Article

Data-Based Stakeholder Identification in Technical Change Management

Fabian Sippl [1,*], Renè Magg [1], Carla Paulina Gil [2], Steffen Düring [2] and Gunther Reinhart [1]

[1]   Institute for Machine Tools and Industrial Management, Boltzmannstraße 15, 85747 Garching, Germany
[2]   BMW Group, Knorrstraße 147, 80788 München, Germany
*   Correspondence: fabian.sippl@iwb.tum.de

Abstract: The efficient and effective handling of technical changes in product and production is seen as an important factor for the long-term success of manufacturing companies. Within the associated processes, the engineering and manufacturing change management, the identification and involvement of all relevant stakeholders, i.e., departments and employees, plays an essential role. Overlooking relevant stakeholders can lead to unforeseen impacts, such as production stops or further necessary changes, and can cause unforseen increased costs. In particular, in large companies, this task is complex and error-prone due to the high number of changes and departments involved, as well as the abundant variety of changes that can take place. Therefore, this contribution introduces an approach for stakeholder identification in technical change management, which allows the automated identification of relevant stakeholders at the beginning of the reactive phases of the change management process. The approach describes all necessary steps from data preparation to the evaluation of the obtained classification models. It is based on a text-classification approach and focuses in particular on the additional integration of expert knowledge to increase model quality. The approach has been successfully applied in cooperation with a German automotive company, and the obtained model quality has been compared to an expert-based classification.

Keywords: stakeholder identification; engineering; manufacturing; change management; text classification

1. Introduction

Manufacturing companies operate in an environment that is often described as increasingly volatile, uncertain, complex, and ambiguous [1]. In recent years, this VUCA world has led to intensive discussions of various paradigms to tackle this complexity in manufacturing, such as flexibility, agility, reconfigurability, and the introduction of concrete approaches such as continuous factory planning or change management. These approaches are intended to enable an efficient and effective handling of the high and permanent need for changes in the product (engineering change) and production (manufacturing change), which are caused, e.g., by rapidly changing customer requirements and legislation or shortened technology life-cycles [2]. Technical change management that is efficient and effective is seen as a critical key factor for the long-term success of manufacturing companies [3]. The associated network of activities refers "to organizing and controlling the process of making alterations to a factory or a product". This includes "the totality of measures to avoid and specifically front-load as well as to efficiently plan, select, process, and control" manufacturing changes (MC) as well as engineering changes (EC) [4–6]. The handling of EC is therefore referred to as engineering change management (ECM) and the handling of MC as manufacturing change management (MCM). Both domains are summarized in the remainder of this paper as technical change management (TCM). Within these processes, numerous activities have to be completed before a change can be finally implemented, such as change description, solution development, change impact

analysis, or the approval of the change. An important prerequisite for the successful execution of these activities is the identification of all relevant stakeholders in order to obtain all necessary assessments of the planned changes and to take the gained knowledge into consideration in the planning process [7,8]. In the case of this contribution, we consider stakeholders at the level of the business units and do not focus on stakeholders at the employee level. The missing or delayed involvement of relevant stakeholders may cause additional costs, unplanned impacts, or even production delays [8]. In industrial practice, stakeholder identification (SI) is currently performed primarily by change coordinators based on their subjective experience, and knowledge [4,9]. Multiple preliminary works already demonstrate the potential of machine learning (ML) and data mining (DM)-based change management support [9–11]. However, these approaches are hardly used so far in industrial practice [12], although the widespread use of digital tools leads to a high availability of change data [12,13]. This change data contains textual information, such as the change title or change description, and also additional information, such as the change trigger, requestor, or affected assemblies [9,14]. Therefore, this contribution introduces a methodology for data-based SI by using text classification approaches and describes all necessary steps from data selection to evaluation and improvement of ML models by the integration of expert knowledge. The methodology focuses on the start of the reactive phase of change management, in which the initial identification of affected stakeholders takes place. The identification of all relevant stakeholders is particularly important in this step, as otherwise the risk of unconsidered change impacts increases in the subsequent step of the change impact analysis.

The remainder of the work is structured as follows: Section 2 introduces the fundamental concepts of change management and describes the basics of natural language processing (NLP) and text classification. Section 3 then provides an overview of the current state of research and identifies existing research gaps. The procedure for data-based SI is then described in detail in Section 4. The industrial application and evaluation is the subject of Section 5. The contribution concludes with a summary and an outlook on future research potential in Section 6.

2. Fundamentals

2.1. Stakeholder Identification

Based on literature from the field of project management and change management, stakeholders in TCM can be defined based on Freeman [15] and Koch [4] as *"any group or individual that can influence or is influenced by the avoidance, front-loading, selection, planning, implementation, and control of manufacturing or engineering changes"*. Beam et al. [16] show that systematic and strategic SI and its prioritization is also a crucial factor in R&D. A distinction can be made between internal and external stakeholders [17]. In the case of manufacturing companies, internal stakeholders refer, for example, to the company management, business units, or individual employees. External stakeholders refer, e.g., to society, legislators, shareholders, or suppliers. This work focuses explicitly on relevant internal stakeholders for the implementation of technical changes because external stakeholders, such as suppliers, are in most cases not represented in IT systems which support the change management process and therefore cannot be directly identified as relevant on the basis of past change data.

2.2. Management of Technical Changes

Manufacturing companies are forced to permanently plan and implement EC as well as MC to adapt their products and the associated production processes to new conditions and requirements. An EC is defined as "an alteration made to parts, drawings, or software that have already been released during the product design process". Closely following this definition, Koch [4] describes MC as "an alteration made to the factory or its elements that have been released for or are already in operations" based on [18–23]. Both change types "can be of any size or type, it can involve any number of people, and take any

length of time of people, and take any length of time" [4]. The process of dealing with EC and MC, the ECM and MCM, can be equivalently divided into three overarching phases. The proactive phase deals with the early identification, avoidance, and front-loading of change demands. In the reactive phase, the need for change manifests itself, and the change is prepared, evaluated, and implemented. Finally, in the retrospective phase, lessons learned are conducted to gain useful knowledge for future changes. This contribution focuses on the reactive phase, as at the beginning of this phase, the information about the intended solution concept for the change demand is captured for the first time in a structured form within the manufacturing change request, or the engineering change request [4,5]. Both are referred to as change request (CR) in the following.

2.3. Text Classification

Text classification is an application of NLP [24]. Overall, NLP comprises approaches to automatically analyze and process human language with special algorithms and to elaborate correlations and patterns [25–28]. Languages are very extensive, contain an infinite number of ways to form sentences, and ambiguities, as words as well as phrases can have different meanings in different contexts [25,29]. In this regard, developing approaches that can automatically understand language and derive insights from textual information is a challenging task [25,29]. To solve this task, NLP combines computational linguistics with statistical, machine learning, and deep learning models [29,30].

Within this contribution, text classification is used to analyze the information provided in CRs (e.g., change description or change cause) to determine whether specific stakeholders in the change process are affected by a change. The target variable of the classification problem is therefore the relevance of a CR for a certain business unit (relevant or not relevant). For this purpose, a hierarchical attention network (HAN) according to Yang et al. [31] is applied. The approach of Yang et al. [31] was developed on the basis of several assumptions. On the one hand, text documents have a hierarchical structure in which words form sentences, and sentences, in turn, form the text document. On the other hand, words within sentences have different meanings depending on the context and thus should not be considered in isolation. Furthermore, different parts of a text document are more or less relevant for the classification task. The aim of the model is to map the hierarchical structure of text documents and thus to identify relevant parts of the text document in order to be able to make meaningful classification decisions. For this purpose, the HAN is divided into two levels, which represent the hierarchical structure of text documents, as visualized in Figure 1. An exemplary implementation of a HAN can be found in Nguyen, Viet [32].

The first level consists of a word encoder and an attention layer on the word level. The second level consists of a sentence encoder and an attention layer on the sentence level. The word and sentence encoders are implemented by bidirectional gated recurrent units (GRU), which are used to map words and sentences with respect to their contexts. Within the attention layer, multilayer perceptrons are applied to identify informative words or phrases. The input variables for the HAN are vector representations (cf. Section 2.4) of the words of a text document, which are converted via the two layers into a vector representation of the text document. A Softmax function is used at the endpoint of the HAN to classify the documents, and the model quality is evaluated on the basis of the negative log-likelihood [31]. As hyperparameters, the maximum number of words and sentences in a document must be defined prior to the application of the HAN.
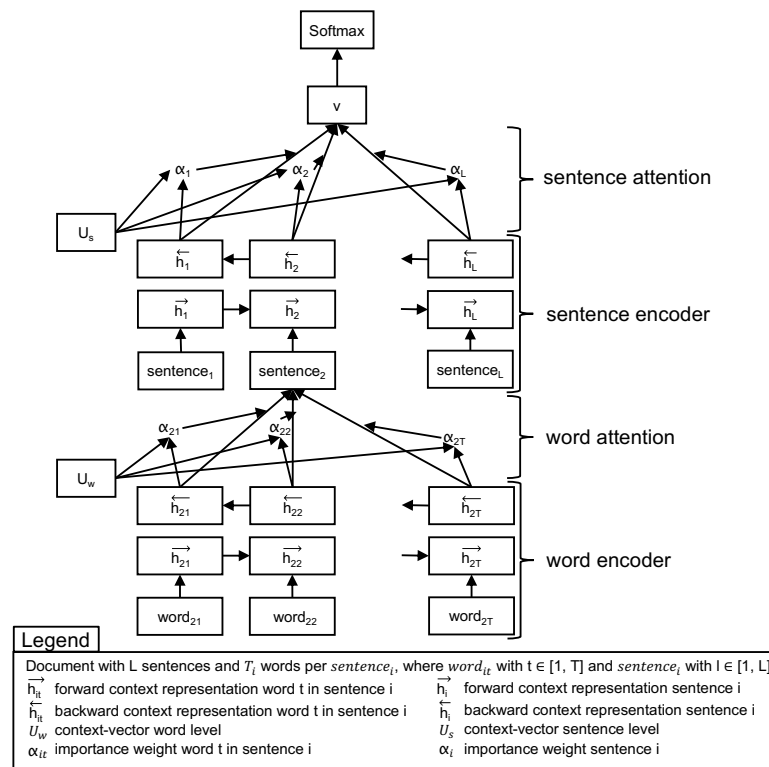
Figure 1. Architecture HAN cf. [31].

2.4. Word Embeddings

In order to enable a classification by using the HAN, the information provided in the CRs must be converted into numerical data [33]. The quality of the resulting models can be improved by considering the context of the words during transformation. Therefore, a word-embedding approach was chosen to convert the textual data into numerical data by using the Word2Vec approach according to Mikolov et al. [34]. The basic assumption behind Word2Vec is that words with similar meanings occur within similar context words. The corresponding algorithm learns word associations of a large text corpus by training a neural network consisting of an input layer, a hidden layer, and an output layer. In Figure 2, the principle of the Continous-Bag-of-Words (CBOW) model is visualized [34]. Prior to training, the words of each individual CR are tokenized. For each CR, a list of the individual words is thus transferred to the Word2Vec algorithm. When using the CBOW approach, each word of the list is used once as a targeted output, and the words surrounding the word are used as input variables. Over the analysis of a multitude of change documents, words occur in different contexts so that they are processed multiple times, allowing the neural network to learn the different contexts. The result of the training phase is a vector space in which each word of the text corpus is mapped to an individual vector. Within the vector space, the vectors of words that occur in the text corpus in a similar text context are close to each other. The vectors formed within the training subsequently provide the opportunity to be used as input data for text classification. The training of word embeddings needs to be performed only once based on a sufficiently large database. Subsequently, the vector representations of the words can be used for different applications [34]. Prior to the training of the word embeddings, the size of the vectors must be defined as an important hyperparameter because larger vectors offer the possibility to describe words more precisely, but at the same time, the dimensionality is increased so that the right measure must be found here.

Example: „Adapting the geometry within component"
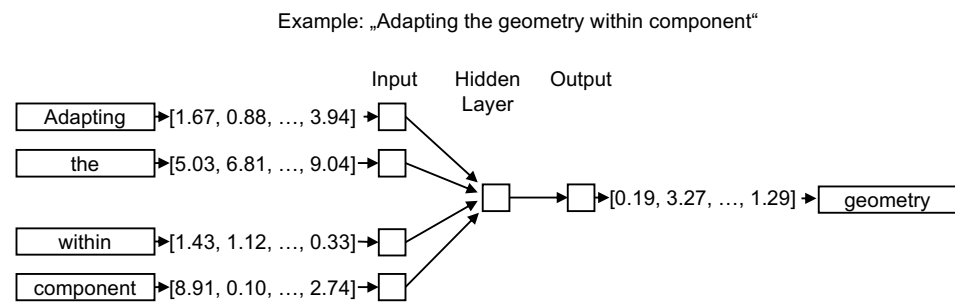
Figure 2. CBOW-Word2Vec cf. [34]

3. State of Research

A systematic literature study according to Webster and Watson [35] was conducted to identify relevant preliminary work. Figure 3 visualizes the methodology of the search strategy.
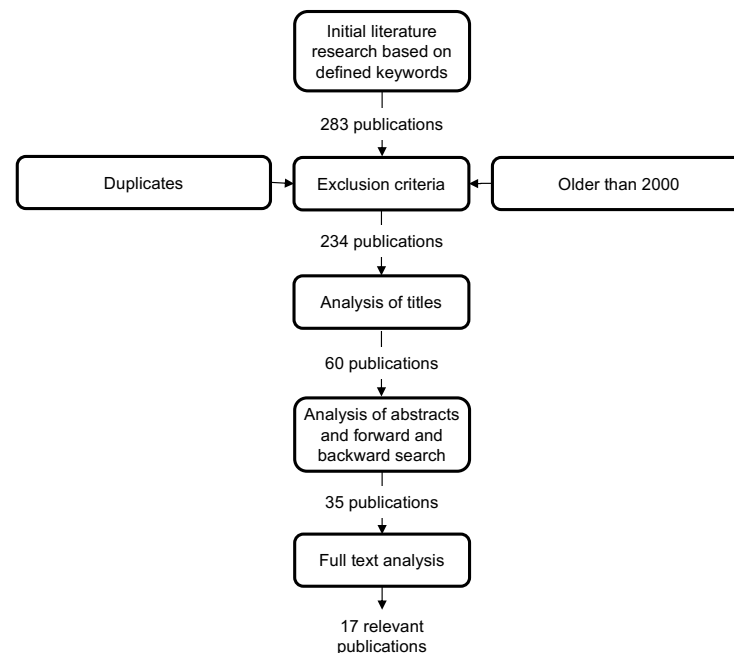
Figure 3. Methodology of the search strategy.

Following the approach of Webster and Watson, a systematic search strategy was applied by using different German and English search terms (cf. Appendix A). For this purpose, subject-specific terms were identified in an initial unstructured literature analysis in the fields relevant for this contribution, as e.g., change management or stakeholder identification. Additionally, similar and synonymous words were collected. All identified terms were used to search the literature databases Scopus and Web of Science, and thus led to the identification of 283 potentially relevant publications. Duplicates and all publications older than the year 2000 were then sorted out, leaving 234 publications. These results were narrowed down to 60 publications by analyzing the titles. The abstracts were then analyzed to decide on the actual relevance. For the publications considered as relevant based on the abstract, the keyword-based search strategy was extended via a forward and backward search. This procedure is intended to ensure that the results of the word-based search can be used to provide an overview of the subject area that is as complete as possible and in which no relevant works are overlooked. A total of 35 publications was thus classified as fundamentally relevant. These were further analyzed by means of a full-text analysis. With this approach, 17 publications were identified that represent the state of research

for the topics under consideration. Appendix A presents a categorization of relevant preliminary work.

3.1. Relevant Preliminary Work

Within the work of Koch [4] and VDA [36], reference processes for change management are presented, and roles are introduced for the individual change process steps. These roles can be used to identify suitable stakeholders when implementing changes. However, the defined roles represent standard information without allowing consideration of change-specific information. Sippl et al. [37] introduce an approach that allows SI based on an employee-specific modeling of responsibilities and a systematic matching with change-specific characteristics. Although this approach supports change-specific SI, it requires high manual effort for the preparation and application. The publications of Schuh et al. [38], Giffin et al. [39], Sharafi [10] , Pasqual and de Weck [40] and Kattner et al. [41], show different approaches to gain descriptive knowledge from past change data, without focusing on SI. For example, in the work of Giffin et al. [39], a network analysis of past CRs is performed by using graph theory and pattern recognition within a change network and a product network. Pasqual and de Weck [40] extend this work by social network analysis. The models presented in the publications of Kocar and Akgunduz [42] , Malak and Aurich [43] , Pan and Stark [44], Do [45] , Wickel [9], and Habhouba et al. [46] use past change data to determine the potential spread of change without considering affected stakeholders. For example, Wickel [9] addresses the challenge of change impacts within complex product structures. In her methodology, change-related dependencies between product components are identified and modeled with data from past changes by using data-mining methods. In the work of Grieco et al. [47], the aim is to group past EC according to their similarity. For the classification of the CRs into different groups, the description texts of the CRs are analyzed automatically by using a text-clustering algorithm. Arnarsson et al. [48] aim in their work to identify relevant past EC by using search engine methods where a search query consisting of single words or entire change documents can be entered. A model using document embedding (cf. Section 2.4) is used to search for similar past EC. Do [11] presents a two-stage approach to find matching experts for new EC. In the first stage, similar EC are selected via an automated analysis of past EC. In the second stage, suitable experts for the EC found in the first stage are identified via automated social network analysis. An operational product data management database is used as the data source for the analyses. Key persons or experts can be identified who communicate more frequently with other actors and have access to more information and resources. With this procedure, experts can be derived for new ECs who have taken on key roles in past similar ECs.

3.2. Shortcomings

The review of the available preliminary work in the areas of data-based approaches in TCM and approaches for SI in TCM reveals several research gaps that are addressed by this contribution. Non-data-based approaches for SI, such as those presented by Koch [4] and Sippl et al. [37], provide support for change-specific selection of relevant stakeholders based on the systematization of change characteristics. These approaches are characterized on the one hand by the high degree of abstraction, which does not allow the details of the change to be considered, and on the other hand by their manual character. As a result, the existing approaches only allow imprecise statements to be made with regard to affected stakeholders, which is also associated with a high level of manual effort for each identification. These approaches, therefore, do not provide the flexibility to efficiently account for the high variety of changes. Of the multitude of data-based approaches in change management (e.g., [9,10,39]), only Do [11] focuses on the change-specific prediction of relevant coordinating stakeholders for EC. However, this approach is limited to the utilization of change process data and therefore does not allow an in-depth analysis and exploitation of the contextual information in the change data. Furthermore, the impact on manufacturing stakeholders, such as assembly or logistics, is not considered. These data-based approaches for SI as

well as the existing approaches for similarity assessment of EC [47,48] are limited to the use of textual data and do not take into account potential additional data features collected during the change process (e.g., as the change costs) or the knowledge of affected stakeholders. However, according to Arnarsson et al. [48], the integration of relevant additional information offers the potential to increase model qualities. An efficient integration of the expertise of involved stakeholders to support data-based models has not been considered so far. The main objective of this contribution is therefore, the introduction of a procedure for data-based SI in TCM based on the approach of text classification, which includes all steps from the initial data selection to the improvement of model quality by the integration of expert knowledge. The approach is intended to serve manufacturing companies as a reference for the efficient introduction of classification models for change-specific prediction of relevant stakeholders. This should relieve coordinating employees in change management and increase the quality of SI. As a result, relevant stakeholders can be identified faster and automatically, allowing more time for planning and implementation. The intended approach should free employees completely from the task of SI. In addition, less relevant stakeholders are overseen, which reduces the risk of not considering relevant change impacts. This prevents possible negative consequences in change planning and implementation, such as additional costs, delays, or production stops.

4. Requirements

The requirements for the approach presented in this contribution were derived from the literature in the fields of change management and industrial practice. The literature-based requirements were derived primarily from Hamraz et al. [49], whose contribution introduces general requirements for ECM methods based on an extensive literature study. Further requirements were described based on the work of Wickel [9], as this work adds the perspective of data-based approaches in change management. In order to enable a later industrial application, the application company was consulted in several meetings to define and review the requirements. In the end, 11 requirements were specified, which are categorized as "data analysis requirements" and "method-oriented requirements". The first category summarizes requirements that relate to the creation and use of the text classification model for data analysis. The second group contains requirements for a generic procedure to introduce and improve text classification models for SI. Table 1 presents an overview of the defined requirements.

Table 1. Overview of requirements.

R1 Data Analysis Requirements	
R1.1	Open source software
R1.2	Semantic text contexts
R1.3	Quantitative evaluation of machine data analysis
R1.4	Prerequisite for automated analysis of new change requests
R1.5	Performance of classifier
R2 Method-Oriented Requirements	
R2.1	Availability of information for model building
R2.2	Procedure for text preprocessing
R2.3	Expert knowledge
R2.4	Objective data selection
R2.5	Industrial applicability
R2.6	Adaptability

It shall be ensured that the software used is not commercial so that it can be replicated by anyone with as few obstacles as possible. Open-source software offers the advantages of easy availability and transparency [50] regarding the code, resulting in easy customizability (R1.1). Data analysis should be performed by using textual contextual analysis rather than based on word frequencies because words and phrases are highly context-dependent

(R1.2). The analysis of the results shall be done by using objective, appropriate metrics in order to select the objectively best model and compare it with other approaches (R1.3). In principle, the automation of SI should be feasible for new CRs to provide not only a potential model but an approach to operationally support TCM (R1.4). The performance quality in identifying relevant CRs shall be at least equal to the performance on expert basis (R1.5).

The accessibility of required data for stakeholder analysis should be taken into account (R2.1). The preprocessing of the textual information for use in the word embeddings and in the HAN should be represented by a defined and reproducible procedure (R2.2). To complement the purely data-driven model with knowledge of the relevant employees, expert knowledge shall be derived and included within the methodology (R2.3). The objective selection of additional change data attributes should be supported to counter the curse of dimensionality and avoid unnecessarily long periods of model training. Therefore, appropriate procedures should be introduced to identify the most relevant data attributes (R2.4). The applicability of the approach in industry and thus in the practical environment shall be assured (R2.5). The adaptability to other business units or companies, and thus the adaptability of the developed approach, should be ensured (R2.6). Section 7 presents the requirements in terms of their fulfillment by the developed methodology.

5. Methodology for Data-Based SI

Figure 4 shows the basic steps of the developed methodology. Four superordinate steps are carried out, which cover different sub-aspects of the methodology. After the initial data selection and preparation follows the central component, the establishment of the text-classification model, which is explained in Section 2.3. This is used to classify CRs in terms of their relevance to individual stakeholders (business units). The text-classification model can only be applied for individual business units and, therefore, must be recreated for every relevant business unit. Therefore, the developed methodology describes the generic procedure for creating classification models for SI of different business units and the individual steps of the methodology were furthermore oriented toward the phases of the cross-industry standard process for data mining (CRISP-DM) according to Shearer [51] as a methodological basis. Accordingly, the methodological step described in Section 5.1 is oriented toward the phases of data understanding with regard to the CRISP-DM, and the methodological steps described in Sections 5.2 and 5.3 are oriented toward the phase of data preparation. Section 5.4 corresponds to the modelling and evaluation phase. An iterative procedure similar to the iterative procedure of the CRISP-DM is used for training of the models. The integration of expert knowledge presented in Section 5.5 is oriented toward the deployment phase. After the model has been successfully trained, the final step of the methodology focuses on the integration of expert knowledge to improve model quality.
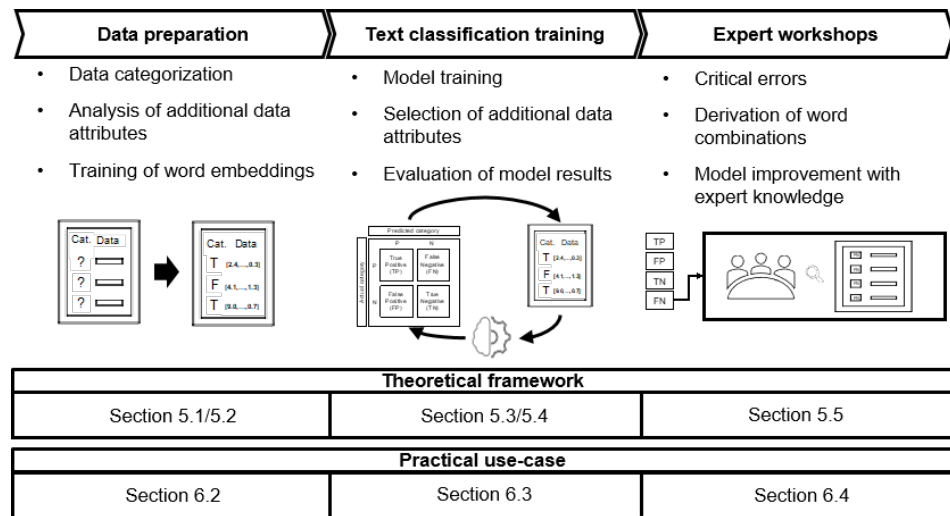
Figure 4. Overview of methodology.

5.1. Data Preparation

The basis for the data-driven identification of stakeholders in change management is a database of past CRs. Within the CRs, different change-related information is usually provided by the change requestor. The essential data for the text classification model are textual information such as the title, problem description, proposed solution, benefits, and comments of the change requestor, structured per CR. In addition, there is further data that can be collected during change creation, such as information about the change cause or a risk evaluation. Table 2 shows which data are collected in the creation of CRs in industry [4,12,14] and which of these are mandatory and which are optional for the approach presented in this contribution. Appendix A provides a more detailed overview of the features of the dataset used.

Table 2. Overview of collected data in TCM.

Data	Description	Exemplary Data Attributes
Mandatory	Textual change data	Title, problem description, solution description, benefit description, comments, …
Optional	Additional data attributes	Reasoncode, change type, change trigger, module group, construction group, variants increase, …

In order to enable the training of the classification model, the database must contain the relevance of the past CRs to the business units. Accordingly, the problem under consideration is a dichotomous classification of CRs as either "relevant to the business unit" or "not relevant to the business unit". This information can be obtained directly from an analysis of change process data. By analyzing the activity logs generated in digital processes, it is possible to determine in detail which employees and business units were involved in which step of the process. Alternatively, by considering additional features, the relevance of stakeholders can also be determined retrospectively. For example, if a feature contains the "cost relevance for logistics", this information can be used to classify the CRs concerning its relevance for logistics. However, these conditions are company-dependent and must be determined in close collaboration with change management experts. The resulting uncertainty of the categorization should be evaluated critically.

5.2. Transformation of Textual Data into Numerical Data

The textual information stored in the database has to be converted into numerical data to make it exploitable for the text classification model. As explained in Section 2.4, for the creation of numerical vector representations of the words, the Word2Vec according to Mikolov et al. [34] is used. This approach was chosen because the resulting word vectors contain contextual information and do not represent a word in isolation. Moreover, Yang et al. [31] used this approach to create the word vectors in the implementation of their HAN. For the implementation of the Word2Vec in Python, the Word2Vec model implemented by Řehůřek [52] was applied. Prior to the implementation, the hyperparameters "size of the word vector" and "minimum occurrence of words" have to be determined based on the available text corpus. The size of the word vectors determines how many entries a vector has. The minimum number of word occurrences ensures that words with a low frequency are not represented by a word vector. The context of low-frequency words cannot be clearly identified as only a very limited number of observations is provided in the data. Therefore, the resulting word vectors cannot be mapped clearly to the context of other words.

5.3. Feature Selection

The additional data available for the changes can complement the textual information with relevant information for the identification of affected stakeholders. The relevance of the data should be evaluated in advance, and non-relevant features should be excluded because training the text classification model is a time-consuming task. The assessment of relevance provides the basis for the selection of actually relevant data for the model training without risking a major loss of model quality and negative effects induced by the curse of dimensionality [53]. For feature selection, the correlation values between the available input and output features are used as a reference. This approach was chosen because of its simplicity, good interpretability, and frequent application [54]. The identification of relevant input features is therefore based on the analysis of the correlation between the available input features and the dichotomous target variable, distinguishing between continuous and categorical data. Continuous data are characterized by an infinite number of different numerical values and categorical data by a fixed number of possible defined values.

The correlation values are determined for all possible input features and thus serve as an indication for their potential added value and, therefore, for an integration into the classification model. A definition of which correlation analyses to use is not generally specified in this paper because it depends on whether continuous data or categorical data are available and whether the categorical data are nominal or ordinal data. Accordingly, Section 6.2 describes the approach adopted in this publication. For further possible correlation analyses, which can be applied in other constellations of input features, please refer to Bortz and Schuster [55].

5.4. Training of Machine Learning Models

After categorizing the CRs, generating training word embeddings, and analyzing the additional data, text-classification models are trained and then evaluated in order to build a target-oriented model. Within the training, an iterative procedure is used. At first, a model is trained, which only considers textual information as an input. The textual information provides the basis for the classification and is therefore integrated into all trained models because textual descriptions are the most precise way to understand the details of the actual change. Based on this basic model, several models are then trained with different combinations of additional features to identify the most successful classification models.

An analysis of the additional data individually, without the text data is not performed because the additional data consists only of single sentences without context. For each model training, the data used must be preprocessed and converted into numerical data. For this purpose, decimal numbers, punctuation, and special characters are removed, abbreviations are replaced with spelled-out words, umlauts are converted, all words are

lowercased, numbers between 0 and 9 are converted into words, and finally tokenization is performed. These steps, as shown in Uysal and Gunal [56], Camacho-Collados and Pilehvar [57], or Kao and Poteet [58], are common steps in preparing text for data analysis. Afterward, each tokenized word is converted to its word vector, which serves as input variables into the HAN. After training, the models can be evaluated and compared by using statistical metrics. By performing this training cycle for different configurations of the data attributes, different models with varying classification quality are generated. The objective is to select the most suitable model for each use case by comparing the metrics of each model. For the training and evaluation of the models, the database data is split into a training dataset, a validation dataset, and a test dataset. The same data split is used for each model training so that the trained models can be compared. The training dataset consists of 70% of the data. The validation dataset and the test data set each consist of 15%. The splitting of the data is done in order to train the text-classification model on the largest possible database and then to objectively evaluate the quality of the model by using unseen data. A trained text-classification model is to decide for new CRs whether they are "relevant to business unit" or "not relevant to business unit". Within the training cycle, the test dataset is used to test whether the trained model delivers target-oriented results. Subsequently, the results are compared with the categories derived from Section 5.1. From this, the number of true positive (TP), true negative (TN), false positive (FP), and false negative (FN) results can be derived and used to evaluate the model quality based on statistical metrics.

In general, the choice of the appropriate evaluation metrics for machine learning models strongly depends on the considered use case. The relevance of incorrect and correct classification varies significantly. In the case of change management, the FN classifications, in particular, are crucial, as they address cases wherein relevant stakeholders are missed. This means that the expertise and assessment of the employees cannot be utilized in the planning and implementation of the change, which increases the risk of unplanned change impacts. FP cases, in contrast, are considered significantly less critical, despite the waste of resources due to unnecessary efforts, as in these cases, there is often only a brief review by the contacted stakeholder, and the non-relevance is quickly recognizable. In addition, it became clear in discussions with industry experts that overall a tendency of the model to classify as "relevant" is to be preferred. In this way, the experts are contacted in uncertain cases and can thus perform a review of the change. Therefore, an optimal balance between FN and FP cases must be considered during the identification of the most suitable model. We have therefore chosen recall as the central parameter of model quality, although it should not be considered in isolation in order to avoid erroneous conclusions. For this purpose, precision, specificity, balanced accuracy, and F1 score are used in addition to the recall. For each trained model, it is to be compared how these metrics change while the recall, as the target variable, is maximized. With this procedure, models should be avoided that show a high recall but at the same time a very high number of FP cases.

5.5. Integration of Expert Knowledge

After the most suitable classification model has been identified, the model quality is to be further improved by integrating expert knowledge. For this purpose, workshops are held with relevant employees on the basis of the classification results obtained. In these workshops, the FN classification cases are analyzed by the employees who were actually affected. For each workshop, lists of FN results are prepared (e.g., in an Excel file) and then reviewed with an employee who would be affected by the changes. Workshops are conducted with individual employees and the duration depends on how many FN cases are to be analyzed, although the objective is to have as many as possible. Based on the use case described in Section 5 in this contribution, it can be assumed that a duration of one minute per FN case should be calculated.

The information on which employees are relevant is obtained from past change data, as described in Section 5.1. This traceability is essential as otherwise employees give their

assessment without being aware of the background and details of the change. Changes that cannot be clearly assigned, e.g., due to reorganizations or changes in the workforce, are not considered in this step. In these workshops, the employees use the textual information of the change under consideration and indicate, based on their expert knowledge, which words or sentence components would have implied the relevance of their business unit. In this way, a vocabulary is derived, which includes words and word combinations that indicate the relevance of the business unit under consideration and therefore offers the potential to support the classification model.

The identified word combinations from all experts are then sorted into categories summarizing terms with similar or synonymous meanings. For example, within a workshop, the word combinations "adjustment steering wheel", "adding screw", and "modification side member" are identified as relevant word combinations. The category "Modification" can then be formed from the terms "addition", "modification", and "adaptation". In contrast, the terms "steering wheel", "screw" and "side member" are clustered as "Product Parts". The terms inside these categories are then automatically combined to obtain all possible word combinations. With the two exemplary lists "Modification" and "Components" a new list of word combinations is created, in which "adjustment steering wheel", "adjustment screw" or "adjustment side member", and "add steering wheel" are contained. However, when creating the combinations, the words within a category are not combined because it is assumed that these combinations (e.g.,"adjustment add") do not offer any added value for the classification due to the similarity of their meanings. The formation of this multitude of word combinations and the subsequent evaluation of their added value enables the identification and inclusion of actually relevant word combinations in the classification model. These workshop-specific lists are then merged per derived category of terms of similar meaning, and duplicates are deleted. For each individual CR, it is then checked which word combinations it contains. The word combinations do not have to occur directly next to each other in the text but may also be distributed throughout the text. The word combinations that do not appear in any CRs are then removed. The remaining word combinations are then combined with the results of the previously trained text-classification model. For each word combination, the effect on the overall model quality of relabeling the text classification model from "not relevant" to "relevant" based on their occurrence is examined. By focusing on the cases recognized as "not relevant" (FN or TN), either a change from FN to TP or from TN to FP can be achieved. It is checked for each word combination whether a reduction in FN is achieved without at the same time disproportionately increasing the number of FP. For this purpose, the change in statistical metrics caused by the result adjustment per word combination is calculated. Only the word combinations that increase the recall and at the same time do not decrease the F1 score are included in the list of available word combinations. All other word combinations are discarded. The actual prediction of the relevant stakeholders is finally made by a two-step approach. The trained model classifies each CR as "relevant" or "not relevant" for the considered business unit. In all cases classified as "not relevant", it is checked whether one or more word combinations occur. If so, the label is changed to "relevant" for the business unit. Thus, the list of word combinations can be used to support the text classification model when analyzing new CRs to reduce the number of FN and thereby improve the overall model result.

6. Industrial Application and Evaluation

In order to evaluate the developed methodology with regard to the quality of the results and the applicability in industrial practice, an application case was carried out in a German automotive company. The insights gained are used to assess the fulfillment of requirements (cf. Section 4) and the fulfillment of the need for action derived from the state of the art (cf. Section 3.2). In the course of this use case, the developed methodology was completely applied and evaluated on the basis of the model quality achieved and the experience gained during its implementation. The resulting model quality throughout the

different process steps is shown in Figure 5 at the end of this section. Here the results of the test data set are shown. The relevant results from Figure 5 are referenced in each section.
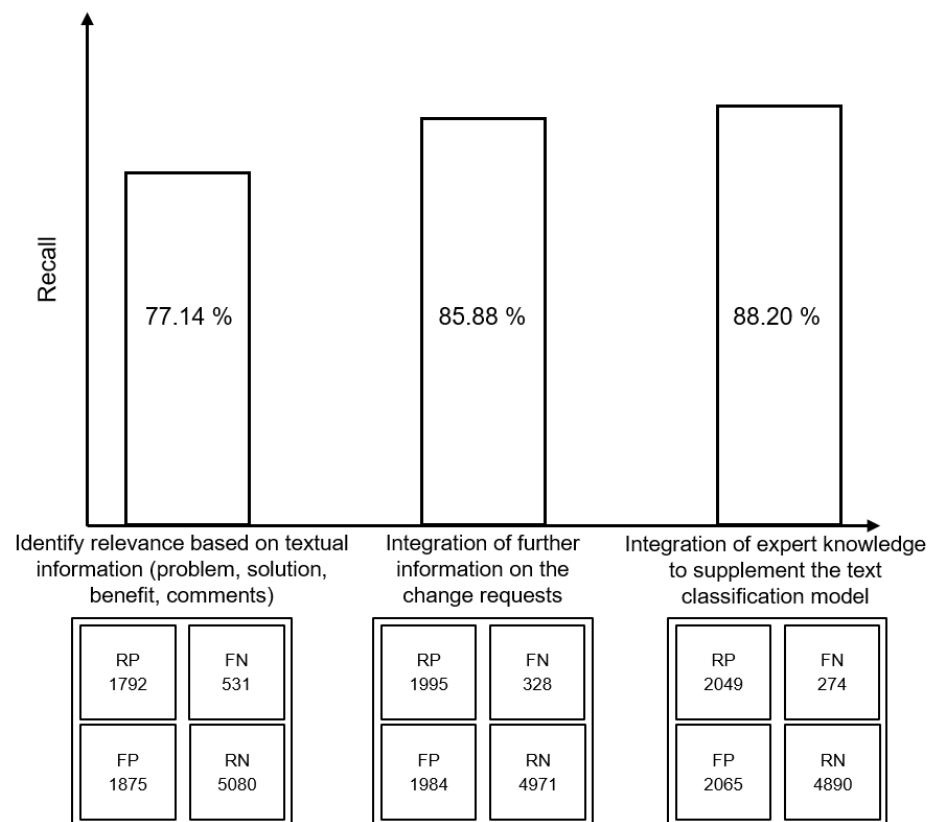

Figure 5. Overview of results.

6.1. Use Case "Assembly"

The goal of this use case was to predict the relevance of individual CRs for assembly. Within this business unit, employees are divided into different areas. Each area is responsible for the assembly processes of different components, such as the body, the interior, or the doors. In the change process, employees in the assembly business unit must be informed of a CR and evaluate it as soon as a CR affects assembly processes and leads, for example, to necessary modifications of the assembly instructions or the elimination of assembly steps. The previously planned assembly processes must then be adapted by the employees to the new boundary conditions. In current industrial practice, the change requestor assesses the relevance of the assembly business unit and asks the relevant employees to analyze and evaluate the planned change. The database used contained 71,506 CRs submitted between January 2018 and April 2021. The CRs include all mandatory textual data (e.g., title, description) and additional optional categorical data attributes (e.g., affected manufacturing technologies), as described in Section 5.1. The use case data and their preparation for use in the HAN are described in detail in the following section.

6.2. Data Preparation

The textual information on the change title, the problem description, the solution description, the description of the benefits of the CRs, and the comments of the change requestor were used as a basic input for the development of the classification model. This textual information was prepared for word-embedding training by using the Word2Vec approach, as described in Section 5.4. In order to remove words with low frequency, words that occur less than five times in the entire text corpus were removed. The database prepared in this way then served as input for the training of word-embedding, which aimed at generating an individual vector for each word in the text corpus. The following

hyperparameters were defined: the size of the word vectors was set to 150, and the CBOW approach was chosen as the model architecture. CBOW was chosen as the Word2Vec model because of its faster training times. Skip-Gram, the alternative option, is better at representing rare words [34]. However, this is not considered relevant for the application of this contribution, as it is more important to be able to represent frequently used words in different contexts in the best possible way. The training of the word embeddings thus resulted in a vector representation for 71,475 words, each containing 150 entries based on the defined set of hyperparameters. These entries represent the word depending on its context and are independent of individual use cases of the methodology (cf. Section 2.4). In addition to the textual information, 13 nominal categorical attributes were provided by the company, such as the design group, the module group, the change type, the reason code, the bill of materials, the cost impact on assembly, and whether an assembly employee was notified due to impact of a change on assembly. The latter two attributes were not used as an input for model generation. However, these attributes were used to a posteriori label all CRs as "relevant" (24.73%) or "not relevant" (75.27%) for assembly. The correlation between the additional attributes and the relevance for the assembly business unit was then assessed by applying the chi-square test, as this test can be used to assess the dependence between nominal categorical attributes and a dichotomously distributed target variable [55]. In order to be able to assess the validity of this approach, individual models were trained, which contained the textual data in combination with all individual additional information as inputs. The exact training specifications are described in Section 6.4. Figure 6 shows the results of the chi-square test (blue) combined with the resulting model qualities (grey) represented by the value of the recall. In this figure, the additional data attributes are sorted by the calculated chi-square values.
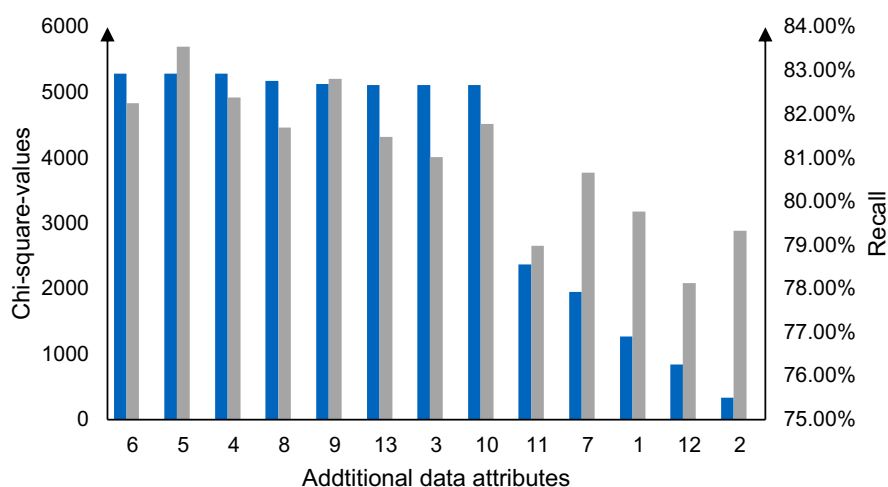
Figure 6. Results of chi-square test and resulting model quality with single additional data.

Figure 6 shows that based on the chi-square test, eight of the thirteen attributes have a comparatively high correlation with the target variable. These attributes also reach a higher recall than the five with a comparable low correlation value. These results indicate that the values of the chi-square test give a good indication of the added value for the classification model and can therefore be used for the initial filtering of the input data.

6.3. Model Training

Based on the derived target variable, the trained word embeddings, and the analysis of the additional data attributes, several text classification models were trained. Before that, however, the hyperparameters of the HAN had to be selected based on the available text corpus. The number of sentences that can be processed was chosen as 20, and the words per sentence were set to 50 (cf. Figure 1). Thus, documents with a maximum of

1000 words could be analyzed. If this limit is exceeded, the document is split evenly until all subdocuments can be processed. For each of these documents, a classification decision was made by the HAN, and if one part of the document was considered "relevant", the whole document was classified as "relevant". First, a classification model was trained, which received only textual information as input and thus served as a reference for the relevance of the additional attributes. Subsequently, further models were trained, which combined the textual information with additional individual attributes. After data preselection, different combinations of the data attributes were evaluated by training separate models. The feature configurations were chosen based on the findings of the chi-square test so that the best results could be expected accordingly.

Figure 5 shows that with the help of the additional data attributes, a significant improvement in model quality could be achieved compared to the model with only the textual information as input. At the same time, the confusion matrices imply that the number of FP cases increases only slightly. With the use of the additional data attributes, the number of FN was reduced by 205, from 531 to 326. At the same time, the number of FP was increased from 1875 to 1984, by 109. Accordingly, an 8.74% improvement in recall was achieved, a 1.27% increase in Precision, and a 1.57% reduction in Specificity. Concurrently, the Balanced Accuracy value increased by 3.59% and the F1 score by 3.48%.

6.4. Expert Knowledge

As described in the developed methodology, the trained models were further improved by integrating expert knowledge. Therefore, 14 employees from different areas of the assembly business unit were asked to specify relevant words and word combinations which indicate the relevance of their business unit. The workshops lasted about one hour. A total of 667 words were named by the experts, which were then grouped into the five superordinate categories "modification", "product component", "assembly", "assembly part", and "plant" (cf. Section 5.5), which unify words with synonymous or similar meanings. From this initial collection, 7930 possible word combinations were derived, which were then automatically evaluated with regard to their potential for increasing the model quality and, at the same time, reducing the number of FN cases. The evaluation resulted in a model-specific list of 630 relevant word combinations whose integration had a positive effect on the recall without reducing the F1 score. Figure 5 shows the impact of integrating expert knowledge, with all results of the text-classification model being considered "relevant to the business area" if one or more of the 630 relevant word combinations occur. The results in Figure 6 show that the two-stage model, which uses the derived word combinations as well as additional data attributes, brings a further improvement in model quality. The word combinations changed the classification results for 135 of 9278 changes. As a result, 81 TN became FP, and 54 FN became TP. Accordingly, a 2.32% improvement in recall was achieved, with a 0.33% reduction in Precision and a 1.16% reduction in Specificity. Concurrently, the Balanced Accuracy value increased by 0.58% and the F1 score by 0.35%.

7. Discussion

Based on the practical application, this section discusses the results versus their impact on the practical use in TCM. In addition, the results obtained are used to evaluate the requirements specified in Section 3. The fulfillment of the requirements is evaluated based on the knowledge gained in the application case, but also on the general characteristics of the developed approach.

7.1. Implications for Industrial Practice

The introduction and application of the methodology are associated with various expenses that must be taken into account. On the one hand, the time required for the training cycles of the models must be taken into account. One training cycle took about 2 h. In the practical use case, 35 models with different configurations of additional data attributes were trained, resulting in a total amount of time required for the training cycles

of about 70 h. On the other hand, the preparation and implementation, as well as the evaluation of the results of the workshops for the collection of expert knowledge, should be considered. The time required for each workshop was one hour of preparation, one hour of execution, and two hours of post-processing of the findings. In total, this resulted in a time expenditure of 70 h across all workshops carried out together. It should be noted that this effort to create and improve the model is only to be undertaken initially and that the company benefits from a long-term perspective by reducing the manual analysis of CRs. However, these efforts can only be justified if an improvement of the model quality can be achieved and therefore have an added value for operational TCM.

In the considered use case, the manual SI by the change requestor in the period from January 2018 to April 2021 reached a recall of 81.32%, with 3303 FN in total or in terms of workdays 23 FN per week on average. Figure 5 shows that higher values for the recall were achieved through the training cycles with different additional data attributes and the selection of the most suitable model. In addition, the integration of the word combinations identified in the workshops with business unit employees led to an increase in recall. Overall, the best model would result in 2112 FN, a total of 1191 fewer than compared to the change requestor equal to a reduction of 8 FN cases per week on average. Compared to the model with just textual data as input, the HAN with additional data attributes can avoid 11 FN per week compared to the HAN without additional data attributes. The additional use of the expert knowledge in the form of word combinations can avoid 3 FN per week. Overall, the application of the method can thus avoid an average of 14 FN per week. Overall, the use of machine data analysis with the aid of HAN and the methodology can reduce the risk of FN and thus of unexpected change impacts.

The overall quality of the results is higher than the benchmark set by manual SI that the task of SI can be completely taken over by the algorithm, which either relieves employees or enables a reduction in the number of employees required for change management. The approach developed is thus distinguished from manual approaches to SI by its complete automation. After the initial efforts for data extraction, data preparation as well as model building and improvement, this allows the identification of affected stakeholders to be carried out almost effortlessly, which can lead to significant economic savings in the long term. Due to the in-depth analysis of the textual information by the HAN, which allows the consideration of all details and nuances of the changes to be analyzed, it can be assumed that the developed approach allows more precise statements about affected stakeholders than all comparable data-based approaches. However, a quantitative comparison cannot be made at this point due to the uniqueness of the use cases in this contribution as well as in the previous contributions. However, a disadvantage of the developed approach is the need for a comprehensive database and thus the need for comparable historical changes. As a result, all data-based approaches are likely to produce imprecise results when analyzing particularly novel changes that have not yet occurred in this form. As a result, it must be concluded that a suitable data-based or non-data-based approach should be selected to support SI in companies, depending on the type of change.

7.2. Fulfillment of Requirements

This section discusses and evaluates the requirements defined in Section 4, grouped into the two categories of "data analysis requirements" and "method-oriented requirements". An overview of the fulfillment of the requirements is presented in Table 3. With regard to the "data analysis requirements", it can be stated that only freely available software, in particular Python libraries, was used to support the developed tool. Therefore, R1.1 ("open-source software") is rated as completely fulfilled. R1.2 ("semantic text contexts") is also considered to be fulfilled since the HAN approach used in combination with word embeddings is focused on recognizing and mapping the hierarchical structure within documents and thus semantic contexts. In contrast, R1.3 ("quantitative evaluation") is considered only partially fulfilled. Although the quality of the results of the developed approach was evaluated on the basis of several established metrics (e.g., recall, precision,

and F1 score), the model performance could not be compared in detail with the expert-based evaluation (e.g., detection of false positives). However, requirement R1.4 ("prerequisite for automated analysis of new change requests") is considered to be fulfilled. If the procedure is applied as described, the result is a classification model that evaluates new change requests by using historical data and expert knowledge. For the final application in industrial practice, suitable interfaces and data connectors should be provided in the next step. R1.5 ("performance quality classifier") can be considered fulfilled with respect to the particularly relevant metric of recall. It can be concluded that the developed approach can assess change requests in slightly better quality and significantly faster than an expert.

The first "method-orientented" requirement R2.1 ("availability of information for model building") can be considered fulfilled. Existing preliminary work proves that a large majority of companies record both textual and additional change data. Thus, the broad industrial availability of the required information can be assumed. However, differences are to be expected with regard to the type and quality of the additional information collected, which is why our approach also considers data preselection. Therefore, R2.2 ("procedure for text preprocessing") and R2.4 ("objective data selection") can also be positively evaluated. The approach contains clear descriptions of the necessary steps for data preprocessing. The review or elimination of data-quality problems was not the focus of the present contribution, which is why reference should be made to established literature in this regard. The use of expert knowledge in change classification was described in detail and successfully applied twice. Therefore, R2.3 is also considered fulfilled. The obtained quality of results, as well as the efforts to implement the approach, show that industrial applicability (R2.5) was sufficiently considered during the development of the approach. Finally, R2.6 ("adaptability") can also be considered as completely fulfilled, as the approach supports the diversity of possible databases and companies by systematically supporting data preparation and selection, as well as the collection and integration of expert knowledge.

Table 3. Overview of the fulfillment of requirements.

R1 Data Analysis Requirements		Fulfillment
R1.1	Open source software	fulfilled
R1.2	Semantic text contexts	fulfilled
R1.3	Quantitative evaluation of machine data analysis	partly fulfilled
R1.4	Prerequisite for automated analysis of new change requests	fulfilled
R1.5	Performance of classifier	fulfilled
R2 Method-Oriented Requirements		
R2.1	Availability of information for model building	fulfilled
R2.2	Procedure for text preprocessing	fulfilled
R2.3	Expert knowledge	fulfilled
R2.4	Objective data selection	fulfilled
R2.5	Industrial applicability	fulfilled
R2.6	Adaptability	fulfilled

8. Conclusions

The complex environment in which manufacturing companies operate places high demands on ECM and MCM. Influencing factors, such as rapidly changing customer requirements, global crises, or changing legal requirements, cause a high and increasing number of permanently necessary technical changes. Their efficient and effective implementation is seen as an important factor in maintaining long-term competitiveness. During the selection and planning of changes, it is essential to identify all relevant stakeholders in the company and in the company's environment and to include them in the change process in order to be able to consider their assessments in the decision-making process, e.g., with regard to expected expenses, risks, or also advantages. In current industrial practice, SI is often based exclusively on the knowledge and experience of individual employees, which has a high potential for error, especially given the high number and variety of changes.

For this reason, a methodology was introduced in the context of this contribution, which is intended to enable the data-based identification of stakeholders in TCM and describes the necessary procedure, from data preparation to integration of expert knowledge to increase model quality. The developed approach was successfully applied and evaluated in cooperation with a German automotive manufacturer. The results show that the achieved model quality is above the performance of a manual SI. By using the developed approach, the decision about relevant stakeholders can be automated and thus performed faster and with higher quality. Despite this very strong indication of industrial applicability and practicability, the approach should also be applied to other companies from different industries. This would allow detailed statements about the necessary effort and limitations. However, it can already be stated that the presented approach requires an extensive database and thus a high number of past changes. Although the absolute majority of companies record the required data in the company [13], there is always the risk of an insufficient number of changes, particularly in the production of long-running products that change only infrequently. This must be considered on a case-by-case basis and therefore depends heavily on the companies' business model, products, and environment. Furthermore, the implementation of the approach requires employees with skills in artificial intelligence and NLP, which are often not available especially in smaller companies. These skills are necessary to successfully guide the creation as well as the interpretation and evaluation of the classification models. The mentioned limitations restrict the number of companies that can potentially introduce the approach for data-based SI. Additionally, it should be noted that the existing approach to data preparation does not take into account organizational changes in the company during the period of data collection. For example, a splitting or merging of departments or changes in responsibility profiles are not taken into account, which can lead to imprecise and incorrect stakeholder predictions. In future work, this problem should therefore be investigated in detail, but with permanent consideration of the cost–benefit ratio. With regard to the validity of the results, it must also be critically considered that only production-related departments were considered in the application case. For example, there was no consideration of the legal or design departments. It therefore remains to be seen whether it is possible to identify these departments as relevant stakeholders with a similar level of quality. Further research work should also focus on reducing the effort required for the preparation and application of the approach. Currently, an individual classification model has to be developed for each business unit. By using a multi-class classification model, this effort could be significantly reduced, whereby the initial effort must be weighed against the achieved model quality on a permanent basis. Furthermore, different techniques of NLP and their effect on the model qualities should be tested. For example, alternatives to the approach of word embeddings such as BERT [59] could allow an improved representation of the textual context. Moreover, the developed approach only focuses on the initial identification of relevant stakeholders. This means that it is currently not possible to dynamically adapt stakeholder predictions on the basis of the actual process flow. The use of pattern-recognition techniques could be a promising approach to tackle this problem. Finally, it can be concluded that the developed approach illustrates the still existing potential of data-driven approaches in MCM and ECM. It is important to consider which information is generated in which process phases and which approaches are suitable for data-based support of change management. The further development and systematic integration into industrial change management processes of these approaches can represent an important contribution to the efficient handling of MC and EC in the future.

Author Contributions: Methodology: F.S.; Writing original draft: F.S. and R.M.; Writing—Review and Editing: C.P.G., S.D. and G.R. All authors have read and agreed to the published version of the manuscript.

Appendix A

	Aspect 1	Aspect 2	Aspect 3	Aspect 4
	AND			
OR	Änderungsmanagement	Engineering Change	Manufacturing Change	Stakeholder
	Change Management	Engineering Change Management	Manufacturing Change Management	Identifikation
	integrated change management	integrated design	manufactur* planning	identificat*
	integriertes Änderungsmanagement	product development project	assembly planning	identifizier*
	Change Propagation	Produktentwicklung	continuous factory planning	Mitarbeiter
	Änderungsausbreitung	integriertes design	Montageplanung	Personal
	modificat*			Betroffene*
	adapt*			Interessensgruppe
	transformat*			group of interest
	reconfiguration			affect*
	Neukonfiguration			Expert*
				stakeholder identification
				planer
				planner
				beteiligt*
				involve
				Team
				participant
				erfass*
				capture
				regist*
				erkenn*
				detect*
				ermittel*
				determin*
				ascertain

Figure A1. Search terms of the literature analysis. * represents any possible word endings

	Giffin et al. 2009	Kocar & Akgunduz 2010	Habhouba et al. 2011	Pasquale & Weck 2011	Sharafi 2013	Malak & Aurich 2013	Do 2015	Grieco et al. 2017	Kattner et al. 2018	Do 2018	Arnarsson et al. 2019	Pan & Stark 2020	Schuh et al. 2021
Data analysis													
Use of semantic text contexts	○	○	○	○	●	○	○	◐	○	○	●	●	○
Quantitative evaluation machine data analysis	◐	○	○	◐	◐	○	◐	●	●	◐	●	●	○
Requirements for automated analysis of new change requests	○	●	●	○	○	◐	●	●	○	●	●	◐	○
Requirements for self-learning system	○	◐	◐	○	○	◐	○	◐	○	◐	○	◐	○
Methodical approach													
Industrial applicability	●	○	◐	●	●	●	○	●	●	◐	●	●	○
Procedure for data pre-processing	○	○	○	○	○	○	◐	●	○	◐	●	●	○
Objective data selection	○	○	○	○	◐	○	◐	○	◐	◐	○	◐	○
Use of expert knowledge	○	○	○	○	○	○	○	○	○	○	◐	○	○
Cross-divisional analysis	○	○	◐	○	●	○	○	○	○	◐	●	◐	●
Identification affected stakeholders	○	○	◐	◐	◐	○	○	○	○	◐	●	○	◐
Legend: ○ = not fulfilled; ◐ = partly fulfilled; ● = completely fulfilled													

Figure A2. Categorization of relevant preliminary work [10,11,38–48].

Feature-Name	Description	Example
Change-ID	Identification number for unique identification	1001
Creation Date	Date of change creation	01.01.2020
Title	Title of the change	Change of Dimensions
Problem description	Description of the problem that triggers the change	Difficulties with the installation of the component
Solution description	Description of the solution to the problem	Changing component size
Benefit description	Description of the benefits of the change	No problems with the installation due to size adjustment
Comment of the creator	Additional information from the creator	-
Rulebook	Description for project supplements	-
Reason code	Code why change is carried out	Reason_01
Derivate	List of derivatives affected	Dev15; Dev27;
Change type	Distinction as to which type of change is carried out	Type_6
Module	List of affected modules	Mod4;
PPG	Identification number of the Product-Process-Group	PPG301
Construction group	Construction group identifier	CG5
MMG	Node naming within module set structure	MMG300
KMG	Node naming within construction set framework	KMG100
Part number	Designation of affected item numbers	PN1022
Commodity	Buyer identifier	BI101
Position type	Nature of the change	3
Part number increase	Number of new part numbers created	0
Part number decrease	Number of discontinued previous part numbers	0
Part numbers quantity	Number of item numbers affected by the change part numbers	1
Cost impact	Cost impact Yes/No	Yes
Call for assessment	Call for assessment Yes/No	Yes

Figure A3. Description of the features in the dataset.

References

1. Mack, O.; Khare, A.; Kramer, A. *Managing in a VUCA World*, 1st ed.; Springer: Berlin/Heidelberg, Germany, 2015. [CrossRef]
2. Wonsak, I.; Bauer, H.; Sippl, F.; Reinhart, G. A scenario-based approach for translating strategic perspectives into input variables for production planning and control. *Procedia CIRP* **2021**, *104*, 429–434. [CrossRef]
3. Wiendahl, H.P.; ElMaraghy, H.A.; Nyhuis, P.; Zäh, M.F.; Wiendahl, H.H.; Duffie, N.; Brieke, M. Changeable Manufacturing-Classification, Design and Operation. *CIRP Ann.* **2007**, *56*, 783–809. [CrossRef]
4. Koch, J. Manufacturing Change Management—A Process-Based Approach for the Management of Manufacturing Changes. Ph.D. Thesis, Technische Universität München, Munchen, Germany, 2017.
5. Lindemann, U.; Reichwald, R. *Integriertes Änderungsmanagement*; Springer: Berlin, Germany, 1998.
6. Jarratt, T.A.W.; Eckert, C.M.; Caldwell, N.H.M.; Clarkson, P.J. Engineering change: An overview and perspective on the literature. *Res. Eng. Des.* **2011**, *22*, 103–124. [CrossRef]
7. Basse, F. *Gestaltung Eines Adaptiven Änderungssystems für Einen Beherrschten Serienhochlauf: Design of an Adaptive Engineering Change System for a Stable Series Ramp-Up*, 1st ed.; Produktionssystematik; Apprimus Verlag: Aachen, Germany, 2019; Volume 2019, Band 22.
8. Malak, R.C. Methode zur Softwarebasierten Planung Technischer Änderungen in der Produktion: Zugl.: Kaiserslautern, Techn. Univ. Produktionstechnische Berichte aus dem FBK. Ph.D. Thesis, Lehrstuhl für Fertigungstechnik und Betriebsorganisation Techn. Univ., Kaiserslautern, Germany, 2013.
9. Wickel, M.C. Änderungen Besser Managen—Eine Datenbasierte Methodik zur Analyse Technischer Änderungen. Ph.D. Thesis, Technische Universität München, Munchen, Germany, 2017.
10. Sharafi, A. *Knowledge Discovery in Databases*; Springer Fachmedien Wiesbaden: Wiesbaden, Germany, 2013. [CrossRef]
11. Do, N. Identifying experts for engineering changes using product data analytics. *Comput. Ind.* **2018**, *95*, 81–92. [CrossRef]
12. Sippl, F.; Schellhaas, L.; Bauer, H. Umfrage zum Änderungsmanagement in der Produktion Status quo, industrielle Anwendung der Änderungsauswirkungsanalyse und Stand der Digitalisierung. *Z. für Wirtsch. Fabr.* **2021**, *116*, 208–212. [CrossRef]
13. Koch, J.; Hofer, A. Änderungsmanagement in der Produktion: Herausforderungen und Anwendungen in der industriellen Praxis. *WT Werkstatttechnik* **2016**, *7/8*, 520–526. [CrossRef]
14. Wickel, M.C.; Lindemann, U. A retrospective analysis of engineering change orders to identify potential for future improvements. In Proceedings of the DS 81: Proceedings of NordDesign 2014, Espoo, Finland, 27–29 August 2014.
15. Freeman, R.E. *Strategic Management: A stakeholder Approach*, 1984th ed.; Cambridge University Press: Cambridge, UK, 2010.
16. Beam, C.; Specking, E.; Parnell, G.S.; Pohl, E.; Goerger, M.N.; Buchanan, J.P.; Gallarno, G.E. Best Practices for Stakeholder Engagement for Government R&D Organizations. *Eng. Manag. J.* **2022**, 1–20. [CrossRef]
17. Hayes, J. *The Theory and Practice of Change Management*, 4th ed.; Palgrave Macmillan: Basingstoke, UK, 2014.
18. Jarratt, T.; Eckert, C.; Clarkson, P. Engineering change. In *Design Process Improvement*; Clarkson, J., Eckert, C., Eds.; Springer: Berlin/Heidelberg, Germany, 2005; pp. 262–285.
19. Reinhart, G.; Pohl, J.; Schindler, S.; Rimpau, C. Cycle-oriented production structure monitoring. In Proceedings of the 3rd International Conference on Changeable, Agile, Reconfigurable and Virtual Production (CARV 2009), München, Germany, 5–7 October 2009; Zaeh, M.F., ElMaraghy, H.A., Eds.; Utz: München, Germany, 2009; pp. 693–701.

20. Malak, R.C.; Yang, X.; Aurich, J.C. Analysing and Planning of Engineering Changes in Manufacturing Systems. In Proceedings of the 44th CIRP Conference on Manufacturing Systems, Madison, WI, USA, 31 May–3 June 2011.

21. Rößing, M. Technische Änderungen in der Produktion—Vorgehensweise zur systematischen Initialisierung, Durchführung und Nachbereitung: Zugl.: Kaiserslautern, Techn. Univ. Ph.D. Thesis, Produktionstechnische Berichte aus dem FBK, Techn. Univ., Kaiserslautern, Germany, 2007.

22. Stanev, S.; Krappe, H.; Ola, H.A.; Georgoulias, K.; Papakostas, N.; Chryssolouris, G.; Ovtcharova, J. Efficient change management for the flexible production of the future. *J. Manuf. Technol. Manag.* **2008**, *19*, 712–726. [CrossRef]

23. ProSTEP iViP e.V. *Manufacturing Change Management (Recommendation): Management of Changes during Production*; ProSTEP iViP e.V.: Darmstadt, Germany, 2015.

24. Chowdhary, K.R. Natural Language Processing for Word Sense Disambiguation and Information Extraction. Ph.D. Thesis, Jai Narain Vyas University, Jodhpur, India, 2004.

25. Chowdhary, K.R. Natural Language Processing. In *Fundamentals of Artificial Intelligence*; Chowdhary, K.R., Ed.; Springer: New Delhi, India, 2020; pp. 603–649. [CrossRef]

26. Vijayarani, S.; Ilamathi, J.; Nithya. Preprocessing Techniques for Text Mining - An Overview. *Int. J. Comput. Sci. Commun. Networks* **2015**, *5*, 7–16.

27. Shah, K.; Patel, H.; Sanghvi, D.; Shah, M. A Comparative Analysis of Logistic Regression, Random Forest and KNN Models for the Text Classification. *Augment. Hum. Res.* **2020**, *5*, 1–16. [CrossRef]

28. Young, T.; Hazarika, D.; Poria, S.; Cambria, E. Recent Trends in Deep Learning Based Natural Language Processing. *IEEE Comput. Intell. Mag.* **2017**, *13*, 55–75. [CrossRef]

29. IBM Cloud Education. Natural Language Processing (NLP). *Fundam. Artif. Intell.* **2020**, 603–649. Available online: https://www.ibm.com/cloud/learn/natural-language-processing (accessed on 12 June 2022).

30. Chowdhary, K.R. (Ed.) *Fundamentals of Artificial Intelligence*; Springer: New Delhi, India, 2020. [CrossRef]

31. Yang, Z.; Yang, D.; Dyer, C.; He, X.; Smola, A.; Hovy, E. Hierarchical Attention Networks for Document Classification. In Proceedings of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, San Diego, CA, USA, 12–17 June 2016; pp. 1480–1489.

32. Nguyen,V. Hierarchical-Attention-Networks-Pytorch. 2019. Available online: https://github.com/uvipen/Hierarchical-attention-networks-pytorch (accessed on 12 June 2022)

33. Li, Y.; Yang, T. Word Embedding for Understanding Natural Language: A Survey. In *Guide to Big Data Applications*; Srinivasan, S., Ed.; Springer International Publishing: Cham, Switzerland, 2018; Volume 26, pp. 83–104. [CrossRef]

34. Mikolov, T.; Chen, K.; Corrado, G.; Dean, J. Efficient Estimation of Word Representations in Vector Space. *arXiv* **2013**, arXiv:1301.3781.

35. Webster, J.; Watson, R. Analyzing the Past to Prepare for the Future: Writing a Literature Review. *Manag. Inf. Syst. Q.* **2002**, *26*, 13–23.

36. VDA: ECM Recommendation Part 0 (ECM), VDA 4965—Part 0. Ed. by VDA Verband der Automobilindustrie; SASIG; ProSTEP iViP e.V. 2010-01. Available online: https://www.prostep.org/fileadmin/downloads/VDA_ECM_Recommendation_-_Part_0__ECM__V2.0.3.pdf (accessed on 12 June 2022).

37. Sippl, F.; Del Rio, B.; Reinhart, G. Approach for stakeholder identification in Manufacturing Change Management. *Procedia CIRP* **2022**, *106*, 191–196. [CrossRef]

38. Schuh, G.; Guetzlaff, A.; Sauermann, F.; Krug, M. Data-based improvement of engineering change impact analyses in manufacturing. *Procedia CIRP* **2021**, *99*, 580–585. [CrossRef]

39. Giffin, M.; de Weck, O.; Bounova, G.; Keller, R.; Eckert, C.; Clarkson, P.J. Change Propagation Analysis in Complex Technical Systems. *J. Mech. Des.* **2009**, *131*, 1–14. [CrossRef]

40. Pasqual, M.C.; de Weck, O.L. Multilayer network model for analysis and management of change propagation. *Res. Eng. Des.* **2012**, *23*, 305–328. [CrossRef]

41. Kattner, N.; Mehlstaeubl, J.; Becerril, L.; Lindemann, U. Data Analysis in Engineering Change Management Improving Collaboration by Assessing Organizational Dependencies Based on Past Engineering Change Information. In Proceedings of the 2018 IEEE International Conference on Industrial Engineering and Engineering Management (IEEM), Bangkok, Thailand, 16–19 December 2018; pp. 617–621.

42. Kocar, V.; Akgunduz, A. ADVICE: A virtual environment for Engineering Change Management. *Comput. Ind.* **2010**, *61*, 15–28. [CrossRef]

43. Malak, R.C.; Aurich, J.C. Software Tool for Planning and Analyzing Engineering Changes in Manufacturing Systems. *Procedia CIRP* **2013**, *12*, 348–353. [CrossRef]

44. Pan, Y.; Stark, R. An Ensemble Learning based Hierarchical Multi-label Classification Approach to Identify Impacts of Engineering Changes. In Proceedings of the 2020 IEEE 32nd International Conference on Tools with Artificial Intelligence (ICTAI), Baltimore, MD, USA, 9–11 November 2020; pp. 1260–1267. [CrossRef]

45. Do, N. Integration of engineering change objects in product data management databases to support engineering change analysis. *Comput. Ind.* **2015**, *73*, 69–81. [CrossRef]

46. Habhouba, D.; Cherkaoui, S.; Desrochers, A. Decision-making assistance in engineering-change management process. *IEEE Trans. Syst. Man Cybern. Part C Appl. Rev.* **2011**, *41*, 344–349. [CrossRef]

47. Grieco, A.; Pacella, M.; Blaco, M. On the Application of Text Clustering in Engineering Change Process. *Procedia CIRP* **2017**, *62*, 187–192. [CrossRef]

48. Arnarsson, I.Ö.; Frost, O.; Gustavsson, E.; Stenholm, D.; Jirstrand, M.; Malmqvist, J. Supporting Knowledge Re-Use with Effective Searches of Related Engineering Documents-A Comparison of Search Engine and Natural Language Processing-Based Algorithms. *Proc. Des. Soc. Int. Conf. Eng. Des.* **2019**, *1*, 2597–2606. [CrossRef]

49. Hamraz, B.; Caldwell, N.H.; Wynn, D.C.; Clarkson, P.J. Requirements-based development of an improved engineering change management method. *J. Eng. Des.* **2013**, *24*, 765–793. [CrossRef]

50. Heron, M.J.; Hanson, V.L.; Ricketts, I. Open Source and Accessibility: Advantages and Limitations. *J. Interact. Sci.* **2013**, *1*, 2. [CrossRef]

51. Shearer, C. The CRISP-DM Model: The New Blueprint for Data Mining. *J. Data Warehous.* **2000**, *5*, 13–22.

52. Řehůřek, R. Word2Vec Model. 2021. Available online: https://radimrehurek.com/gensim/auto_examples/tutorials/run_word2vec.html (accessed on 12 June 2022).

53. Verleysen, M.; François, D. The Curse of Dimensionality in Data Mining and Time Series Prediction. In *Computational Intelligence and Bioinspired Systems, Proceedings of the International Work-Conference on Artificial Neural Networks, Barcelona, Spain, 8–10 June 2005*; Cabestany, J., Prieto, A., Sandoval, F., Eds.; Springer: Berlin/Heidelberg, Germany, 2005; pp. 758–770.

54. Hall, M.A. Correlation Based Feature Selection for Machine Learning. Ph.D. Thesis , University of Waikato, Waikato, New Zeeland, 1999.

55. Bortz, J.; Schuster, C. *Statistik für Human- und Sozialwissenschaftler*, 7th ed.; Springer: Berlin/Heidelberg, Germany, 2010.

56. Uysal, A.K.; Gunal, S. The impact of preprocessing on text classification. *Inf. Process. Manag.* **2014**, *50*, 104–112. [CrossRef]

57. Camacho-Collados, J.; Pilehvar, M.T. On the Role of Text Preprocessing in Neural Network Architectures: An Evaluation Study on Text Categorization and Sentiment Analysis. *arXiv* **2017**, arXiv:1707.01780.

58. Kao, A.; Poteet, S.R. *Natural Language Processing and Text Mining*; Springer: London, UK, 2007.

59. Devlin, J.; Chang, M.W.; Lee, K.; Toutanova, K. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. *arXiv* **2018**, arXiv:1810.04805.