Supporting Business Process Improvement with Natural Language Processing: A Model-based Approach

Martin Bergner[1] Hans-Georg Fill[2] Florian Johannsen[3]

Abstract: The research area of business process improvement (BPI) focuses on the development of methods and techniques for redesigning processes as well as improving the process performance accordingly. Thereby one particular challenge lies in the classification of customer feedback that is received via various online and offline channels. In the paper at hand we present an approach how customer feedback that is issued via social media can be classified using natural language processing. For this purpose, a domain-specific modeling method has been designed that permits to specify the processing steps in visual form and that can integrate the results in existing modeling methods for BPI. For a first evaluation, the approach has been prototypically implemented on the ADOxx-based SeMFIS platform and assessed using a SWOT analysis.

Keywords: Business Process Improvement, Natural Language Processing, Conceptual Modeling

1 Introduction

One of the current challenges in research on business process management is to provide manageable and practicable approaches that support the goal-oriented conduction of business process improvement (BPI) projects [Va16]. On the one hand, this is regarded as an academic need to close the gap between the areas of business process management and process-oriented quality management [St06]. On the other hand, corresponding support in terms of methods and tools is also strongly sought after by practitioners who are often overwhelmed by the large quantity and complexity of available techniques [JF14a, Ha15]. Besides traditional approaches such as the numerical analysis and simulation of business processes [Ba11], a wide range of methods and tools have been specifically developed for BPI (e.g., Six Sigma) [Ze11].

In a recent attempt to provide a manageable set of BPI techniques and establish a best-practice approach for the systematic conduction of BPI projects, a *roadmap for business process improvement (BPI)* has been developed [JF14a]. The roadmap guides the user during the systematic development of process improvement opportunities, starting with a first analysis of a business process in the form of Supplier-Input-Process-Output-Customer (SIPOC) diagrams, the investigation of customer feedback to a detailed analysis of the

[1] University of Vienna, Faculty of Computer Science, Waehringerstrasse 29, 1090 Vienna, martin.bergner@univie.ac.at
[2] University of Vienna, Faculty of Computer Science, Waehringerstrasse 29, 1090 Vienna, hans-georg.fill@univie.ac.at
[3] University of Regensburg, Chair of Business Engineering, Universitätsstraße 31, 93053 Regensburg, florian.johannsen@wiwi.uni-regensburg.de

process performance by means of key performance indicators (KPIs). In the following, the emphasis is put on the investigation of customer feedback, which represents the base for the definition of project goals [Me13].

With the upcoming of social media platforms such as Twitter, Facebook, LinkedIn, or Xing [HKP12], organizations started to collect customer opinions via these channels complementing traditional sources such as CRM systems or online and offline surveys. In this regard, customer posts, e.g., on an enterprise's Facebook page, represent the so-called ”Voice of the Customer (VOC)” [PNC00], capturing consumers' current attitude towards the product and service offerings or the company in general. This information provides valuable insights into customers' behavior and serves as a base for triggering marketing efforts or BPI projects [GR10]. The information in social media channels can be analyzed in real-time and is - compared to the collection of primary data - accessible at low costs. Further, customers' sentiment as reflected by the posts is highly up-to-date, eliminating the risk of receiving outdated data, which is given when analyzing secondary data exclusively (e.g., complaint reports) [Li12].

The challenge when retrieving customer feedback from social media platforms is that the provided data is largely unstructured. Especially customer utterances in the form of comments are expressed in natural language, which needs to be processed by humans before it can be used for management decisions. Therefore, the following two research questions can be derived:

1. How can support be provided and technically realized for BPI initiatives to classify natural language expressions?
2. How can this support be integrated with model-based approaches for BPI?

For providing support in classifying these natural language expressions and assigning them to BPI initiatives, a model-based approach for the machine-based processing of natural language has been designed and prototypically implemented [Be15]. With this approach it is aimed for a solution that can be adapted to multiple scenarios in BPI. Furthermore, the approach can be integrated with existing modeling methods for BPI. The addressees of the modeling method are BPI experts requiring support in managing BPI projects.

The remainder of the paper is organized as follows: In section 2 we will discuss foundations in the area of model-based BPI, social media platforms, and natural language processing. Then, in section 3 the modeling method will be described. This will include the description of the modeling language and the implementation. Subsequently, a first evaluation using a SWOT analysis is given in section 4. The paper will be concluded by an outlook on future work.

2 Foundations

Before we continue with the description of the modeling method, we first briefly outline some foundations that are necessary to understand the design of the approach.

This includes a short overview on the topic of model-based BPI, a characterization of social media platforms, and an outline of the core techniques that we re-used from the area of natural language processing.

2.1 Model-Based Business Process Improvement

The field of BPI is closely related to the process-oriented quality management discipline [St06]. Besides the domains of business management and information technology, also process-oriented quality management has for a long time been concerned with the simplification of business processes, the control of the quality of the outputs and the corresponding IT support [Ha15, St06]. Considering the multitude of BPI techniques existing [An99, Me13, HGJ06], the challenge lies in selecting the most suitable and effective techniques for supporting employees in the conduction of particular BPI projects [Ze11, JF14a]. This becomes even more important because practitioners increasingly refrain from using extensive methods for performing BPI projects (e.g., Six Sigma) and prefer a manageable set of BPI techniques instead [Da13]. Thereby, in the course of a project, it is not only essential to make use of adequate data sources for investigating the current functioning of a process but also to codify the BPI project participants' knowledge adequately [SM09].

In [JF14a] a proposal for such a set of techniques has been generated in the form of a roadmap – see figure 1. The roadmap contains 11 techniques for business process improvement that are ordered according to the phases "Define", "Measure", "Analyse", "Improve", and "Control" (DMAIC), which were derived from the Six Sigma cycle [PNC00]. For reasons of brevity we only discuss the first phase "Define" in the following and refer interested readers to the details in [JF14a].

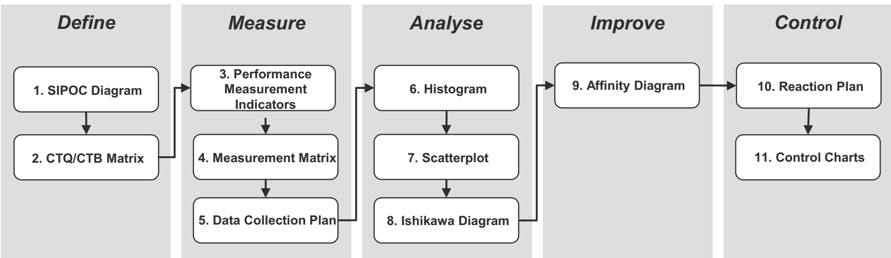

Figure 1: Roadmap for Business Process Improvement [JF14a]

The first BPI technique of the roadmap is the so-called SIPOC diagram. It gives a high-level overview of the inputs and outputs of a business process as well as the process steps and the corresponding process customers [Me13]. Subsequently, the Critical-to-Business (CTB) and Critical-to-Quality (CTQ) factors are elaborated by help of the CTQ/CTB Matrix. For this purpose, customer (VOCs) and employee requirements (Voice of the Business - VOBs), e.g., extracted from CRM systems or social media channels, are collected and condensed to core statements from which CTQ and CTB factors are derived that determine the goals of a BPI project. For instance the VOC statements *"the processing times are quite long"* and *"it took a long time until I received feedback"* can be consolidated to a core statement *"process cycle times"* from which the CTQ factor (project goal) *"reduction of process cycle times by 2 working days"* may be derived. So far, this condensation has usually been accomplished manually in BPI projects. However, considering the large amount of data processed (e.g., data retrieved from social media channels) and limited human resources in BPI projects [GR05] machine-based support is required. All BPI techniques of the roadmap, as shown, were designed as conceptual model types [JF14a] allowing to efficiently codify, communicate and process the results emerging in BPI projects. Accordindly, we speak of an approach for *model-based BPI* in that context [JF14a].

2.2 Social Media Platforms

Social media are today not only used for private communication purposes but are of increasing importance for the communication between enterprises and customers [KA14]. Via social media channels customer inquiries can be efficiently handled, marketing material widely shared or complaints solved quickly amongst others [PM12]. According to Berthon et al. social media thereby fulfill three functions [Be12]: a. the dissemination of information, b. the interaction with people, and c. the interpretation / sense-making of information. Regarding the collection of customer requirements the second and the third function are relevant. The interaction thereby includes both personal interaction with individual users, e.g., via personal messages, as well as in the form of posts and comments on social media content such as Twitter tweets or Facebook posts. The interpretation of information and sense-making refers to the discussion and reflection of phenomena by social media users.

Social media platforms support these functions by providing IT support for constructing a public or semi-public profile, connecting to other users, and exchanging information with them on various levels of interaction [BE08]. For most of these platforms it is possible to retrieve information programmatically via web interfaces. For example, the Facebook Graph API[4] lists more than 50 root nodes as entry points for querying information. This includes aspects such as friend lists, notifications, messages, comments, or photos as well as system information such as copyright rules or debug information. Some of this information is structured, e.g., the number of likes received for an object, but most information is expressed in natural language.

2.3 Natural Language Processing (NLP)

The machine-based processing of natural language information or computational linguistics comprises a large number of different activities that are today supported by a wide

[4] Facebook Graph API Version 2.6: https://developers.facebook.com/docs/graph-api/reference/ accessed 04-05-2016

range of libraries [IMF13]. In the following we briefly discuss the most important concepts used in NLP. For more detailed information we refer to the descriptions in [JM00, IMF13].

One of the first steps is the *identification of the used language*. Thereby, the language is identified in the form of a score to determine the confidence of the identification algorithm.

Part-of-Speech Tagging (POS) assigns labels for grammatical classes of words, i.e., nouns, verbs, adjectives, etc. POS is the basis for resolving ambiguities, e.g., for homonyms, i.e., lexical equivalents with different meanings.

Named Entity Recognition (NER) classifies the mentions of people, organizations, locations, and other named entities. It is essential for aspects such as text summarization or the identification of contextual information.

Sentence Boundary Disambiguation stands for the splitting of sentences into smaller parts. Although this may seem trivial at first sight, the punctuation especially in German and such used in social media contexts requires specific procedures, e.g., to identify emoticons.

Lemmatization refers to the task of finding the base form of a word, i.e., the form contained in a dictionary. It is required for subsequent tasks such as query expansion and finding similar words. It shall not be mixed up with *stemming* which does not take the context of a word into account. *Word Splitting* stands for the splitting of compound words into their individual parts. This applies to both orthographic compounds (e.g., ’school-bus’) as well as semantic compounds (e.g., ’bittersweet’).

Query Expansion is a common technique in information retrieval for enhancing the query performance by adding additional words. A typical approach is to add synonyms and hypernyms by using a thesaurus to increase the probability of finding words with related meaning in other texts.

Term Weighting measures the occurrence of terms in a text and uses it to assign weights to the terms. The assumption is that more important terms occur more frequently. Based on these weights, comparisons between texts can be conducted to measure their similarity. This is accomplished by *Similarity and Dissimilarity Measures*, e.g., the often used cosine similarity between two vectors of term frequencies.

Another important technique that is often used in natural language processing is *Stop Word Removal*. Thereby, words that carry little semantic meaning are removed from a text before it is being processed. However, it has to be carefully chosen which words are to be removed as they are sometimes decisive for semantic meaning in combination with other words.

3 A Modeling Method for Natural Language Processing in Business Process Improvement

With the above described foundations we can now advance to the description of a modeling method for supporting natural language processing in the context of BPI. Thereby we follow a traditional design science research process in which we focus on the design and development of the artifact in form of the modeling method [Pe07, KK02]. The goal of the modeling method is to close the gap between business requirements and technological opportunities by simplifying the processing of natural language information from social media [De08]. We therefore focused in particular on the second step of the BPI roadmap where the utterances of customers (VOCs) are classified and used for deriving project goals.

In this context, the major challenge is to assign customer utterances in social media channels to the affected business processes. For humans who are familiar with both, the business processes of a company and the way customers express feedback via social media, this can be achieved quite easily. However, for machines that have to process this information automatically this is a rather complex issue. One of the reasons for this complexity is the fact that the knowledge about the business processes is typically not accessible by machines. Although one might argue that it could be referred to business process models or internal process descriptions that document this knowledge [KJS96], the problem is the *semantic gap* between descriptions of business processes that are used by a company internally and such that customers perceive beyond company borders. This also manifests itself in analyses of the language used by customers in social media channels and that used in business process descriptions. We thus had to find a mediator between these two language worlds.

The way we addressed it was to revert to so-called 'FAQ' sections on company websites. These sections document questions that are frequently asked by customers including the answers returned by employees of that company. Although these sections do not directly refer to actual business processes of a company, the fact that they are often well maintained and structured permits to easily infer which business processes are affected by company professionals. In addition, they include enough quantity of textual information as it is required for successful natural language processing.

3.1 Design of the Modeling Method

When building NLP applications a large number of APIs, services, and platforms are available today [IMF13]. Some of these platforms include visual modeling languages for specifying the processing of natural language information by invoking distinct processing components. Two notable examples include the visual editor included in the current release of Microsoft's Azure Machine Learning Studio (AzureML) [5] and Unitex [Pa16]. Besides several machine learning algorithms, AzureML permits to visually specify steps for natural language processing such as the pre-processing and filtering of datasets, format conversions as well named entity recognition or different methods for term frequency calculations. Unitex on the other hand is a collection of programs that have been developed for analyzing natural language texts. It is a low-level tool for processing natural language that follows a graph-based approach. Unitex graphs can handle the automatic inflection of dictionaries, the pre-processing of texts, the normalization of text automata, the search for patterns and several more [Pa16][p.119ff.].

[5] See https://studio.azureml.net/ last accessed 06-05-2016

For our approach we chose to take a high-level approach that can be integrated with existing enterprise modeling methods in the area of business process management and BPI. For this purpose we specified three model/diagram types as shown in the meta model in figure 2.

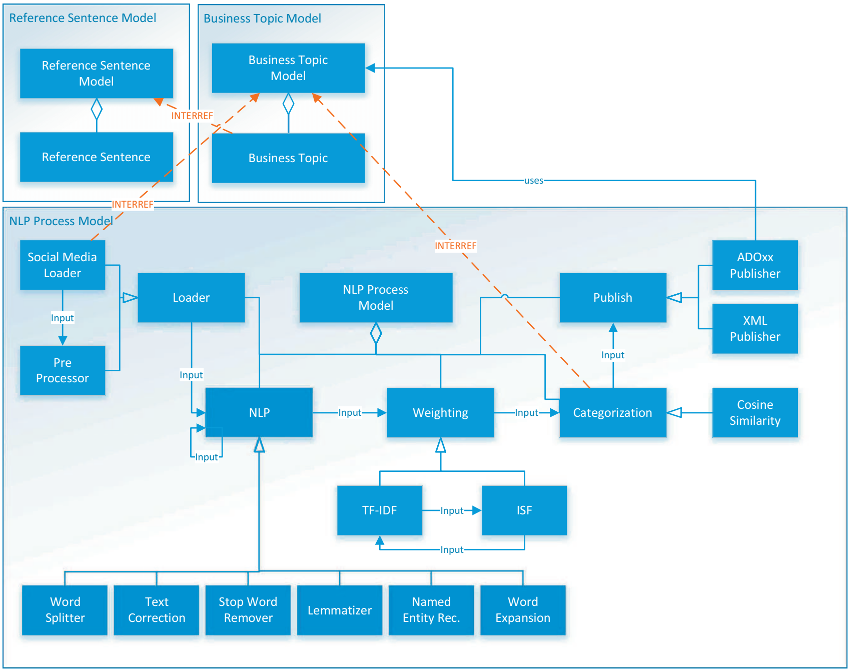

Figure 2: Meta Model Showing the Three Model Types: Reference Sentence Model, Business Topic Model, and NLP Process Model

The *business topic model* and the *reference sentence model* are used to represent categories of topics to which information from social media shall be assigned. This can for example be used to represent the headings of a company's FAQ sections together with the corresponding questions and answers given in that section. Another use case would be to define arbitrary business sections that correspond to a company's business areas and assign textual information, e.g., from the company website or other available documents.

The *NLP process model* is responsible for specifying the steps for processing natural language information. It is linked to the business topic model type to receive the information about the categories to which the natural language information shall be assigned. Furthermore, it contains the following modeling classes that a user can arrange to determine the processing steps. These classes generally follow the process of *loading* natural language data, *preparing* the data for categorization using NLP techniques, *categorizing* the data based on the user-defined categories in the business topic model, and *publishing* the results.

The *Social Media Loader* is responsible for the settings of the data loading process. It enables the user to choose between several social media platforms and to define the corresponding settings, e.g., to retrieve posts from a company's Facebook page. The *Pre Processor* class is responsible for the splitting of posts into sentences and several filterings prior to processing. Users are able to filter posts by length and language, replace abbreviations, hyperlinks, etc. The *Word Splitter* class splits compound words into individual words. The *Stop Word Remover* class handles the removing of stop words based on pre-defined stop word lists. The *Word Lemmatizer* is concerned with the the lemmatization and correction of the capitalization of words. The *Named-Entity Recognition* class is used to identify locations and persons using a corpus. The *Word Expansion* class handles the addition of synonyms to each word by using a thesaurus file in the sense of a query expansion. The *TF-IDF* weighs each term in a given document by using the TF-IDF algorithm. The *ISF* class weighs all sentences of a document, using the the inverse sentence frequency, a modified form of the Inverse Document Frequency (IDF), which is used to identify and remove unnecessary sentences in a document. The *Cosine Similarity* class categorizes the incoming natural language information depending on the prior weighting values. The *ADOxx Publisher* class is responsible for exporting an ADOxx compatible XML file, which contains the results of the categorization in a Business Topic and Reference Sentence Model. The *XML Publisher* class exports an XML file with the analysis results that can be processed by a specifically created web application called Social Media Categorization Viewer, which will be discussed in the next section.

3.2 Prototypical Implementation

The modeling method has been implemented as a prototype on the ADOxx meta modeling platform [FK13] – see the screenshot in figure 3. In order to enable later extensions for semantic processing with ontologies, it was chosen to extend the SeMFIS library [Fi16]. SeMFIS includes several model types for representing ontologies which will be used in the future to extend the descriptions used in the business topic and the reference sentence model.

In addition to the creation of visual model editors for the model types described in the previous section, several additional software components and services were implemented to realize the NLP steps. As the intended target application area of the system were German social media posts and tweets, the corresponding versions for German were selected.

For pre-processing, the Apache Commons Lang library[6] was used, for the language detection the language detection library for Java[7], for word splitting the jWordSplitter library by Daniel Naber[8], for lemmatizing the mate-tools[9], and for named entity recognition the Stanford CoreNLP toolkit [Ma14].

[6] Apache Commons Lang: https://commons.apache.org/proper/commons-lang/download_lang.cgi last accessed 06-05-2016

[7] https://github.com/shuyo/language-detection/blob/wiki/ProjectHome.md last accessed 06-05-2016

[8] Decomposition of German compound words with jWordSplitter http://www.danielnaber.de/jwordsplitter/index_en.html last accessed 06-05-2016

[9] https://code.google.com/archive/p/mate-tools/ last accessed 06-05-2016

To establish the interfaces to social media platforms the RestFB library[10] for Facebook and the Twitter4j library for twitter[11] were included.

For several of the natural language processing components external datasets and corpora are required. For the lemmatizer and the POS model we used the datasets of the mate-tools, for the stop word removal the stop word list provided on Google Code[12], for the text correction the 'incorrect word list' from the German Wikipedia[13], for the detection of emoticons by the pre-processor the emoticon list from the English Wikipedia[14], for the Named Entity Recognition the WaCky Corpus [Ba09], and for the Word Expander the thesaurus provided by Open Thesaurus[15].

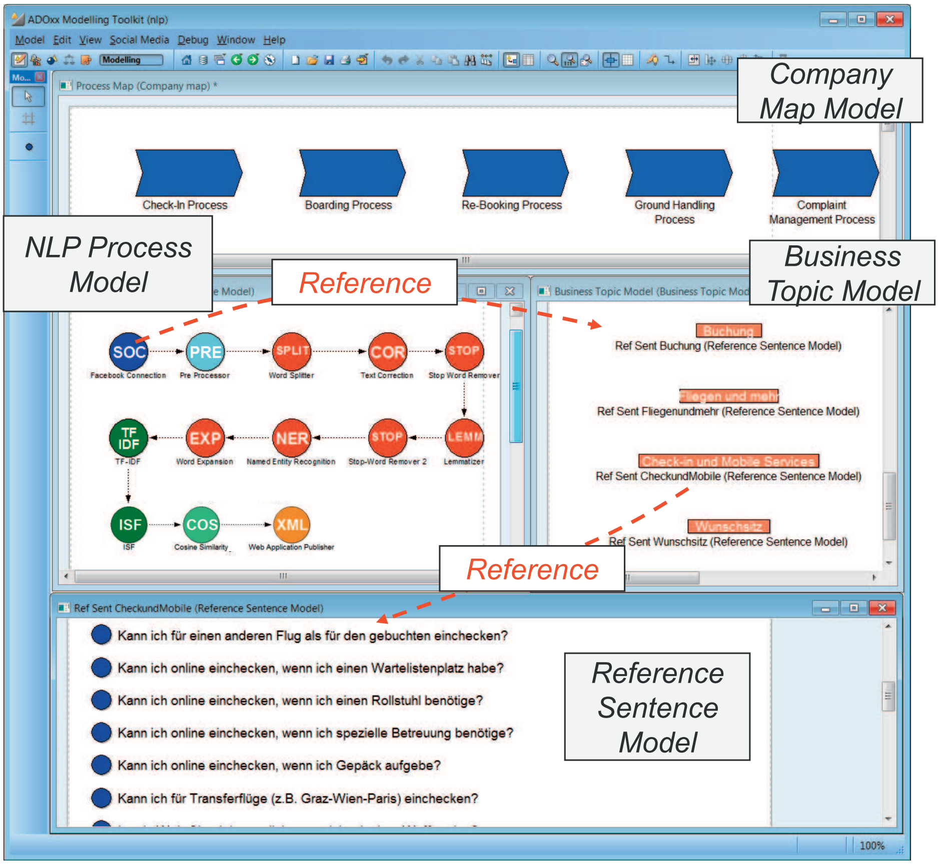

Figure 3: Screenshot of the Implementation of the Modeling Method on ADOxx

[10] http://restfb.com/ last accessed 06-05-2016
[11] http://twitter4j.org/ last accessed 06-05-2016
[12] https://code.google.com/archive/p/stop-words/ last accessed 06-05-2016
[13] https://de.wikipedia.org/wiki/Kategorie:Wikipedia:Falschschreibung last accessed 06-05-2016
[14] https://en.wikipedia.org/wiki/List_of_emoticons last accessed 06-05-2016
[15] https://www.openthesaurus.de/ last accessed 06-05-2016

In figure 3 at the top, a sample model of a process map is shown that is included in the standard distribution of SeMFIS and below the three new model types. The sample NLP process model shows a standard NLP process for processing posts that are retrieved via the Facebook connection. It includes steps such as pre-processing, stop word removal, lemmatization, named entity recognition, word expansion and cosine similarity calculation based on the TF-IDF and ISF measures. The final step in the process is the publication of the results via a web application. On the right hand side an example for a business topic model is shown. The model is referenced from the Social Media Loader class. The topics in the model contain references to several reference sentence models. An example is shown at the bottom. The information from these models is used in the NLP process model to assign the Facebook posts to the categories in the business topic model.

In figure 4 the web application 'Social Media Categorization Viewer' for analyzing the results of the execution of the NLP process are shown. It receives the data from the main application running on ADOxx via an XML interface. Via this application, users can inspect the assignment of the Facebook posts to the categories specified in the business topic model. The current version of the web application also includes the possibility for users to give information about the adequacy of the assignment of the posts. For this purpose, a mechanism has been added in the implementation where users can manually classify a number of posts to the defined categories. For assessing the accuracy of the automatic assignment, the standard measures for precision and recall can then be shown in addition to the classification for each category and for all posts in total.

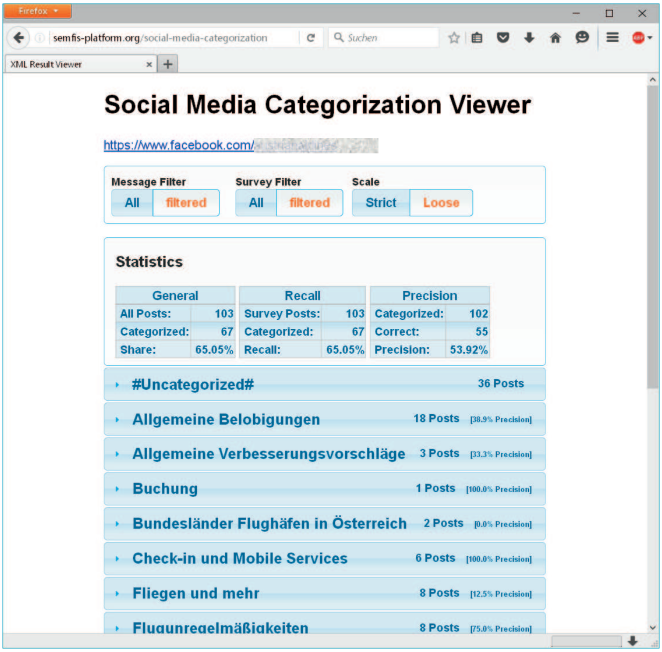

Figure 4: Screenshot of the Social Media Categorization Viewer Web Application

4 Evaluation Using a SWOT Analysis

The approach has undergone a first evaluation in the form of a SWOT analysis. Based on the application of the approach to social media data that has been retrieved from the Facebook presence of an Austrian airline company, the following strengths, weaknesses, opportunities, and threats could be assessed. More detailed evaluations including detailed user evaluations and empirical analyses are already planned and will be conducted in the near future.

The *strength* of the approach lies in the integration with existing enterprise modeling approaches. Especially for the support of BPI projects, the visual specification of topics and reference sentences is easy enough to be accomplished also by domain experts. As the technological details of the natural language processing steps are abstracted through the modeling language, users do not need in-depth technological know-how for applying the approach. The standard NLP process as shown in the sample above does not necessarily need to be modified and can thus be applied to various NLP classification tasks.

The major *weakness* of the approach is that it is currently tied to natural language processing of posts and comments in German. Although plenty of libraries exist for conducting a similar procedure in English or other languages, technical know-how about the NLP configuration process is required to achieve multi-language capabilities. However, the modeling language as implemented on ADOxx already offers extension points for this case so that for example English thesauri or stop-word lists can be assigned via changes in attribute values of the model.

An *opportunity* of the presented approach is its extensibility with further NLP processing features and its integration in further modeling methods. Besides the mentioned algorithms and NLP services, a multitude of other NLP algorithms and libraries exist. By extending the NLP process model, these can be integrated into the processing pipeline. In addition, the implementation of the modeling method based on the SeMFIS library offers the opportunity of using more formal corpora for the NLP tasks. For example, by reverting to the OWL ontology model type of SeMFIS, formal specifications between terms can be expressed that can be subsequently used by the NLP components. Regarding the integration with other modeling methods, not only the integration with methods directed towards BPI such as the RUPERT modeling tool [JF14b, JF15] become possible. Also other modeling approaches e.g., in the area of strategic or innovation management could profit from the integration with classification results from social media. Generally, regarding BPI, the automatized analysis of social media data enables to get detailed insights into those issues that negatively or positively shape customers' current sentiment (e.g., service quality, employee friendliness, etc.). This allows a firm to trigger corresponding BPI efforts in consequence. At present, the classification of customer posts mostly is done manually by the workforce (e.g., by reverting to the CTQ/CTB Matrix), which is a time-consuming and resource-intense process. Thus, the solution introduced above represents a significant contribution for the goal-oriented conduction of BPI projects.

Although the proposed approach does not lead to any direct *threats* in our view, there are some indirect consequences that may result from it. For example, users of the NLP processing may be tempted to rely too much on its results and may not inspect any more how these results have been achieved. Although first test runs with sample data have shown that well-configured NLP processes can lead to satisfying results, these are far from perfect - i.e., they never reach 100% accuracy in terms of precision and recall. Despite the great advances also in the area of NLP there will always be some uncertainty that users have to deal with.

5 Conclusion and Outlook

In this paper we have presented an approach for simplifying the use of machine-based natural language processing with a modeling method. The goal of the method is to support BPI by classifying in particular the utterances of customers (VOCs) in social media and by highlighting the affected business processes accordingly. The approach has undergone a first evaluation in the form of a SWOT analysis.

However, there are several aspects that need to be evaluated in more detail. These can be grouped into technical / architectural, NLP and BPI-specific aspects. Regarding technical aspects, the performance of the presented approach will have to be evaluated in detail. This affects how many posts, e.g., from Facebook or Twitter, can be retrieved and correctly classified in a given time frame. For any practical usage this is highly relevant as the amounts of information in social media channels increase dramatically every day. In this context also the architecture consisting of the meta modeling platform and the NLP components will have to be assessed. The current prototype may not yet fully exploit the optimal exchange of information between all components and will thus have to be optimized also in this aspect.

The use of the NLP components will have to undergo a separate evaluation. Here, it will need to be checked if the components that are currently used lead to the optimal performance of the overall system or whether some of them can be replaced with more powerful versions. For this purpose, a number of test runs with different test data will have to be performed. The results will then have to be compared with manually classified natural language data. This is already technically supported by the presented application so that detailed performance results on the classification will be delivered.

Finally, aspects that are specific to the domain of BPI will have to be evaluated. As BPI projects are conducted by practitioners it will need to be tested whether the approach meets their requirements or not. This concerns the usability of the tool, the embedding into BPI projects as well as the quality of the generated results. Therefore, in-depth investigations with real-life social media data from companies will have to be performed and evaluated together with practitioners. If necessary, this will lead to refinements of the modeling as well as the natural language processing approach.

The immediate next steps include the further development of the prototype to advance it to a level that can be tested with practitioners and academics alike. To permit the evaluation in the context of BPI it is also aimed for the integration of the modeling method in the RUPERT modeling toolkit that has been previously applied to BPI use cases [JF14a, JF14b].

References

[An99] Andersen, B.: Business process improvement toolbox. ASQ Quality Press, 1999.

[Ba09] Baroni, M.; Bernardini, S.; Ferraresi, A.; Zanchetta, E.: The WaCky wide web: a collection of very large linguistically processed web-crawled corpora. Language Resources and Evaluation, 43(3):209–226, 2009.

[Ba11] Barjis, J.: Enterprise Modeling and Simulation with Enterprise Engineering. Journal of Enterprise Transformation, 1(3):185–207, 2011.

[BE08] Boyd, D.M.; Ellison, N.B.: Social Network Sites: Definition, History, and Scholarship. Journal of Computer-Mediated Communication, 13:210–230, 2008.

[Be12] Berthon, P.R.; Pitt, L. F.; Plangger, K.; Shapiro, D.: Marketing meets Web 2.0, social media, and creative consumers: Implications for international marketing strategy. Business Horizons, 55:261–271, 2012.

[Be15] Bergner, M.: Integrating Natural Language Processing with Semantic-based Modeling. Master's thesis, University of Vienna, 2015.

[Da13] Davis, D.: 3rd Biennial PEX Network Report: State of the Industry - Trends and Success Factors in Business Process Excellence. Technical report, PEX Network, 2013.

[De08] Demirkan, H.; Kauffman, R. J.; Vayghan, J. A.; Fill, H.-G.; Karagiannis, D.; Maglio, P.P.: Service-oriented technology and management: Perspectives on research and practice for the coming decade. Electronic Commerce Research and Applications, 7(4):356–376, 2008.

[Fi16] Fill, H.-G.: SeMFIS: A Flexible Engineering Platform for Semantic Annotations of Conceptual Models. accepted for Semantic Web (SWJ), 2016.

[FK13] Fill, H.-G.; Karagiannis, D.: On the Conceptualisation of Modelling Methods Using the ADOxx Meta Modelling Platform. Enterprise Modelling and Information Systems Architectures, 8(1):4–25, 2013.

[GR05] Gijo, E.V.; Rao, T. S.: Six Sigma Implementation - Hurdles and More Hurdles. Total Quality Management & Business Excellence, 16(6):721–725, 2005.

[GR10] Gallaugher, J.; Ransbotham, S.: Social Media and Customer Dialog Management at Starbucks. MIS Quarterly Executive, 9(4):197–212, 2010.

[Ha15] Harmon, P.: The Scope and Evolution of Business Process Management. In (Vom Brocke, J.; Roseman, M., eds): Handbook on Business Process Management 1. Springer, pp. 37–80, 2015.

[HGJ06] Hagemeyer, C.; Gershenson, J.K.; Johnson, D.M.: Classification and application of problem solving quality tools: A manufacturing case study. The TQM Magazine, 18(5):455–483, 2006.

[HKP12] Heidemann, J.; Klier, M.; Probst, F.: Online social networks: A survey of a global phenomenon. Computer Networks, 56(18):3866–3878, 2012.

[IMF13] Ingersoll, G.S.; Morton, T.S.; Farris, A.L.: Taming Text - How to Find, Organize and Manipulate it. Manning, 2013.

[JF14a] Johannsen, F.; Fill, H.-G.: Codification of Knowledge in Business Process Improvement Projects. In: European Conference on Information Systems (ECIS'2014). AIS, 2014.

[JF14b] Johannsen, F.; Fill, H.-G.: RUPERT: A modelling tool for supporting business process improvement initiatives. In: Proceedings of DESRIST'2014. Springer, pp. 418–422, 2014.

[JF15] Johannsen, F.; Fill, H.-G.: Supporting Knowledge Elicitation and Analysis for Business Process Improvement through a Modeling Tool. In: International Conference on Business Informatics 2015. AIS, 2015.

[JM00] Jurafsky, D.; Martin, J.: Speech and language processing: An introduction to natural language processing, computational linguistics, and speech recognition. Prentice Hall, 2000.

[KA14] Kane, G.C; Alavi, M.: What's Different About Social Media Networks? A Framework and Research Agenda. MIS Quarterly, 38(1):275–304, 2014.

[KJS96] Karagiannis, D.; Junginger, S.; Strobl, R.: Introduction to Business Process Management Systems Concepts. In (Scholz-Reiter, B.; Stickel, E., eds): Business Process Modelling. Springer, Berlin et al., pp. 81–106, 1996.

[KK02] Karagiannis, D.; Kuehn, H.: Metamodeling Platforms. In: Third International Conference EC-Web 2002 â Dexa 2002. Springer, Aix-en-Provence, France, p. 182, 2002.

[Li12] Liu, B.: Sentiment analysis and opinion mining. Synthesis Lectures on Human Language Technologies, 5(1):1–167, 2012.

[Ma14] Mannning, C.; Surdeanu, M.; Bauer, J.; Finkel, J.; Bethard, S.J.; McClosky, D.: The Stanford CoreNLP Natural Language Processing Toolkit. In: 52nd Annual Meeting of the Association for Computational Linguistics: System Demonstrations. pp. 55–60, 2014.

[Me13] Meran, R.; John, A.; Roenpage, O.; Staudter, C.: Six Sigma+Lean Toolset. Springer, 2013.

[Pa16] Paumier, S.: UNITEX User Manual 3.1 - March 27, 2016. Universite Paris-Est Marne-la-Vallee, 2016.

[Pe07] Peffers, K.; Tuunanen, T.; Rothenberger, M.A.; Chatterjee, S.: A Design Science Research Methodology for Information Systems Research. Journal of Management Information Systems, 24(3):45–77, 2007.

[PM12] Pinto, M. B.; Mansfield, P.: Facebook as a complaint mechanism: An investigation of millennials. Journal of Behavioral Studies in Business, 5, 2012.

[PNC00] Pande, P.S.; Neuman, R.P.; Cavanagh, R.R.: The Six Sigma Way: How GE, Motorola, and other top companies are honing their performance. McGraw-Hill, 2000.

[SM09] Seethamraju, R.; Marjanovic, O.: Role of process knowledge in business process improvement methodology: a case study. Business Process Management Journal, 15(6):920–936, 2009.

[St06] Stracke, C.: Process-oriented quality management. In (Ehlers, U.D.; Pawlowski, J.M, eds): Handbook on Quality and Standardisation in E-Learning. Springer, pp. 79–96, 2006.

[Va16] Vanwersch, Rob J. B.; Shahzad, K.; Vanderfeesten, I.; Vanhaecht, K.; Grefen, P.; Pintelon, L.; Mendling, J.; Merode, G.G.; Reijers, H.A.: A Critical Evaluation and Framework of Business Process Improvement Methods. Business & Information Systems Engineering, 58(1):43–53, 2016.

[Ze11] Zellner, G.: A structured evaluation of business process improvement approaches. Business Process Management Journal, 17(2):203–237, 2011.