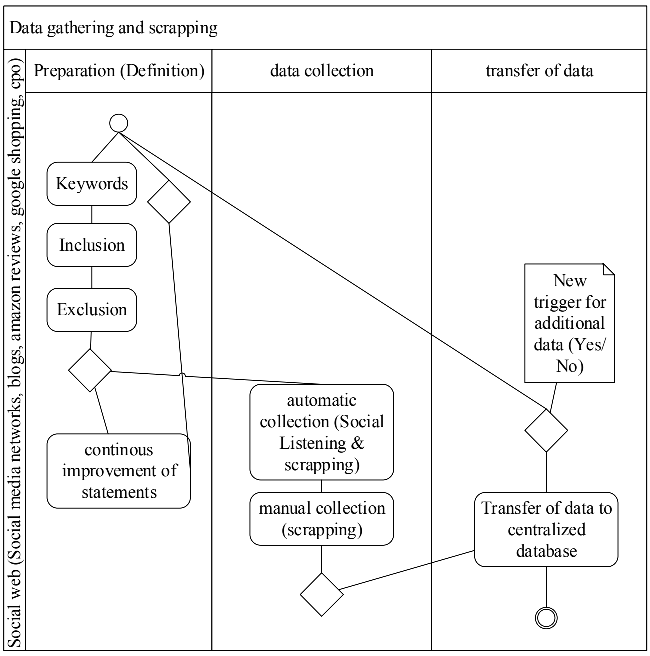

Fig. 4: UML data-capturing, 2019.

In the second stage the gathered data must be preprocessed. There are several steps needed to process the data and make it available for the analytical part. First, there will be SQL statements executed which remove the Social Media data from unnecessary links, html tags and stop words. After this cleansing the SQL data will be transformed into a .csv file. The second main source of data are the product reviews from different shopping or review websites. Additionally, there are several actions executed to clear out unnecessary metadata like stop words, html tags and emojis. A normalization of the text data will be done as the algorithms can work more efficiently in this lower-case text modules. In the last step the preprocessed data is exported to a .csv file. As Azure Machine Learning Studio (AML) is mainly used in this use case, the gathered data must be converted into the .csv file format for transferring and processing the data. Fig. 5 additionally shows webservice as a possible last step. A webservice can be used as part of an automated process that transfers data from one to next step.

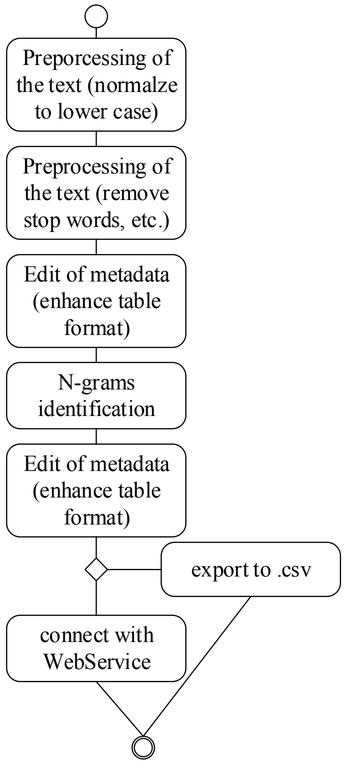

Fig. 5: UML data-preprocessing, 2019.

The third process is a simple word count process that analyzes the size of the Social Media posts or the product reviews through counting the number of words. The number of words in a single data record is important for the analyzation in a later stage as there can be correlations deduced and causalities be confirmed or declined. The analyzation of the number of words is executed by simplified SQL scripts as written in Fig. 6.

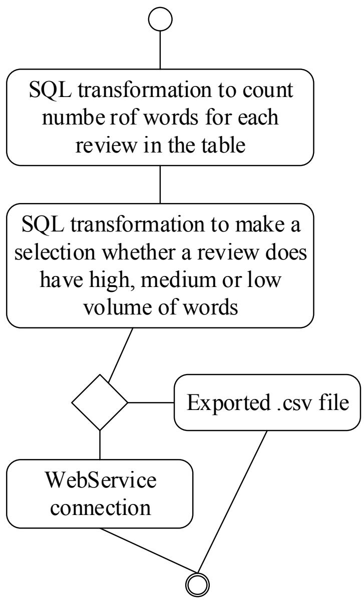

Fig. 6: UML word-counts, 2019.

The application of the machine learning algorithms is based on the preprocessed data. In the social web many posts are not directly related to specific products. But more often they are related to a supplier, another vendor or the logistic company who delivers the product at the end. Therefore, it needs to be identified whether Social Media post, a review or any other data record is assigned to the product, the seller or other topics. The last process visualization in Fig. 7 shows the process steps needed for this identification. For this use case of a power tool producer a differentiation between the "product" and the "seller" is done. If there is no clear assignment possible, the post can be related to "both".

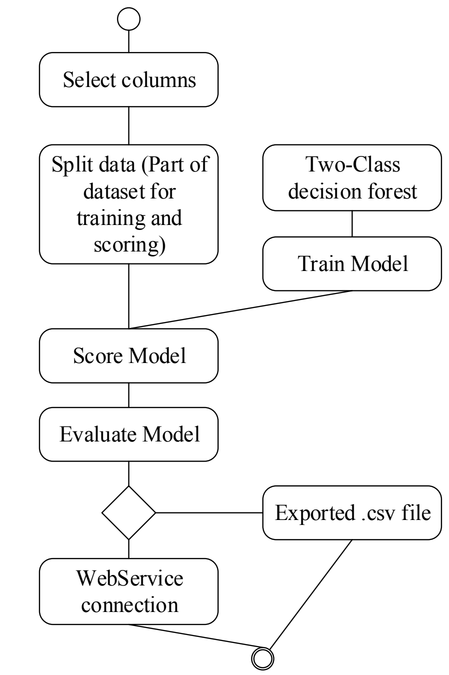

Fig. 7: UML seller or product identification, 2019.

After the identification of the product or seller data records a sentiment analysis will be done. The sentiment analysis is based on a common sentiment database called "SENTIWORDNET" [15]. This database is one of the most common sentiment databases and open source based. The benefit of analyzing the sentiment with one common sentiment database is that one aligned base of the sentiment can be used. For the products of the power tool producer there are two main languages for the data records, English and German.

For an enhanced PM approach new intelligent analytical method must be designed considering knowledge from existing methods. To include Social Media data in a PM model this data needs to be combined with IoT. It is important to identify and understand the cause-effect relationship from the Social Media and IoT data perspective. The PM Model will analyze continuously incoming product reviews as well as Social Media data like posts, tweets, statements and further posts that are related to errors, failures or damages of products. The words used for the analysis are defined in a new list which also includes words from a hierarchical reclamation list. It is important to update this list as soon as there is a new word used to describe certain circumstances of defects. The model analyzes data on a product level and determines how much and which posts are written in the web. If the number of posts will increase or decrease significantly, a notification will be sent immediately to the customer service team. Furthermore, the second PM process which is running is continuously analyzing the IoT data and compares those swings of the Social Media data with the ones from the mechanical data. The combination of those two data sources is then used for the visualization and the predictive approach of forecasting different cases like the needed service hours, the needed parts and the needed hyper care phase if a new product is released.

IV. PREDICTIVE MAINTENACE FRAMEWORK

Grambau outlines in the State-of-the-Art paper that the use of Social Media data for PM approaches is quite rare [1]. Moreover, most of these approaches use only machine-generated data or only Social Media data. Additionally, there is no common method, but each approach uses different technologies and analyzation models to gain new insights from their data. Overall it is clearly shown, the more data is available the better the analyzation result is.

In this section the generic PM Framework is described which basically is a new concept to describe how to use Social Media data in a PM service cases.

A. *PM Cockpit*

Most of the existing approaches and technologies are aiming for gathering as much data as possible from technical sources, like sensor data, and to present an overview of a certain topic. The approach within this paper is to combine different data sources especially Social Media data with technical data. For this a new so-called PM Cockpit architecture is needed. The architecture in Fig. 8 describes the reference architecture transformed in the framework approach which consist out of four basic parts. The external data sources, the internal data sources and databases, the application software and the end-user application.

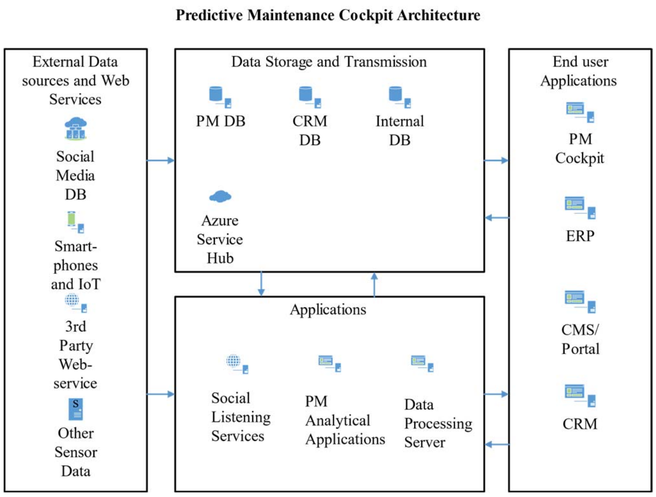

Fig. 8: Technical architecture of the Predictive Maintenance Cockpit, 2019.

On the left side are the external data sources and web services. These are all publicly available information within and outside of Social Media networks and the world wide web itself. Moreover, sensor data from smartphones and IoT and other sensor data is considered. With different enterprise and application interfaces, Azure Service Hub technology and Social Listening tools the necessary data can be pulled and saved in the internal databases shown in the top box. Furthermore, the box shows the internal data sources: CRM and ERP as well as the custom internal databases, the PM database and specific databases for products and services of the company.

The second internal part is the application box. In here the data of the external and internal sources are processed. Social Listening applications crawl data from Social Media networks, the data processing application consolidate all the available data which is then transferred and processed within the PM application. In more detail the gathered data from the external and internal sources will be processed with specific software and algorithms and saved in a centralized database. The merged data will again be processed from machine learning software and algorithms, and the results are provided to the end-user through CRM or the PM Cockpit.

In the End-user Application box on the right all applications which can visualize the processed data are shown. Additionally, here the user can interact with the system and for example trigger a new request of a prediction.

Furthermore, it is highly important that compliance and data regulations as well as data laws within the European Union (EU) and the single countries where the tool should be used are considered. Many of the tools and applications are only available as a cloud solution and are located outside of the EU through which legal situation changes. This might be a risk for an implementation of a PM solution with the mentioned technologies in this paper. Therefore, an analysis has to be done to check the behavior and rights of the end user which work with the tool and which are providing information through Social Media or any other channel to a company. Moreover, there could be certain concerns and behaviors about the liberality of data of the people which differ from country to country.

For example, in Germany people are very skeptical about data transfers but not everybody reads the data privacy and the general terms and conditions or don't think about using applications and websites. It's a paradoxical situation which makes it necessary that the people are good informed what happen with their data, how it is processed and what is finally done with it.

For the use of PM technology an efficient way of using available resources is mandatory. In most of the cases the resources from the cloud are used and rented for a certain time when the applications are running. For this it is necessary to plan the analytical processes to have to most appropriate costs in place.

Companies, their customers and end users benefit from the PM Cockpit in regards of products and services as they have more information available which results in an increasing quality of customer service activities. Service waiting times can be reduced and error sources are identified much faster. Furthermore, maintenance is better predictable, and the costs are reduced as unnecessary maintenance can be excluded in fact of eliminated error sources. End users also have the possibility to influence future product and service developments as their opinions and shared experiences are considered.

With this PM cockpit architecture following important facts are implemented:

- Connection of heterogeneous data sources - Social Media data and the internal company data.
- Identification of Social Media data's influence on predictions which only used machine generated data so far.
- Usage of new information to improve products and services through the experience of the customers.
- Determination of improved predictions for service orientated topics of products or services.
- Visualization of the filtered data in a so-called Service Cockpit or a CRM system.

B. PM Framework Process

The PM Cockpit is the base for the new PM Framework process which generically defines how the data is gathered, transferred via EAI/ API's to the internal databases, processed and visualized. If information and data is shared and created in real time in Social Media networks and analyzed afterwards, new insights and error sources might be detected. These errors could not be detected by analyzing only technical data.

The generic PM Framework consist out of four major process steps:

- The gathering and filtering of the data from the different sources.
- The processing of the data. Reducing redundancies and missing or invalid data.
- Validation and training of the prediction algorithm based on specific parameters. This validation is an iterative process and thus, it can be reacted very fast to changing parameters and the model is trained.
- In the last step all the validated data will be pushed and visualized to the PM database and be visualized in the cockpit or in a CRM system.

The four major process steps are also reflecting the base architecture of the PM Cockpit. The gathering of data from external sources, the internal storage and processing of the gathered data and the visualization to the end-user.

The main process describes the overall picture of an enhanced process model for Social Media data in the manner of Predictive Models. This overview not shows the individual processes which are preprocessing and analyzing the data to be ready for the output.

C. Enrichment of existing predictive maintenance model

The enrichment of an already existing PM method is useful as soon as there is an IoT integration set up. The idea behind this enhancement is a product related analyzation of data from the tool and data from the Social Media sources.

To implement a combined solution the following components are needed.

- IoT Infrastructure
- Registered user and data collection of in field tools
- In place analytics to preprocess data
- Product group data hosting
- API's
- Agreement with user to process data

If those major components are already in place or taken into consideration the data from Social Media can be gathered and implemented.

V. Application of Reference Architecture

In this section the experimental setup is described. The case study defines in which industry sector and in which specific company the reference model will be evaluated. The following table outline the participants, the considered data and the specific activities of the practical evaluation.

The experimental setup is split into two phases and contains the necessary data sources which contain data over a defined timeframe beginning in January 2017 to April 2019. In phase one mainly the data preprocessing as well as a first analyzation with the tool Power BI is done. Power BI is used as it is a powerful tool which helps to analyze data in a visual way. There are already a lot of different examples how Power BI is used like for Sales Pipeline tracking in bigger companies or even in Universities to help students measure their performance [17].

TABLE 1: Case study definition

Setting	Power tool producer, a German company which produces power tools
Participants	• Social Media user • Online shop user • Company employees
Included data	• Internal Customer data • Social Media Posts • Shopping review data • Product data
Activities	• Gather data • Preprocess data • Analyze data • Visualize and transfer data

The second phase has the focus on the analyzation of the data sets with machine learning algorithms. One of the tools is the Azure Machine Learning Studio which provides a huge amount of Machine Learning components [18]. The output and the results will then again be imported in Power BI for the better visualization.

VI. Results and Expected Result Hypotheses for the Application Use Case

The findings taken from this work are formulated into hypotheses, correlations, causalities and KPI's. These aspects are evaluated in the final step of the dissertation. These statements have been derived from the first analysis with the anonymized data sets by using Power BI and the Azure Machine Learning Studio including the defined processes from section III. Furthermore, those statements help to understand the benefit of incorporating Social Media data in PM models.

- With the analysis of the Social Media and review data over a certain timeframe (1 year, ½ year, 3 months) which is related to a product table the lifecycle of the products can be identified as well as patterns can be recognized which identify the occurrence of product failures.
- A critical mass of posts can be identified which might indicate a product failure or defect and thus, the necessity of a service case.
- The sentiment of a post is not necessarily in line with its rating. Thus, the sentiment analysis reveals additional information which might have been overlooked as the review is rated pretty good.

In the analyzation part of the framework different correlations between the datasets themselves or between the data and the behavior of people in Social Media networks can be identified.

- If in Social Media the number of posts for a specific product increases, the number of product related reviews will increase.
- If the sentiment of Social Media posts is mainly negative, the amount of negative written reviews and received service cases will increase as well.

Moreover, based on the mentioned correlations and analysis of the data several causalities can be derived.

- If the sentiment of Social Media posts is mainly positive, the number of written reviews and created service cases will decrease or stabilize on a low value.
- If the number of Social Media posts with a positive sentiment increase, the number of positive written reviews will increase whereby the total number of reviews and the number of "negative" service requests will decrease.
- The more product reviews and Social Media posts are generated the higher the potential for innovative product development and problem solutions is.
- If the product reviews have a negative sentiment, the product rating will be low.
- If the sentiment is negative, the number of words used in a single review will be higher than of a review with positive sentiment.

The use of different technology tools enables the identification of different clusters inside the base data. With this possibility new clusters or even patterns in the base data as well as in the analyzed data can be detected.

- There will be a cluster where Influencer for power tools can be identified.
- There will be a cluster which identifies the most responsive nations in the world.

From a business perspective it is important to measure the analyzed data as well as interpret several conditions like "sentiment" as a KPI. The defined KPI's can be triggered from the business to measure the data which is available in the specific use case and additionally to identify the value of the data itself in the manner of a return on invest approach.

- PR-Ratio: Number of posts in relation to the number of reviews.
- PRSent-Ratio: The overall sentiment of Social Media posts in relation to overall sentiment of reviews.
- PoRe-Div: The number of the Social Media post divided with the reviews of the specific product. The higher this value is, the higher is the potential that something happened with the product.

- TTFI: Time to first service incident. Relation of IoT data to the amount over time of Social Media and review data.

The mentioned KPI’s are newly defined in the progress of this paper. The following two KPI’s will be developed in the Azure Machine Learning Studio.

- Rating of Keywords from different data sources like Twitter, Facebook, Amazon reviews, Google shopping reviews.

- Time to first incident. Included in this mathematical calculation are the date of purchase, social media data, average value of the sentiment, number of service cases and average product review rating

VII. Conclusion

The paper illustrates a framework to gather, process and analyze Social Media data in the manner of service-related topics. In this work the established framework was applied to sample and anonymized product data from a power tool producer. However, the data was limited, significant insights could be gathered, and correlations and causalities be derived.

To verify the defined processes in this framework paper and evaluate the findings of section VI, the subsequent research will use the processes in a real-world use case with a power tool producer. New data sources as sensor data and further mechanical data will be integrated and combined with processes from this framework of the enhanced PM model.

Future research will incorporate following components:

- Cooperation with a power tool producer to define further use cases for the beneficial application of the enhanced PM model.

- Sentiment analysis on a defined data base.

- Evaluation and automation of the framework processes.

- Validation of the Reference Architecture and the derived correlations and causalities.

- Deriving prediction for specific products.

Thus, continuous research will take the output of this work as a basis in order to establish a final and verified Predictive Maintenance model that incorporates Social Media data and aims in enhancement of service processes.

References

[1] J. Grambau, A. Hitzges, and B. Otto, “Predictive Maintenance in the Context of Service - A State-of-the-Art Analysis of Predictive Models and the Role of Social Media Data in this Context,” in *ICEIS*, 2018, pp. 223–230.

[2] M. Rosemann and R. Schütte, “Multiperspektivische Referenzmodellierung: Scientific Figure on ResearchGate,” in *Referenzmodellierung*, J. Becker, M. Rosemann, and R. Schütte, Eds., Heidelberg: Physica-Verlag HD, 1999, pp. 22–44.

[3] M. Maicher and H.-J. Scheruhn, *Informationsmodellierung: Referenzmodelle und Werkzeuge*. Wiesbaden: Deutscher Universitätsverlag, 1998.

[4] R. Schütte, *Grundsätze ordnungsmäßiger Referenzmodellierung*. Wiesbaden: Gabler Verlag, 1998.

[5] J. Vom Brocke, *Referenzmodellierung: Gestaltung und Verteilung von Konstruktionsprozessen*. Zugl.: Münster, Univ., Diss., 2002, 2nd ed. Berlin: Logos, 2015.

[6] G. Patel, “Study Of The ANSI/SPARC Architecture,” in *International Journal of Modern Trends in Engineering and Research [Kurztitel fehlt!]*.

[7] *ANSI-SPARC*, 1975.

[8] I. Ong, P. Siew, and S. Wong, “A Five-Layered Business Intelligence Architecture,” *CIBIMA*, pp. 1–11, 2011.

[9] J. F. de Paz *et al., Eds., An Architecture for Proactive Maintenance in the Machinery Industry: Ambient Intelligence– Software and Applications – 8th International Symposium on Ambient Intelligence (ISAmI 2017)*: Springer International Publishing, 2017.

[10] H.-G. Kemper, W. Mehanna, and H. Baars, *Business Intelligence - Grundlagen und praktische Anwendungen: Eine Einführung in die IT-basierte Managementunterstützung,* 3rd ed. Wiesbaden: Vieweg + Teubner, 2010.

[11] L. Wang and X. V. Wang, Eds., *Cloud-Based Cyber-Physical Systems in Manufacturing*. Cham: Springer International Publishing, 2018.

[12] G. Heinrich and K. Mairon, *Objektorientierte Systemanalyse*. München: Oldenbourg, 2008.

[13] A. Kemper and A. Eickler, *Datenbanksysteme: Eine Einführung,* 8th ed. München: Oldenbourg, 2011.

[14] K. C. Laudon, J. P. Laudon, and D. Schoder, *Wirtschaftsinformatik: Eine Einführung,* 3rd ed. Hallbergmoos: Pearson, 2016.

[15] S. Baccianella, A. Esuli, and F. Sebastiani, “SentiWordNet 3.0: An Enhanced Lexical Resource for Sentiment Analysis and Opinion Mining,” in *LREC*, 2010.

[16] J. Yang and A. M. Anwar, Eds., *Social Media Analysis on Evaluating Organisational Performance a Railway Service Management Context*. 2016 IEEE 14th Intl Conf on Dependable, Autonomic and Secure Computing, 14th Intl Conf on Pervasive Intelligence and Computing, 2nd Intl Conf on Big Data Intelligence and Computing and Cyber Science and Technology Congress(DASC/PiCom/DataCom/CyberSciTech), 2016.

[17] J. Sluijter and M. Otten, “Business intelligence (BI) for personalized student dashboards,” in *Proceedings of the Seventh International Learning Analytics & Knowledge Conference*, Vancouver, British Columbia, Canada: ACM, 2017, pp. 562–563.

[18] Microsoft Corporation, *Microsoft Azure Machine Learning Studio.* [Online] Available: https://azure.microsoft.com/de-de/services/machine-learning-studio/. Accessed on: Oct. 15 2018.

978-1-7281-3401-7/18/$31.00 ©2019 IEEE

2019 IEEE International Conference on Engineering, Technology and Innovation (ICE/ITMC)