Towards an Algorithm-Based Automatic Differentiation of Liability Cases by Analyzing Complaint Texts

Insa Lemke and Nadine Schlüter[a]
Research Group Product Safety and Quality Engineering, University of Wuppertal, Gaußstraße 20, Wuppertal, Germany

Keywords: Complaint Management, Liability Management, Algorithm, Decision Support System.

Abstract: Effective complaints management is important to maintain customer loyalty and offers the opportunity to feed knowledge back into product development and production. However, with products and supply chains becoming increasingly complex, the picture is often unclear when it comes to handling complaints. This applies in particular to the handling of legal liability issues. The challenge arises from the correct classification of the various legal bases in connection with the receipt of a customer complaint. To this end, we introduce the concept of an algorithm that uses automatic text recognition to analyze the text of a complaint and determine whether a liability case may exist under German law. This paper presents the different development steps and phase components of the algorithm as well as the current implementation status.

1 INTRODUCTION

Inadequate complaint management can result in significant financial losses and risks for manufacturing companies (Hardin, 2015). Defective products that have already been launched on the market not only jeopardize customer loyalty but also increase the likelihood of legal consequences. (Cieśla, 2023; Stauss & Seidel, 2019). Both manufacturers and sellers face multiple challenges concerning product liability-related repercussions. On the one hand, it is important to identify any liability risks that could arise from complaints as quickly as possible. On the other hand, those responsible in the supply chain as well as the correct countermeasures must be identified and initiated as quickly as possible (Schmitt & Linder, 2013; Stauss & Seidel, 2019; Yilmaz, Varnali & Kasnakoglu, 2016).

To tackle this issue, this paper presents the development of an algorithm that automates the assessment of complaint texts for potential product liability cases resulting from customer complaints. The "AlGeWert" project aims to automate aspects of complaints management in industry to enhance the certainty of action in complaints processing. The following sections deal with the development and presentation of the current prototype status.

Section 2 introduces the problem and discusses relevant obstacles and criteria. Section 3 examines the current state of the art pertaining to automated complaints processing, specifically exploring the incorporation of liability-relevant considerations in research. Section 4 discusses the algorithm's various components and their methodological contexts. In section 5, we present the current state of implementation and explain how the separate phases are executed. We discuss the achievements and limitations of the current state of work in section 6. Finally, section 7 provides a summary and an outlook on the following project components.

2 PROBLEM DEFINITION AND RESEARCH BACKGROUND

In the modern era of industrialization, demand for faster product availability, and international supply chains, the requirements for product development are changing (Anagun, Bolel, Isik & Ozkan, 2022; Cieśla, 2023). Additionally, increasingly complex products not only challenge their design and manufacturing, but also complaint management

[a] https://orcid.org/0000-0001-6008-9797

Lemke, I. and Schlüter, N.
Towards an Algorithm-Based Automatic Differentiation of Liability Cases by Analyzing Complaint Texts.
DOI: 10.5220/0012569100003690

In *Proceedings of the 26th International Conference on Enterprise Information Systems (ICEIS 2024) - Volume 1*, pages 603-611
ISBN: 978-989-758-692-7; ISSN: 2184-4992

(Agren et al., 2019; Anagun et al., 2022). Effective processing and assessment of customer complaints during the use phase can bring numerous benefits. However, this task is frequently viewed as tedious and unprofitable (Stauss & Seidel, 2019). Inadequate complaint management not only misses opportunities to improve customer loyalty and use feedback to enhance products, but also poses risks in identifying liability issues for manufacturers and retailers due to faulty products in the market. Personnel without training in legal matters may face challenges in differentiating between various scenarios.

The presented problem focuses on liability cases arising from defective products in Germany and German law, which have three legal bases. Liability cases are considered under the statutory warranty that applies when a purchase contract is concluded, as well as the Product Liability Act (based on the European Product Liability Directive) and the German Civil Code (BGB). These legal bases differ in terms of time limits, types of defects, and severity, among other factors. Additionally, each of these bases considers varying responsibilities and draws dependent responsibility. Although a distinction between these three legal bases is theoretically possible, the practical application proves challenging.

To address the issue of automated complaint processing and the identification of potential liability cases, the "AlGeWert" project aims to develop an algorithm capable of conducting precise analysis of complaint texts. By performing an automated analysis of the complaint text, algorithms should be able to identify potential liability cases quickly and accurately. This approach should not only expedite prompt responses to customer inquiries but also assist in the early identification of legal liabilities. Through the reduction of human errors and subjective interpretations, objective and standardized analysis is made certain.

Before explaining the concept and practical application of the algorithm to be developed, the next section examines the current state of automated complaints management and the handling of liability cases in complaints processing.

3 STATE OF THE ART IN AUTOMATED COMPLAINTS MANAGEMENT

Responding quickly and handling complaints efficiently can provide important insights for product development and customer retention. However, the focus is not always on the legal or liability implications. This section examines various projects and publications that have already dealt with improving the processing of complaints in industrial companies.

In a literature review, Zaby and Wilde (2018) examine previous research on complaint management, particularly from a customer relationship management perspective. Despite an extensive literature review, they conclude that there is a great need for a comprehensive review of complaint management, but only a few publications address the topic. However, there are a handful of publications that address the need for complaint management to improve customer satisfaction and product safety and quality.

Behrens, Wilde, and Hoffmann (2007) recognize the need to include suppliers and customers in value chains. This is the only way to establish a product quality control process. They combine their approach with the so-called 8D method, which is a common standard in Germany, especially in the automotive industry, to improve complaints. The 8D method uses a fixed sequence of steps that lead to the identification of the causes of problems and complaints, but it does not yet offer the possibility of specifically querying aspects that would consider the possibility of liability cases occurring.

Schmitt and Lindner (2013) point out that an examination of complaint management can also provide valuable information beyond customer relationship management, particularly regarding product quality and continuous improvement. They present their own approach to technical complaint management, but do not specify legal sources and liability-related requirements.

Hake, Rehse, and Fettke (2021) analyze the potential for automation of the 8D method in medical technology. They consider legal regulations, but do not address product liability perspectives due to the higher standards applied to this field.

Hedge (2023) emphasizes the significance of reliability-focused product development in curbing customer complaints, lowering warranty expenses, and mitigating negative publicity resulting from defective products. The concept also factors in product liability, yet it does not delve further into the practicality of complaints and their relevant assessment.

Stauss and Seidel (2019) highlight the importance of incorporating liability-relevant data in complaints management. Nonetheless, they fail to specify the practical implementation and disregard automated applications.

Based on the research presented, it is evident that automated complaint processing can offer substantial assistance in handling procedures. Additionally, sources indicate that considering liability-related criteria is critical to the success of complaint management.

However, no explicit examples detailing the implementation of this consideration exist. The AlGeWert project aims to close this gap by showing how an algorithmic process can be implemented to ensure automatic text analysis to check complaints for liability-relevant factors.

4 CONCEPT OF THE ALGORITHM

The AlGeWert algorithm is based on a structure of different processes where several elements influence each other. The interplay of these elements affects the automated analysis outcome. The subsequent code is not the sole determinant of the analysis result. To ensure a secure and reliable data interpretation system, the algorithm is to be linked to external databases, including a well-curated information base. For practical purposes, we assume that an accurately maintained database of customer and order information serves as the primary source of data for the algorithm. Equally significant is the integration with an organizational system that charts the value chain of each potential product under consideration and clearly outlines the roles, as well as references to collaborating suppliers, vendors, etc. The latter point is a distinct work package of the project, not to be discussed in detail at present due to the current focus and scope of this paper. For now, the algorithm's explanation assumes availability of a system capable of identifying the manufacturers and suppliers involved in an affected product. As input text for the algorithm, complaint messages in the form of freely worded email messages are considered. It is crucial to combine these messages with a well-maintained database of the current project status, as the quality of such text can vary greatly in terms of content, spelling, and grammar.

After presenting basic information, subsequent sections will address the individual algorithm phases that the AlGeWert algorithm undergoes when analyzing a complaint text.

4.1 Determine Customer Data

At the start of the analytical process, the algorithm uses simple comparison mechanisms to determine whether the order number is the first indicator, since it is the primary key that allows a direct link to the relevant order. If the order number is absent, the algorithm executes several iterations to locate the relevant order through other means. First, it examines the complaint text for the customer number, then for the name, and lastly for an address. Figure 1 depicts the structured hierarchy for matching customer data in the text.

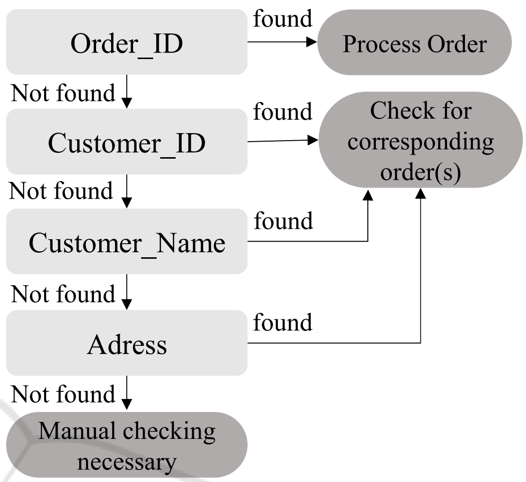

Figure 1: Matching hierarchy to identify the order data.

If customer data is not found or the information does not match the database, there may be two reasons: the complaint lacks necessary data, or the product was purchased from another vendor. However, this does not automatically eliminate liability, making a manual review or consultation with the customer necessary. In this case, the algorithm stops here and provides a note to the processing staff.

If the algorithm identifies customer data successfully, it proceeds to the next stage.

4.2 Synchronizing Order Data

In the second stage, the algorithm retrieves information from the database by utilizing the data obtained after comparing the complaint text with the database. The data searched for may vary slightly depending on whether the customer or order information has previously been identified. The process is illustrated in Figure 2.

The most straightforward way to obtain information is by identifying an order number. In this scenario, the algorithm can access the corresponding order and product directly. If there are numerous products purchased under an order, the complaint text is used to compare with their names. If no product

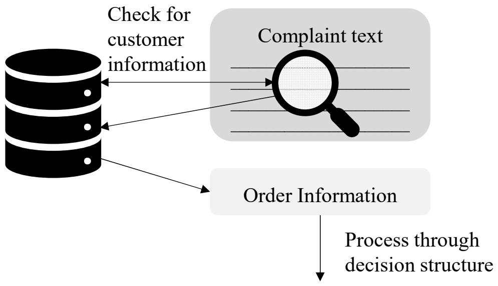

Figure 2: Processing the found customer information.

name is found within the complaint text, all purchased products are verified against the latter criteria and a note is presented at the end for all potential cases. If no order number is found, the scenario becomes similar. In the case of multiple orders found, the complaint text is compared or checked for a purchase date. If the algorithm still cannot uniquely identify an order, it will check all customer orders and provide relevant user information at the end.

4.3 First Decision Phase

After extracting the customer and order data in the first two phases, the algorithm can proceed with evaluating the complaint and determining the liability risk. The algorithm consists of two distinct decision phases, each based on different methods. Technical term abbreviations will be explained when introduced. The first decision phase utilizes a comparison, similar to the preceding steps, which involves a straightforward comparison of information. The AlGeWert algorithm's first decision phase, as displayed in Figure 3, utilizes the customer's residential location (extracted from customer information) and purchase date (extracted from order information).

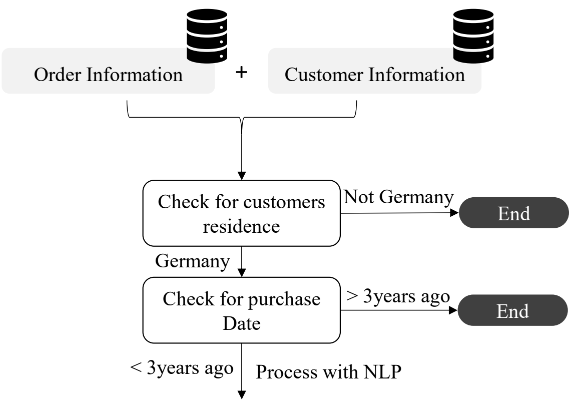

Figure 3: First Decision Phase of the AlGeWert Algorihtm.

As Section 2 has already illustrated, there are three distinct laws that establish criteria for manufacturers and sellers in Germany to be held liable in the case of defective products. However, all of these laws can only be fully enforced if the buyer is also living in Germany. It should be noted that this does not automatically imply an absolute absence of liability if the customer does not reside in Germany. In this situation, however, it is impractical for the algorithm to make an automated decision as it necessitates consideration of various marginal criteria. Consequently, the process concludes by referring to the requirement for a manual check. Likewise, if the purchase date surpasses three years, it is a crucial period where the expiration of this duration generally implies that the manufacturer and seller are no longer accountable for defects within the product. However, it cannot be completely ruled out that certain product groups may have different time limits. If the purchase date is more than three years ago, the algorithm will require a manual check. If neither of the above two points apply, the algorithm proceeds to the next step of determining the error information.

4.4 Determining Failure Information

Previous project research indicates the relevance of identifying the specific defect type for classifying different legal bases. While it is feasible to examine a product liability case under multiple legal bases, considering the defect type remains the most practical approach for an automated preliminary assessment. Nevertheless, straightforward comparison mechanisms prove insufficient for this purpose; it is inadequate to rely solely on keyword identification. For this reason, the algorithm requires a Natural Language Processing (NLP) model to identify the type of complaint. The aim is to distinguish three main aspects and determine whether the faulty product

1. is damaged and no longer functions,
2. has damaged other property of the customer due to the fault or
3. has led to personal injury.

A rather small NLP structure is required, since only the failure categories "Product Defect", "Property Damage" & "Personal Injury" need to be determined. Initially, a bag-of-words model is used in combination with the pre-trained SpaCy pipeline, eliminating the need for pre-processing or tagging. This NLP model is then used to examine the entire complaint text received in order to identify the failure based on the problem description. This requires that the defect is at least partially described - i.e., whether the product is defective or has already caused other damage. The complaint text will then be categorized into one of the three mentioned failure types, along with the purchase date, to facilitate the production of meaningful assessments during the second decision phase.

4.5 Second Decision Phase

The NLP framework analysis offers insights into the type of failure and the required action. Relevant criteria from various liability cases guide the algorithm's assessment of potential risks. The algorithm determines the liability scenarios associated with the defect and the link with the date of purchase. In this manner, objective criteria can be utilized to evaluate and analyze the potential occurrence of various legal grounds. The process for the ultimate determination of potential liability cases is depicted in Figure 4.

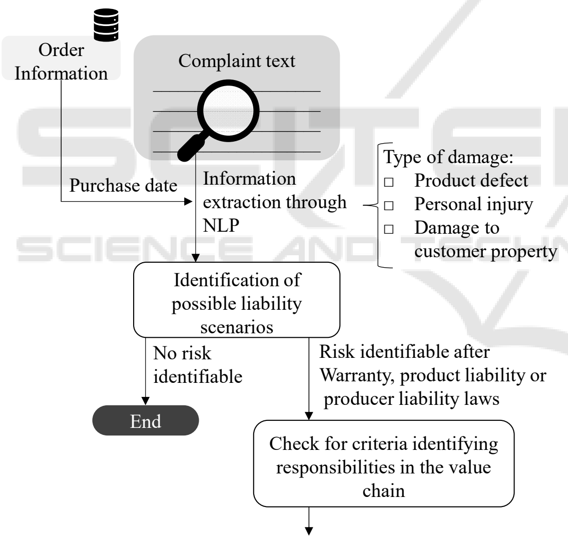

Figure 4: Second decision phase using an NLP framework.

After the second decision phase, there are two courses of action. If the algorithm links the pertinent information and concludes that there is no potential risk of liability, it will terminate and display the message "No liability risk could be identified". However, if there is a potential risk of liability according to one of the three legal bases, the algorithm will proceed to the subsequent iteration. To identify potential liability cases more precisely, liability risks that have been identified are temporarily stored and compared with the process or value creation model during the following phase.

4.6 Identification of Responsibilities

When the algorithm detects a potential liability case, it is crucial to determine which party in the value chain may be held accountable or affected. Under German law, liability for defects depends on their nature and severity. The manufacturer, retailer, or seller may be held responsible. The primary determining factor, based on legal grounds, is who is primarily responsible and to what extent the affected party was informed. This is detailed information that the algorithm cannot assess, as it is uncertain whether documentation on these aspects even exists. However, the algorithm can determine the product involved, the suppliers and other manufacturers involved in the value creation process, and which party would be responsible in the event of a liability case. As previously stated in Section 4's introduction, this publication and project presentation assume that a documented value chain process exists for each product, including various companies and suppliers involved.

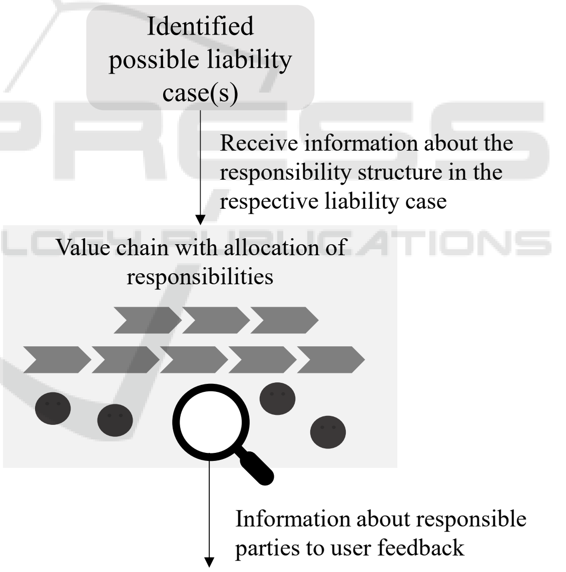

Figure 5: Connecting the identified liability risks to the responsible parties in the value chain.

The responsible parties are linked not by specific individuals, but through the submission of filings made by individual companies involved in the value chain or departments dealing with respective liability cases. The structure of this step can vary depending on the number of companies involved in the value chain and the possibility of duplication for the manufacturer and seller. The primary objective of the assignment is not to undertake a binding allocation of tasks. It is primarily about presenting information in the subsequent edition of the guide, specifying the departments to notify when addressing complaints and the pertinent factors to consider when sharing information. The algorithm's analytical process concludes after this sixth phase, and this data is transmitted to the user in the next and decisive step.

4.7 Assistance for Users

After the algorithm completes its previous phases, it generates a recommendation or information to be sent to the user. The project aims to increase action certainty in complaint handling and processing. Therefore, the main target group consists of employees responsible for processing complaints. The algorithm operates by having an employee process multiple complaints and also check for any corresponding liability risks. The process is completed through an interface where the employee inputs the raw complaint text and receives procedural instructions at the end. These procedural instructions are based on the results determined in the previous six phases of the algorithm and vary based on the number of review loops conducted. If the first phase of customer data determination yields no results, the user receives a corresponding message as detailed in section 4.1. Should a complaint risk be identified, the user is notified with all relevant information at the final customer information:

1. order information
2. date of purchase
3. identified product
4. identified failure
5. liability risks
6. responsible party in processing

This information can provide an initial indication of the severity of the complaint and the triggering failure, as well as next steps. It is important to note that the AlGeWert algorithm is not a substitute for legal expertise, but rather a tool to aid in gathering information. If a liability risk is identified through a complaint, all relevant information is compiled and made available to the appropriate legal department for review, if necessary. Section 5 will present the practical implementation of the algorithm and its current processing status, following a detailed conceptual overview in the previous sections.

5 IMPLEMENTATION AND APPLICABILITY OF THE ALGEWERT PROTOTYPE

The individual components of the AlGeWert algorithm were presented in the previous sections. Once the conceptual phase has been completed, the next step is to implement it in practice and test its functionality. According to the seven-phase concept, the AlGeWert algorithm has been partially implemented as a functional prototype. The practical implementation involves linking multiple software systems that interact with each other. Python serves as the programming environment for transferring the algorithm into a practical application. The initial stage of extracting customer data involves using basic matching mechanisms to compare an exemplar database for the project and its contained information. In this case, a SQL database is utilized, which is managed through the open-source and freely available program, MySQL. One reason for selecting this database is that it can be effortlessly integrated and queried with Python. On the contrary, MySQL is a widely employed software, indicating that the algorithm can be utilized partially in practical scenarios. At present, the database encompasses customer and order information that is easily readable by the algorithm.

Table 1: Exemplary Order Table in Database.

Customer_ID	Order_Date	Product_Name	Product_Number	Order_Number
KN01	2020-04-30	SweepSentry	1014	202009
KN02	2019-01-20	SparkleBot	1006	201902
KN03	2019-10-24	SweepSentry	1014	201917
KN04	2022-07-06	DustDevour	1008	202209
…	…	…	…	…

Table 2: Exemplary Customer Data in Database.

Customer_ID	Last_Name	First_Name	Street_Adress	Post_Code	City	Country
KN01	Mehl	Pascal	Engelskirchener Straße 49	55490	Woppenroth	(DE) Germany
KN02	Gäpfert	Jonas	Birlinghovener Straße 152	55288	Gabsheim	(DE) Germany
KN03	Kläckner	Marietheres	Am Leerberg 37	42879	Remscheid	(DE) Germany
KN04	Schmidt	Emma	Scheibeng 14	1190	Wien	(AT) Austria
…	…	…	…	…	…	…

Both Tables 1 and 2 show sample extracts from the database used. For privacy reasons and to simplify the presentation, only imaginary example data is used here. The databases initially follow a simple structure, and that the customer number is decisive for creating the link between customer and order.

As for the completion of the AlGeWert algorithm, all phases are now being evaluated in detail. Information extraction and the first decision phase are already implemented. The second decision phase can also be considered implemented, since the decision mechanism and the relevant criteria have already been developed and can be stored in the source code. However, the NLP structure has not been fully implemented at this stage of the research. In initial tests, the bag-of-words model has successfully recognized the different types of failure. Systematic validation through corresponding tests is being planned, but cannot yet be published at the current stage of the project. Therefore, the phases of failure detection and the second decision phase have not yet been completed. To enhance the NLP model, we will gather a set of genuine and/or realistic raw complaint texts to train the model. This will be done without any use of machine translation tools. The only phase yet to be implemented is the identification of value chain responsibilities, as the type of failure must be fully functional before proceeding. Moreover, a suitable file format for updating the related value chain processes with data input, without any manual source code changes, must also be identified. The "Assistance for Users" phase is comparable to the second decision phase. The implementation has already occurred in principle. When the algorithm categorizes the existing complaint during the first, second, or third phase, it offers feedback to the user on how to proceed. However, this phase is only regarded as complete if all information is effectively utilized and processed. Hence, the status assigned to this phase is "ongoing."

Figure 6 displays the current stage of development, referencing both the implemented processes that are currently operational and the algorithm phases that have not yet been practically implemented.

A significant part of the algorithm has already been or is being implemented in a Python environment to create a user-friendly application. However, a few more steps need to be implemented before the algorithm is complete and can be validated as a whole.

The previous sections have outlined the structure and practical implementation, as well as the current project status, of the AlGeWert algorithm. Section six will discuss how these findings can be contextualized in scientific literature, detailing both challenges and successes attained.

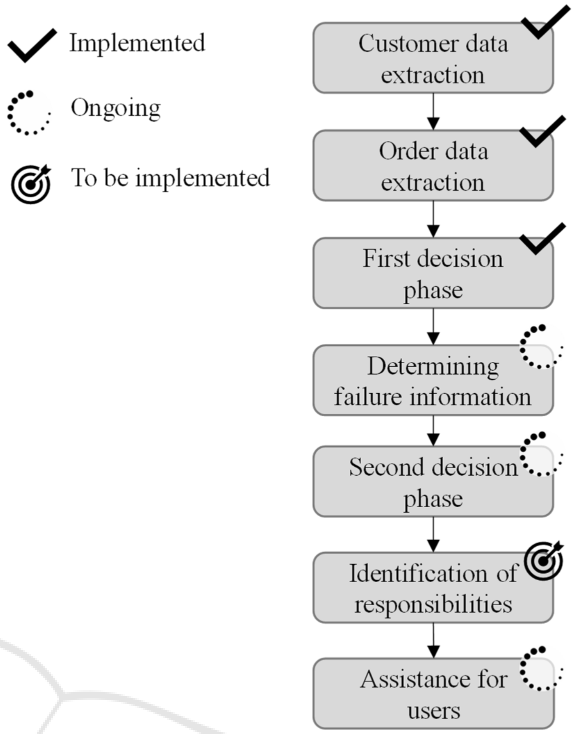

Figure 6: Current implementation status of the AlGeWert phases in practical application.

6 DISCUSSION

The design of the AlGeWert algorithm is intended to provide the user with information about the liability-relevant risk of a complaint by means of a fast and automated text analysis. In the last sections, both the structure and the practical implementation status were presented.

The algorithm is basically feasible and can already provide initial assistance to users. The technical implementation is not yet complete, as one of the most complex aspects - the detection of failure types by the NLP structure - is still under development. Nevertheless, the project team sees the biggest challenge in extracting the right information from legal texts and correctly feeding it to the decision algorithm. A clear demarcation of the terms of use is necessary in this context. It is important to note that the current use case is limited to the internal processing of data based on the liability principles applicable in Germany.

Although it has already been mentioned that the algorithm cannot replace legal expertise anyway, the project aims to produce the most concrete and accurate information possible and to ensure that the algorithm does not produce false statements. For this reason, it has been decided to initially filter according to the objective criteria of "place of residence" and "date of purchase", and to always add a disclaimer when assessing the type of failure, which excludes complete legal certainty.

For the time being, the extraction of information considering legal criteria to distinguish the types of failures can be seen as complete. However, it remains a task within the project to always keep an eye on current developments and changes in the relevant case law and to revise the decision structure in case of doubt.

What is also missing from the algorithm's process flow is the integration of a retrospective process step that allows the information gained from a liability-relevant complaint to be fed back into the product development process. The potential to use such data as a source of information for product development is recognized here; the main task is how to process the information accordingly.

7 CONCLUSIONS AND RESEARCH PROSPECT

This paper presents a concept for the development of an algorithm for the automated evaluation and verification of information with respect to liability risks. After identifying the problem, the current state of the art was reviewed and the lack of possibilities for concrete retrieval of liability-relevant information in complaint processing was highlighted. Subsequently, the conceptual structure of the algorithm was presented, which is divided into the following phases:

1. Customer data extraction
2. Order data extraction
3. First decision phase
4. Determining failure information
5. Second decision phase
6. Identification of responsibilities
7. Assistance for users

The result is an algorithm, already partially implemented, that makes it possible to check complaint texts for liability-relevant content. It does this by comparing information extracted from a written customer complaint with an underlying decision structure, which allows it to differentiate between different liability scenarios based on objective criteria such as "customer's residence" and "date of purchase". An NLP framework is used to implement a decision structure that identifies failure information and categorizes the type of failure. This links the relevant legal bases to the given complaint. The algorithm is designed to suggest user instructions to employees working with it, ensuring fast and efficient handling of customer complaints. For further prospects, the research project aims to implement and complete the failure detection in such a way that it can be used reliably. This requires testing the NLP model with sufficient data and validating it through systematic tests under the same conditions. Integrating the value chain into work processes is necessary to ensure the algorithm's full functionality.

ACKNOWLEDGEMENTS

The authors appreciate the German Research Foundation (DFG) for the support of the Project "AlGeWert" (funding code: SCHL 2225/7-1).

REFERENCES

Agren, S. M., Knauss, E., Heldal, R., Pelliccione, P., Malmqvist, G. & Boden, J. (2019). The impact of requirements on systems development speed: a multiple-case study in automotive. *Requirements Engineering*, *24*(3), 315–340. https://doi.org/10.1007/s00766-019-00319-8

Anagun, Y., Bolel, N. S., Isik, S. & Ozkan, S. E. (2022). Deep Learning-Based Customer Complaint Management. *Journal of Organizational Computing and Electronic Commerce*, *32*(3-4), 217–231. https://doi.org/10.1080/10919392.2023.2210049

Behrens, B.-A., Wilde, I. & Hoffmann, M. (2007). Complaint management using the extended 8D-method along the automotive supply chain. *Production Engineering*, *1*(1), 91–95. https://doi.org/10.1007/s11740-007-0028-6

Cieśla, M. (2023). Complaint management system in building material factory. *Management and Production Engineering Review*. https://doi.org/10.24425/mper.2019.128243

Hake, P., Rehse, J.-R. & Fettke, P. (2021). Toward Automated Support of Complaint Handling Processes: An Application in the Medical Technology Industry. *Journal on Data Semantics*, *10*(1-2), 41–56. https://doi.org/10.1007/s13740-021-00124-z

Hardin, W. (2015, 23. Juli). Minimizing Liability Risk in Engineering Product Design. *GlobalSpec*. Verfügbar unter: https://insights.globalspec.com/article/1257/minimizing-liability-risk-in-engineering-product-design

Hegde, V. (2023). Improve Warranty Failures in Original Equipment Manufacturing via Design for Reliability. In

2023 Annual Reliability and Maintainability Symposium (RAMS) (S. 1–4). IEEE.

Schmitt, R. & Linder, A. (2013). Technical complaint management as a lever for product and process improvement. *CIRP Annals*, *62*(1), 435–438. https://doi.org/10.1016/j.cirp.2013.03.040

Stauss, B. & Seidel, W. (2019). *Effective complaint management. The business case for customer satisfaction* (Management for professionals, Second edition). Cham: Springer. https://doi.org/10.1007/978-3-319-98705-7

Yilmaz, C., Varnali, K. & Kasnakoglu, B. T. (2016). How do firms benefit from customer complaints? *Journal of Business Research*, *69*(2), 944–955. https://doi.org/10.1016/j.jbusres.2015.08.038

Zaby, C. & Wilde, K. D. (2018). Intelligent Business Processes in CRM. *Business & Information Systems Engineering*, *60*(4), 289–304. https://doi.org/10.1007/s12599-017-0480-6