INTERNATIONAL JOURNAL OF COMPUTER SCIENCE & INFORMATION SYSTEM

Volume10 Issue06, June-2025, pg. 1-9

Published Date: - 01-06-2025

E-ISSN: 2536-7919
P-ISSN: 2536-7900
SJIF 2019: 4.58 2020: 5.046 2021: 5.328

Integrating Failure Mode and Effect Analysis with Technical Language Processing for Enhanced Maintenance Insights

Dr. Radmila Juric

Senior Lecturer in Computer Science, University of Westminster, UK

Abstract: In technical domains, particularly asset management and maintenance, vast amounts of valuable information are captured in unstructured text formats such as work orders, maintenance logs, and inspection reports [1, 4, 28, 30]. Extracting actionable insights from this "technical language" is challenging using traditional methods [1, 15]. Concurrently, Failure Mode and Effect Analysis (FMEA) is a widely adopted systematic approach for identifying potential failure modes within a system, analyzing their causes and effects, and prioritizing risks [2, 6]. While powerful, traditional FMEA is often a manual and labor-intensive process, relying on expert knowledge and structured data [8]. This article explores the integration of FMEA principles with Technical Language Processing (TLP) techniques to automate and enhance the extraction, analysis, and utilization of failure-related information from unstructured technical text. We propose a conceptual framework where TLP methods are employed to identify potential failure modes, effects, and causes within technical documentation. These extracted elements can then be structured and analyzed using FMEA principles to provide data-driven insights into asset reliability and maintenance needs [5]. This integration has the potential to improve the efficiency and effectiveness of FMEA, unlock hidden knowledge within maintenance data, and support proactive maintenance strategies.

Keywords: Failure Mode and Effect Analysis (FMEA), Technical Language Processing (TLP), Natural Language Processing (NLP), Maintenance, Asset Management, Reliability, Text Mining, Data Analytics.

INTRODUCTION

Modern industrial and infrastructure systems generate enormous volumes of data, a significant portion of which exists in the form of unstructured text [1, 4, 28, 30]. Within asset management and maintenance operations, this includes detailed descriptions of equipment failures, maintenance actions performed, observed anomalies, and inspection findings, often recorded in maintenance work orders and operational logs [4, 31, 44]. This "technical language" contains critical knowledge about the behavior of assets, common failure patterns, and the effectiveness of maintenance interventions [1, 15]. However, the

Volume10 Issue06, June-2025, pg. 1-9

Published Date: - 01-06-2025

E-ISSN: 2536-7919
P-ISSN: 2536-7900
SJIF 2019: 4.58 2020: 5.046 2021: 5.328

unstructured nature of this text makes it difficult to systematically analyze and leverage this information for improving reliability and maintenance strategies [44, 45].

Natural Language Processing (NLP) is a field of artificial intelligence concerned with enabling computers to understand, interpret, and generate human language [9, 10, 11, 12, 13, 14]. While traditional NLP has focused heavily on general language, Technical Language Processing (TLP) is an emerging subfield specifically adapted for the unique characteristics of technical documentation, which often involves specialized terminology, jargon, abbreviations, and a less formal grammatical structure compared to standard text [1, 15]. Adapting NLP techniques for technical text presents specific challenges [15].

Failure Mode and Effect Analysis (FMEA) is a foundational methodology in reliability engineering and risk management [2, 6]. It is a systematic, proactive method for identifying potential single points of failure in a system or process, determining the causes and effects of these failures, and evaluating the risk associated with each failure mode based on its severity, occurrence, and detection [2, 8]. FMEA is widely used across various industries to improve product design, manufacturing processes, and maintenance planning [7, 8]. Despite its value, traditional FMEA can be a time-consuming and resource-intensive process, often relying on expert workshops and manual data analysis [8]. The effectiveness of FMEA is also dependent on the availability and quality of structured data, which is often lacking in real-world maintenance environments [4, 54].

The confluence of the increasing availability of digital maintenance records and advancements in TLP presents an opportunity to bridge the gap between unstructured technical text and structured reliability analysis methods like FMEA. By applying TLP techniques to technical documentation, it may be possible to automatically extract key information related to failure modes, their causes, and their effects, which are the core components of an FMEA [55]. This integration has the potential to significantly enhance the efficiency, scalability, and data-driven nature of FMEA in maintenance contexts.

This article proposes a framework for integrating FMEA principles with TLP techniques to unlock valuable insights from unstructured technical maintenance data. We discuss how TLP can be used to automate the identification and structuring of FMEA elements from text, the potential benefits of this integrated approach for maintenance and reliability, and the challenges and future directions in this emerging area.

METHODS

(Note: As this is a conceptual article outlining a methodology rather than presenting experimental results, the "Methods" section describes the proposed framework and the TLP techniques that can be applied, rather than detailing experimental procedures.)

The proposed methodology integrates principles of Failure Mode and Effect Analysis (FMEA) with Technical Language Processing (TLP) techniques to extract and structure failure-related information from

Volume10 Issue06, June-2025, pg. 1-9

Published Date: - 01-06-2025

E-ISSN: 2536-7919
P-ISSN: 2536-7900
SJIF 2019: 4.58 2020: 5.046 2021: 5.328

unstructured technical text, primarily focusing on maintenance work orders and logs. The process involves several key steps:

2.1. Data Collection and Preprocessing: The primary data source consists of unstructured text from maintenance records, such as work order descriptions, technician notes, and failure reports [4, 31, 44]. This raw text data often contains noise, inconsistencies, abbreviations, and domain-specific jargon [1, 15, 45]. Preprocessing steps are crucial to prepare the text for TLP analysis. These include:

- Text Cleaning: Removing irrelevant characters, symbols, and formatting.
- Tokenization: Splitting the text into individual words or sub-word units [10, 20].
- Normalization: Handling variations in spelling, capitalization, and abbreviations [45].
- Stop Word Removal: Eliminating common words that do not carry significant meaning in the technical context.
- Stemming or Lemmatization: Reducing words to their root form to group related terms [20].

2.2. Identification of FMEA Elements using TLP: The core of the proposed method involves using TLP techniques to automatically identify and extract the key components of an FMEA from the preprocessed text:

- Failure Mode Identification: TLP techniques, such as Named Entity Recognition (NER) and keyword extraction, can be trained or configured to identify phrases or terms that describe how an asset or component fails [46, 47]. This might involve recognizing specific failure symptoms or observed malfunctions [48]. Domain-specific dictionaries and ontologies can be used to guide this process [15].
- Effect Analysis: TLP can be used to identify the consequences or impacts of the identified failure modes. This might involve analyzing the text for descriptions of operational disruptions, safety hazards, environmental damage, or repair actions taken [30, 49, 50]. Relation extraction techniques can be employed to identify causal relationships between failure modes and their effects [30].
- Cause Analysis: TLP can also be applied to extract the underlying reasons or root causes of the failure modes. This could involve identifying mentions of faulty parts, operational errors, environmental factors, or wear and tear [30, 49]. Again, relation extraction can be valuable in linking causes to specific failure modes.

Advanced TLP techniques, including those based on deep learning architectures like Transformers (e.g., BERT, GPT) [26, 33, 38, 39, 40, 41, 42, 43], can be particularly effective in capturing the semantic relationships and context within technical text, improving the accuracy of identifying FMEA elements [27, 55]. Transfer learning from models pretrained on large text corpora can be beneficial, even with limited labeled technical data [17, 55].

Volume10 Issue06, June-2025, pg. 1-9

Published Date: - 01-06-2025

E-ISSN: 2536-7919
P-ISSN: 2536-7900
SJIF 2019: 4.58 2020: 5.046 2021: 5.328

2.3. Structuring and Prioritization based on FMEA Principles: Once the failure modes, effects, and causes are extracted, they can be structured into a format compatible with FMEA. This involves:

- Mapping to FMEA Categories: Assigning the extracted information to the relevant FMEA columns (Failure Mode, Effect of Failure, Cause of Failure).

- Risk Prioritization: While traditional FMEA uses expert-assigned severity, occurrence, and detection ratings [2], TLP can potentially provide data-driven input for these factors. For example:

  - Occurrence: The frequency of mentions of a specific failure mode or cause in the maintenance logs can provide an indication of its occurrence rate [48].

  - Severity: The language used to describe the effects (e.g., "major disruption," "safety critical") could be analyzed using sentiment analysis or severity scoring techniques to estimate severity [51].

  - Detection: Information about how the failure was identified (e.g., "operator noticed," "alarm triggered") might provide insights into detection [31].

Alternatively, TLP can be used to extract information that helps human experts assign these ratings more consistently and efficiently. The resulting structured data can then be used to calculate a Risk Priority Number (RPN) or other risk metrics [2, 8].

2.4. Analysis and Reporting: The structured FMEA data derived from TLP can be analyzed to identify high-risk failure modes, common causes, and frequent effects. This analysis can inform various maintenance and reliability activities, including:

- Updating FMEA Documents: Automatically populating or updating FMEA worksheets with real-world data from maintenance logs [55].

- Identifying Emerging Failure Modes: Detecting new or previously unanticipated failure modes based on patterns in the text data.

- Improving Maintenance Procedures: Using insights into common causes and effective repairs mentioned in the logs to refine maintenance tasks.

- Predicting Maintenance Needs: Leveraging the extracted information to build predictive models for equipment failures [32].

- Benchmarking and Performance Monitoring: Comparing failure patterns and frequencies across similar assets or sites [53].

Volume10 Issue06, June-2025, pg. 1-9

Published Date: - 01-06-2025

E-ISSN: 2536-7919
P-ISSN: 2536-7900
SJIF 2019: 4.58 2020: 5.046 2021: 5.328

Tools and libraries for NLP and TLP, such as NLTK [20], spaCy [21], Gensim [22], TextBlob [23], Stanford CoreNLP [24], AllenNLP [25], and the Hugging Face Transformers library [26], can be utilized to implement the TLP steps of this methodology.

RESULTS

(Note: As this is a conceptual article outlining a methodology, this section describes the anticipated outcomes and benefits of applying the proposed integrated approach, rather than presenting specific experimental data.)

Implementing the proposed integration of FMEA principles with TLP techniques is expected to yield significant improvements in the analysis of technical maintenance data and the effectiveness of reliability programs. The primary anticipated results include:

- Automated Extraction of FMEA Elements: TLP techniques are expected to successfully identify and extract key failure modes, effects, and causes from unstructured maintenance text with a measurable level of accuracy. This automation will significantly reduce the manual effort currently required for this task in traditional FMEA [55].

- Creation of Data-Driven FMEA Inputs: The analysis of extracted information is anticipated to provide quantitative or semi-quantitative data points related to the occurrence, severity, and detection of failure modes. This data-driven input can complement or potentially replace subjective expert estimations, leading to more objective and consistent risk prioritization [48, 51].

- Enhanced Identification of Failure Patterns: By processing large volumes of historical maintenance data, the integrated approach is expected to reveal recurring failure modes and common causal factors that might not be readily apparent through manual review or analysis of structured data alone [44, 47]. This includes identifying subtle or emerging failure patterns.

- Improved Maintenance Data Quality: The process of applying TLP and structuring the extracted information can highlight inconsistencies, ambiguities, and errors in the original text data [45, 54], providing valuable feedback for improving data entry practices and overall data quality [4].

- Support for Proactive Maintenance: The insights gained from the automated FMEA analysis can directly support the development and refinement of proactive maintenance strategies, such as optimizing preventive maintenance schedules, implementing condition monitoring, and improving spare parts management [5, 7].

- Faster and More Scalable Analysis: Automating the extraction and initial analysis of FMEA elements through TLP will enable faster processing of large datasets and facilitate the application of FMEA principles across a wider range of assets and systems [55].

Volume10 Issue06, June-2025, pg. 1-9

Published Date: - 01-06-2025

E-ISSN: 2536-7919
P-ISSN: 2536-7900
SJIF 2019: 4.58 2020: 5.046 2021: 5.328

• Unlocking Hidden Knowledge: The integrated approach is expected to unlock valuable knowledge embedded within the free-text descriptions of maintenance activities, making this information accessible for reliability analysis and decision-making [1, 30]. This includes insights into the context of failures and the specific actions taken by technicians [28].

Examples from related work in applying NLP to maintenance data, such as categorizing maintenance work orders [4], extracting keywords [46], identifying fault frequencies [47], predicting maintenance time [32], and analyzing defect logs [29], support the feasibility and potential impact of applying TLP for FMEA element extraction. Recent studies on leveraging BERT for maintenance text analysis further demonstrate the applicability of advanced TLP models in this domain [55, 56, 57].

DISCUSSION

The integration of Failure Mode and Effect Analysis (FMEA) with Technical Language Processing (TLP) offers a compelling approach to enhance reliability and maintenance practices in technical domains. By automating the extraction and structuring of failure-related information from unstructured text, this methodology addresses key limitations of traditional FMEA and unlocks the vast potential of historical maintenance data.

One of the primary benefits of this integrated approach is the ability to move towards a more data-driven FMEA. While expert knowledge remains invaluable, supplementing it with insights derived directly from real-world operational and maintenance data can lead to more accurate risk assessments and prioritization [5, 55]. The frequency of reported failures in maintenance logs can provide a more objective measure of occurrence compared to subjective estimations [48]. Similarly, analyzing the language used to describe the consequences of failures can provide a data-informed basis for estimating severity [51].

Furthermore, this approach can significantly improve the efficiency and scalability of FMEA. Manually reviewing and analyzing thousands of maintenance records to identify failure modes and effects is a time-consuming task [44]. Automating this process with TLP allows for the analysis of much larger datasets, potentially enabling the application of FMEA principles to a broader range of assets and at a higher frequency [55]. This is particularly relevant for organizations managing large portfolios of assets [53].

The integration also facilitates the identification of emerging or subtle failure modes that might be missed in traditional FMEA or structured data analysis. Free-text descriptions often contain nuanced information about how assets are behaving that can serve as early indicators of potential issues. TLP techniques capable of capturing semantic meaning and context are well-suited for identifying these patterns [27, 30].

However, several challenges need to be addressed for successful implementation. The quality and consistency of the unstructured text data are critical [4, 45, 54]. Variations in terminology, abbreviations, and the level of detail provided by different technicians can impact the accuracy of TLP extraction [15].

Volume10 Issue06, June-2025, pg. 1-9

Published Date: - 01-06-2025

E-ISSN: 2536-7919
P-ISSN: 2536-7900
SJIF 2019: 4.58 2020: 5.046 2021: 5.328

Domain adaptation of general NLP models to the specific language of maintenance is also necessary [15, 17]. This often requires domain-specific training data and expertise in both TLP and maintenance.

Another challenge lies in accurately interpreting the context and nuances within the technical text. For example, distinguishing between a reported failure mode and a successful repair action requires sophisticated TLP capabilities. Resolving ambiguity and handling negations or conditional statements are also important considerations.

Future research should focus on developing more robust and domain-adaptive TLP models for technical text. Exploring advanced techniques for relation extraction and causal analysis from unstructured data can further enhance the ability to link failure modes, effects, and causes. Developing methods for automatically estimating FMEA risk parameters (severity, occurrence, detection) directly from text data, potentially using techniques like reinforcement learning or expert-guided machine learning, is another promising area. Finally, validating the effectiveness of this integrated approach through case studies and empirical evaluations on real-world maintenance datasets is crucial to demonstrate its practical value [55].

The integration of FMEA with TLP represents a significant step towards leveraging the wealth of information contained within unstructured maintenance data for improved reliability and asset management. By transforming raw text into structured, actionable insights, this approach can empower organizations to make more informed decisions, optimize maintenance strategies, and ultimately enhance the performance and longevity of their assets.

CONCLUSION

Unstructured technical text within maintenance records holds invaluable information regarding asset failure modes, causes, and effects. Traditional methods for reliability analysis, such as FMEA, while systematic, often struggle to effectively utilize this data due to its unstructured nature. This article proposed a conceptual framework for integrating the principles of FMEA with Technical Language Processing (TLP) techniques to automate the extraction and structuring of failure-related information from technical text. By employing TLP for identifying and extracting failure modes, effects, and causes, and then applying FMEA principles for structuring and prioritization, organizations can gain data-driven insights into asset reliability. This integrated approach has the potential to significantly improve the efficiency, scalability, and effectiveness of FMEA in maintenance contexts, unlock hidden knowledge within maintenance data, and support the development of more proactive and data-informed maintenance strategies. While challenges related to data quality and TLP model adaptation exist, the potential benefits make this an important area for continued research and development.

REFERENCES

1. Brundage, M.P.; Sexton, T.; Hodkiewicz, M.; Dima, A.; Lukens, S. Technical language processing: Unlocking maintenance knowledge. Manuf. Lett. 2021, 27, 42–46.

Volume10 Issue06, June-2025, pg. 1-9

Published Date: - 01-06-2025

E-ISSN: 2536-7919
P-ISSN: 2536-7900
SJIF 2019: 4.58 2020: 5.046 2021: 5.328

2. International Electrotechnical Commission. IEC 60812: Failure Modes and Effects Analysis (FMEA and FMECA), Edition 3.0; International Electrotechnical Commission: Geneva, Switzerland, 2018.
3. Unsworth, K.; Adriasola, E.; Johnston-Billings, A.; Dmitrieva, A.; Hodkiewicz, M. Goal hierarchy: Improving asset data quality by improving motivation. Reliab. Eng. Syst. Saf. 2011, 96, 1474–1481.
4. Sexton, T.; Hodkiewicz, M.; Brundage, M.P. Categorization errors for data entry in maintenance work-orders. In Proceedings of the Annual Conference of the PHM Society, Scottsdale, AZ, USA, 2–5 September 2019; Volume 11.
5. Payette, M.; Abdul-Nour, G. Asset Management, Reliability and Prognostics Modeling Techniques. Sustainability 2023, 15, 7493.
6. Zio, E. An Introduction to the Basics of Reliability and Risk Analysis; Series on Quality, Reliability and Engineering Statistics; World Scientific: Singapore, 2007.
7. Islam, H. Reliability-centered maintenance methodology and application: A case study. Engineering 2010, 2, 863–873.
8. Boral, S.; Howard, I.; Chaturvedi, S.K.; McKee, K.; Naikan, V.N.A. An integrated approach for fuzzy failure modes and effects analysis using fuzzy AHP and fuzzy MAIRCA. Eng. Fail. Anal. 2020, 108, 104195.
9. Joseph, S.R.; Hlomani, H.; Letsholo, K.; Kaniwa, F.; Sedimo, K. Natural language processing: A review. Int. J. Res. Eng. Appl. Sci. 2016, 6, 207–210.
10. Jurafsky, D. Speech & Language Processing; Pearson Education India: Hoboken, NJ, USA, 2000.
11. Nadkarni, P.M.; Ohno-Machado, L.; Chapman, W.W. Natural language processing: An introduction. J. Am. Med. Inform. Assoc. 2011, 18, 544–551.
12. Thanaki, J. Python Natural Language Processing; Packt Publishing Ltd.: Birmingham, UK, 2017.
13. Bansal, A. Advanced Natural Language Processing with TensorFlow 2: Build Effective Real-World NLP Applications Using NER, RNNs, Seq2Seq Models, Transformers, and More; Packt Publishing Ltd.: Birmingham, UK, 2021.
14. Beysolow, T., II. Applied Natural Language Processing with Python: Implementing Machine Learning and Deep Learning Algorithms for Natural Language Processing; Apress: New York, NY, USA, 2018.
15. Dima, A.; Lukens, S.; Hodkiewicz, M.; Sexton, T.; Brundage, M.P. Adapting natural language processing for technical text. Appl. AI Lett. 2021, 2, e33.
16. Fort, K. Collaborative Annotation for Reliable Natural Language Processing: Technical and Sociological Aspects; John Wiley & Sons: Hoboken, NJ, USA, 2016.
17. Yang, X.; He, X.; Liang, Y.; Yang, Y.; Zhang, S.; Xie, P. Transfer learning or self-supervised learning? A tale of two pretraining paradigms. arXiv 2020, arXiv:2007.04234.
18. De Marneffe, M.C.; Manning, C.D.; Nivre, J.; Zeman, D. Universal dependencies. Comput. Linguist. 2021, 47, 255–308.
19. Nivre, J.; De Marneffe, M.C.; Ginter, F.; Goldberg, Y.; Hajic, J.; Manning, C.D.; McDonald, R.; Petrov, S.; Pyysalo, S.; Silveira, N.; et al. Universal dependencies v1: A multilingual treebank collection. In Proceedings of the Tenth International Conference on Language Resources and Evaluation (LREC'16), Portorož, Slovenia, 23–28 May 2016; pp. 1659–1666.

Volume10 Issue06, June-2025, pg. 1-9

Published Date: - 01-06-2025

E-ISSN: 2536-7919
P-ISSN: 2536-7900
SJIF 2019: 4.58 2020: 5.046 2021: 5.328

20. Bird, S.; Klein, E.; Loper, E. Natural Language Processing with Python: Analyzing Text with the Natural Language Toolkit; O'Reilly Media, Inc.: Sebastopol, CA, USA, 2009.

21. Honnibal, M.; Johnson, M. An improved non-monotonic transition system for dependency parsing. In Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing, Lisbon, Portugal, 17–21 September 2015; pp. 1373–1378.

22. Řehůřek, R.; Sojka, P. Software Framework for Topic Modelling with Large Corpora. In Proceedings of the LREC 2010 Workshop on New Challenges for NLP Frameworks, Valletta, Malta, 22 May 2010; pp. 45–50.

23. Loria, S. TextBlob: Simplified Text Processing. 2018. Version 0.15.1. Available online: https://textblob.readthedocs.io/ (accessed on 2 February 2024).

24. Manning, C.D.; Surdeanu, M.; Bauer, J.; Finkel, J.R.; Bethard, S.; McClosky, D. The Stanford CoreNLP natural language processing toolkit. In Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics: System Demonstrations, Baltimore, MD, USA, 23–24 June 2014; pp. 55–60.

25. Gardner, M.; Grus, J.; Neumann, M.; Tafjord, O.; Dasigi, P.; Liu, N.F.; Peters, M.; Schmitz, M.; Zettlemoyer, L.S. AllenNLP: A Deep Semantic Natural Language Processing Platform. arXiv 2017, arXiv:1803.07640.

26. Wolf, T.; Debut, L.; Sanh, V.; Chaumond, J.; Delangue, C.; Moi, A.; Cistac, P.; Rault, T.; Louf, R.; Funtowicz, M.; et al. Transformers: State-of-the-art natural language processing. In Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing: System Demonstrations, Online, 16–20 November 2020; pp. 38–45.

27. Liu, Z.; Lin, Y.; Sun, M. Representation Learning for Natural Language Processing; Springer Nature: Berlin/Heidelberg, Germany, 2020.

28. Giordano, V.; Fantoni, G. Decomposing maintenance actions into sub-tasks using natural language processing: A case study in an Italian automotive company. Comput. Ind. 2025, 164, 104186.

29. Li, H.; Deng, F.; Lu, J.; Zhang, T.; Li, H. An Application of Automatic Text Revision for Power Defect Log. J. Phys. Conf. Ser. 2021, 1757, 012027.