2.1

AI-Based Knowledge Management System for Risk Assessment and Root Cause Analysis in Semiconductor Industry

Houssam Razouk[1,4], Roman Kern[3,4], Martin Mischitz[1], Josef Moser[1], Mirhad Memic[1], Lan Liu[3], Christian Burmer[2] and Anna Safont-Andreu[1]

[1]Infineon Technologies Austria AG, Austria
[2]Infineon Technologies AG, Germany
[3]Know-Center GmbH, Austria
[4]Graz University of Technology, Austria

Abstract

Due to the increasing technical complexity of products and market pressure, the demands in the semiconductor industry are rising with respect to quality, performance, and time to market. Root cause analysis and risk assessment are crucial elements for success in fulfilling these demands. As a result, there is an ever-growing number of technical documents, which potentially contain valuable information serving as a base to inform development and production. Experts need to cope with this large number of technical documents, for example, to generate new hypotheses to identify possible root causes of deviations or potential risks in the ramp-up and production phase of new products. Unfortunately, most of the technical documents are unstructured, making processing them even more tedious. New advances in computer science, specifically artificial intelligence (AI), open the door for a higher degree of automation of knowledge management tools to support experts. Knowledge bases such as knowledge graphs allow for representing complex information but need to be created for each domain. Novel state-of-the-art graph embedding algorithms showed promising results in complementing knowledge bases with new relations. Complementary to knowledge base completion, language models trained on large textual corpora have demonstrated their ability to capture complex semantics. This paper proposes a new expert system concept for failure root cause analysis and risk assessment in the semiconductor industry, which leverages the advanced graph embeddings in combination with language models. The main challenges in this setting are the type of relations of interest, which are causal, and the language being used, which is highly domain-specific. Thus, we devised AI for consistency improvement of the data, predicting new links, and information extraction from unstructured data. The information extraction is conducted by levaraging domain specific ontologies and by focusing on presence of causal language.

Keywords: expert system, root cause analysis and risk assessment, knowledge representation, semiconductor industry, natural language processing, information extraction, knowledge graph, convolutional neural network, recurrent neural network, machine learning, link prediction, text classification, consistency improvement.

2.1.1 Introduction and Background

In the last decades, more than ever, high-tech microelectronic-based products consolidate as part of everyday life. Thus, expectations concerning functionality, reliability, and competitive prices are growing. As a response, more functions are integrated, facilitating products' performance continuous growth. Consequently, the technological complexity of microelectronic components and the amount of data are constantly increasing due to Industry 4.0 applications in the production facilities. Moreover, the fierce, competitive market situation for the industrial semiconductor companies, which are the leading supplier of such high-tech products, is inevitably increasing the time to market and price pressure. Therefore, knowledge and experience are necessary to enable innovation, stable production, and cope with market dynamics.

In the semiconductor manufacturing industry, knowledge and experience refer to domain-specific know-how in chip design, operation and control of highly sophisticated infrastructure, metrology, quality assurance, verification, and validation. An effective and efficient knowledge management system that, on-demand, applies and rolls out existing know-how and allows rapid learning from the failures has to be in place.

The necessary knowledge is usually accumulated over long periods and reflects in the practical experience of the human domain experts. It is common to document human domain experts' knowledge in the written form using the domain-specific language. One of the main challenges is how to make this knowledge continuously more accessible to all potential users, respectively, i.e., the engineering teams working in semiconductor manufacturing.

To guarantee the high quality of the products in the semiconductor industry, human domain experts thoroughly investigate deviations in the manufacturing process or in the products' characteristics. Mainly, two standard processes triggered after deviation identification to answer causal questions: (i) risk assessment: what will be the effect of the observed deviation? (ii) root cause analysis: what is the cause of the observed deviation? Therefore, it is highly intriguing to identify causal relations along the whole production process.

In the semiconductor manufacturing with many hundred subsequent process steps, it is effort-intensive and time-consuming to keep up with the detailed information required for successful semiconductor manufacturing.

The following chapters describe our concept system, which leverages recent development in computer science and artificial intelligence for automated information extraction methods from relevant text documents transferring it into a form that allows the inference of additional causal relationships.

2.1.2 Research Areas

This chapter proposes a knowledge management system for risk assessment and root cause analysis in the semiconductor industry. In specific, this chapter discusses the various system components and functionalities. Lastly, this chapter highlights the different challenges and research areas addressed for the proposed system's use cases.

The defined use case of risk assessment and root cause analysis relies on information about previous experiences. This information is documented in different data sources. The human experts' abilities to interpret various data sources, extract information, and intelligently combine the information, formulating hypotheses, are the fundamental motivation of the proposed system.

To address the motivation mentioned above, we opted for a star schema system design. The knowledge representation is the core component of the systems. The knowledge representation is responsible for the storage of

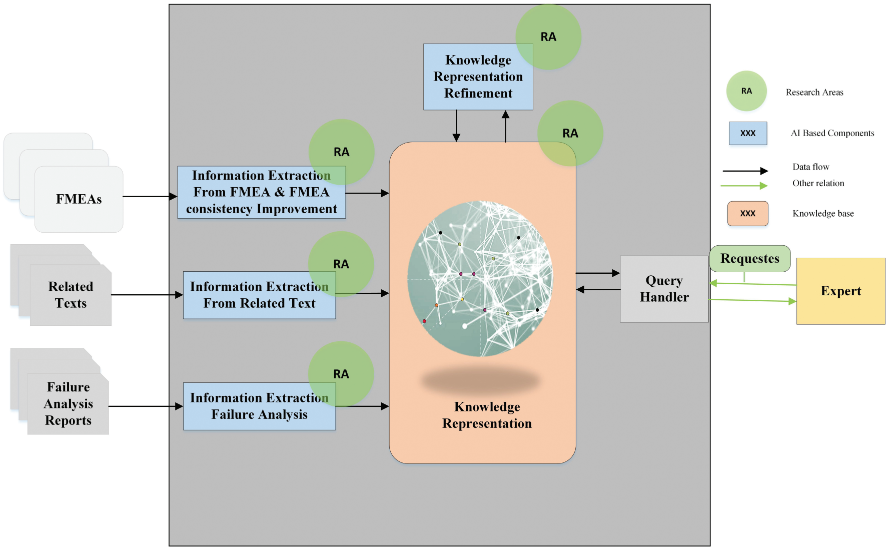

Figure 2.1.1 AI-based knowledge management concept system for risk assessment and root cause analysis in semiconductor industry.

information extracted from different sources. The failure mode effect analysis (FMEA) documents library, the Failure Analysis reports (FA reports), and other related text documents were selected as primary data sources for the proposed system. Different information extraction algorithms to account for the heterogeneity of the data sources were incorporated.

Finally, the refinement algorithms interact with the knowledge representation predicting new links that indicate causal relations, which are not present in the data sources. This is achieved by combining information extracted from different sources in an intelligent way that mimics the human experts' ability for reasoning and inference.

Figure 2.1.1 illustrates the general structure for the proposed system concept.

The proposed system concept is implemented as a demonstrator that uses three different types of documents as the data sources. However, the proposed system can be extended to accommodate more types of data sources.

Several challenges were identified and addressed as follows. The article addresses first the consistency improvement of the selected FMEA documents. Then, it presents the information extraction from FA reports. Third, the information extraction from free text is analysed. Fourth, the knowledge representation of causal information that enables reasoning and inference is described, and finally, the refinement algorithm is addressed.

2.1.2.1 FMEA and FMEA Consistency Improvement

The well-known FMEA (failure mode effect analysis) process is performed by identifying a process map and visualizing components of the process with a multidisciplinary team. At each step in the process map, challenges that may lead to errors or potentially unsafe conditions are identified. Each of the ways that the workflow can fail is called a failure mode. A list of these failure modes and their possible causes and effects is compiled [1–3], formulating an FMEA document. In FMEA documents, relations between the columns (Effect/ Failure mode/ Root cause) are interpreted as causal relations, where designated cells in the same row are causally related, effectively representing a causal chain that identifies a single, known risk. The FMEA process supports the experts to document possible failure mechanisms reaching from the effects on the (final) product back to its potential root causes. Ideally, in a given cell, a short, descriptive text represents a single concept. In this case, a concept is a separable (identifiable) phenomenon that acts as either as (i) a root cause, (ii) an effect observed in the product characteristics, or (iii) an intermediate state in the causal chain. Since the FMEA documents are manually compiled by domain experts and the complex nature of causal relations (e.g., many to many relations with transitive perception [4]), inconsistencies are likely to occur. Crucially, any ambiguity anywhere in the FMEA documents will ultimately affect the whole causal chain due to the nature of causal relations. This is worsened by the collaborative manner in how such documents are being crafted. Typically, a group of experts from different fields (e.g., physicists, technology experts, designers) work together to formalize FMEAs, following techniques like brainstorming meetings and workshops. Due to the heterogeneity of these groups and their vastly different scopes, divergent interpretations of individual causal roles (i.e., the respective cells) are more likely to occur, contributing to the inconsistencies experienced in FMEA documents. Based on the analysis of existing FMEA documents, the majority of data inconsistencies is attributed to one of three main categories:

1. Cases of conflict in the direction of the relations, i.e., the direction of the causal effect relation, are reversed.
2. Cases of merged cells, wherein the short text of a single cell comprises multiple concepts or even relations between numerous concepts, e.g., a causal chain of multiple causes and effects.

3. Cases of missing information in the causal chain, where the documented relation describes a subsequent effect of the cause but skips its direct effect.

Additionally, anyone not closely familiar with the production process will struggle to interpret the documents, not only because of the presence of inconsistencies as mentioned above but also due to the language used in the short text, containing domain-specific terms including many abbreviations. To sum up, FMEA documents in their current state are designed to be created and maintained by experts and to be interpreted exclusively by experts. However, automatic data analysis methods lack the ability to correct for impairments in data quality [5], with merged and mixed-up data being a prime example. As such, methods for improving FMEAs data consistency and quality are highly required.

In our research, we found that just addressing the consistency purely via data-driven methods on a document level does not alleviate the problem, i.e., building a classifier to detect the content of the columns. Hence, domain knowledge needs to be considered to define a classification schema that mimics human domain experts' perception of the short text. As a response, we systematically defined: (i) a domain-specific model, consisting of concepts schema (i.e., types of cause/effects) and a relationship consistency scheme (i.e., valid relations between the concepts), and (ii) a model of inconsistencies, consisting of reverse direction of the causal relations, missing information, and merged cells. Based on real-world data, we found that the case of merged cells is in fact, the most prevalent cause of inconsistency. Moreover, the manual classification of the complete documents' library is time-consuming. Thus, our research provides a systematic study that addresses the effectiveness of variant AI-based approaches for short text classification in the semiconductors industry where domain-specific language is used. Also, our research extends the intersentential pattern mining algorithm presented in [6] to address the cases of merged cells.

Our research aims to transform the FMEAs from their current state of "experts only" to a more machine-friendly form that could achieve a higher automation degree of expert systems' methods for root cause analysis and risk assessment.

2.1.2.2 Causal Information Extracting from Free Text

Our research in causal extraction from the free text, contained in documents with no predefined structures (unstructured documents), was initiated via an initial literature survey. As a result, it has been found that there are two main approaches, namely: (1) rule-based approaches, (2) machine-learning-based approaches. The first category is based on several hypotheses, which are briefly outlined below.

Hume's comments on causal relation: Based on the observation that cause and effect often co-occur and thus have a higher likelihood to be part of a causal relationship [7].

Transitivity: Preserving transitivity is an essential desideratum for an adequate analysis of causation [8]. For example, if e_i causes e_j, and e_j causes e_k, then the transitivity states that e_i causes e_k. Moreover, if there is another cause for e_j, e.g. e_l, it also follows that e_l causes e_k. This property is beneficial in a textual setting since causal statements are expected to be infrequent.

Suppes' probabilistic theory of causation [9]: If entity e_i causes e_j, then we will likely observe that the conditional probability given e_i exceeds the marginal probability of e_j: $P(e_j \mid e_i) > P(e_j)$.

In addition to these hypotheses, several lexical cues are indicative of causal relationships. Radinsky et al. [10] propose three types of lexical cues: (1) causal connectives (e.g., because, as, and after), (2) causal preposition (e.g., due to, because of), (3) periphrastic causative verbs (e.g., cause, lead to). Another key insight from literature is that such lexical cues are domain-dependent and thus are required to be specifically tailored towards the target text.

As an alternative to manually constructed rule-based methods, there are also machine-learning-based methods for extraction causality. However, corpora need to be annotated with ground truth for these methods to work, which is typically conducted as manual work by domain experts [11]. An example for such annotation is depicted in Figure 2.1.2, also highlighting the importance of additionally include hints for co-reference resolution (relation from "it" to "something").

Once sufficiently many sentences have been annotated with their contained causal relationships, methods from the field of supervised machine learning can be applied. Traditionally, this has been considered a sequence classification task and approached via Conditional Random Fields [12].

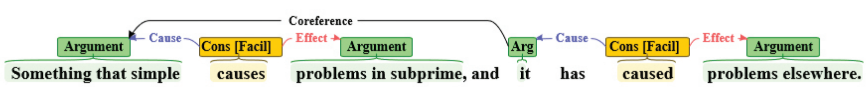

Figure 2.1.2 Manual corpora annotation example.

More recently, deep learning approaches have been adapted for these tasks, including LSTMs [13] and CNNs [14].

We identified two works as promising starting points for our setting, with the first work by Zhao et al. [15] considering extraction of pseudo causal relations from medical text. They make clear that without in-depth domain knowledge, one can only identify candidates that need to be screened by domain experts. Another aspect that makes this work of particular interest for our setting is the inclusion of intre-sentential relationships. Yu et al. [15] provided the second source of methods to extract causal relationships from the scientific literature. Here the language from scientific publications is expected to be closer to the text available as technical reports. They use the contextual word-embedding model BERT and classify sentences into four types of causal relationships as the first stage. The source code of their approach is also available for download[1].

In the semiconductor industry, many unstructured documents containing a significant amount of information are available. Extracting relevant structured information from such documents in a (semi-) automated fashion helps engineers in processing these documents more efficiently. Moreover, allow for more automation. However, the task is extremely challenging given the entirely different style in which each document is produced. Moreover, data annotating is highly effort and resource-intensive. Thus, our research aims to extract causal relations from a different source of unstructured documents in an unsupervised approach. The resulting cause and effect pairs will be used to populate the proposed system concept knowledge representation for further querying, reasoning, or inference. We consider the following sources of domain-specific unstructured texts, production tools manuals, handbooks, and PowerPoint presentations.

We devise a two-step approach, which combines rule-based and machine learning based approaches.

- Connective discovery: in this step, a rule-based approach is leveraged to distinguish sentences that contain causal cues. An initial test on publicly available causal IE dataset [16] shows that the rule-based approach achieves 80% accuracy.
- Entity extraction: in this step, a machine learning based approach is devised to identify the exact phrases for the cause and effect within a sentence that is extracted in the previous step. Specifically, we

[1]Detecting Causal Language Use in Science Findings, https://github.com/junwang4/causal-language-use-in-science

extract candidate phrases per sentence using part-of-speech tagging [17], followed by scoring using a pre-trained BERT model [18, 19]. Experiments on dataset [16] show that this simple approach achieves 40% hits@5.

The method has been tested on the real-life dataset. Initial analysis has revealed that the current assumption that a cause-effect pair will be mentioned within a single sentence is invalid. In addition, given the un-labelled data, a lack of automatic evaluation means is also an issue. Our follow-up work will address these two issues.

2.1.2.3 Failure Analysis Process, Failure Analysis Reports, and Ontologies

The Failure Analysis process aims to trace back detected failures in functional characteristics of a device to its corresponding physical defects. The Failure Analysis process is complex and requires significant knowledge about the device and the different diagnostic tools. Moreover, The Failure Analysis process includes performing experiments and analyzing their outcomes. Finally, the Failure Analysis process outcome is documented in a Failure Analysis report (FA report) report.

The FA report is an unstructured text document that summarizes the entire investigation process of a single device, i.e., the set of hypotheses, experiments, obtained measurements, and their implications. A set of possible hypotheses can be interpreted as the causal model that human domain experts rely on while conducting the Failure Analysis process. Our research aims to extract this causal information from experts and reports, boosting expert systems methods for the Failure Analysis process. The FA report primarily consists of unstructured text. Hence, human domain experts are free to report their work according to their personal preferences. However, the articulation of findings might vary considerably. For example, complete sentences, paragraphs, bullet points, and tables are commonly used in FA reports.

In the semiconductors industry, the acquisition of human domain experts' knowledge and its storage in a machine-readable form paves the way for applying AI methods that consider domain knowledge automatically. Various knowledge representation methods can be used to encode human domain experts' knowledge, e.g. standard definitions of terms used in the reports. In the Failure Analysis domain, we opted to formalize human domain experts' knowledge as an ontology (a well-studied knowledge representation approach designed to store terminological definitions, allowing to structure them in a hierarchical manner) [20]. The Web Ontology Language (OWL), commonly used to define ontologies, allows for the articulation of human domain experts' knowledge. Moreover, given its formal logic-based semantics, OWL ensures that the formulated statements are interpreted unambiguously. Thus, the Failure Analysis ontology formulated using OWL language includes three main definitions:

1. Individuals describe real-world entities, like a job's integrated circuits sample or tools available in a lab.
2. Classes define parts of the Failure Analysis domain by summarizing properties of a collection of individuals, e.g., class OpticalMicroscope comprises all individual microscopes installed in the lab. Also, class IncomingInspection including all individuals describing applications of this method.
3. Properties determine relations between two individuals, e.g., property uses tool links class method, a superclass of class IncomingInspection, with class tools, a superclass of class OpticalMicroscope.

By analyzing a given FA report, human domain experts are able to trace all the laboratory processes, from the first visual inspections of the device to the key method leveraged to identify the fault. Moreover, the ontology provides complementary information that describes the human domain experts' knowledge (not existing in the FA reports). In consequence, the goal of the proposed artificial intelligence tool, which is used in the FA laboratory, is to map the written report to the conceptual knowledge contained in the ontology. We hypothesize that mapping the written report to the conceptual knowledge contained in the ontology allows for further AI-based algorithms. Moreover, as future work algorithms, which capture FA knowledge (reports mapped to the ontology), could be leveraged for diverse tasks such as:

1. Incorporate the language model for FA reports consistency improvements.
2. Offer a centralized search tool for all FA-related knowledge – previous reports, knowledge management database, etc.
3. Assess during the different stages of the FA job, applying the knowledge acquired from previous reports through suggestions and statistical knowledge.

2.1.2.4 Knowledge Representation

No-SQL databases, including graph databases, showed their effectiveness while working with distributed data [21]. Moreover, recent advances in graph embedding algorithms show promising results for downstream tasks such as node classification and link prediction [22]. Therefore, we opted for No-SQL databases in specific graph databases as a framework for the knowledge representation for the proposed system concept.

Moreover, the proposed system concept use case addresses risk Assessment and root cause analysis. In the proposed system concept use case, causal relations are the main relation types of interest. However, humans' ability to interpret causal information makes causal relations deceptively simple. Hence, causal relations are context-dependent [4]. Causal information representation is gaining more interest in many research disciplines to increase the understandability of automated decision-making systems [23].

In the semiconductor industry, the context information is highly variant from one product to another and from one process to another. Therefore, an event that occurs in multiple contexts (i.e., different manufacturing processes or different products) might have a completely different meaning. Consequently, the causal relation between two events might change or even disappear depending on the context. However, human domain experts are able to judge the possibility of the extension of the causal relations between the contexts.

Our research addresses causal knowledge representation in the semiconductor industry manufacturing studying the effectiveness of incorporating context information with regards to the inferencing and reasoning algorithms.

2.1.2.5 Refinement Algorithm

Real-life knowledge graphs, such as those constructed in this work, are often orders of magnitude sparser than benchmark knowledge bases like Freebase, especially if they are built via (automatic) extraction methods from textual corpora. The textual information stored in the text attributes is commonly used to compensate for the lack of structure in a sparse graph. For a recent example, see ref [22]. Language models such as BERT are preferred choices for generating low dimensional representations for the textual information, as they have been shown to consistently improve the performance of a large variety of Natural Language Processing (NLP) tasks [24].

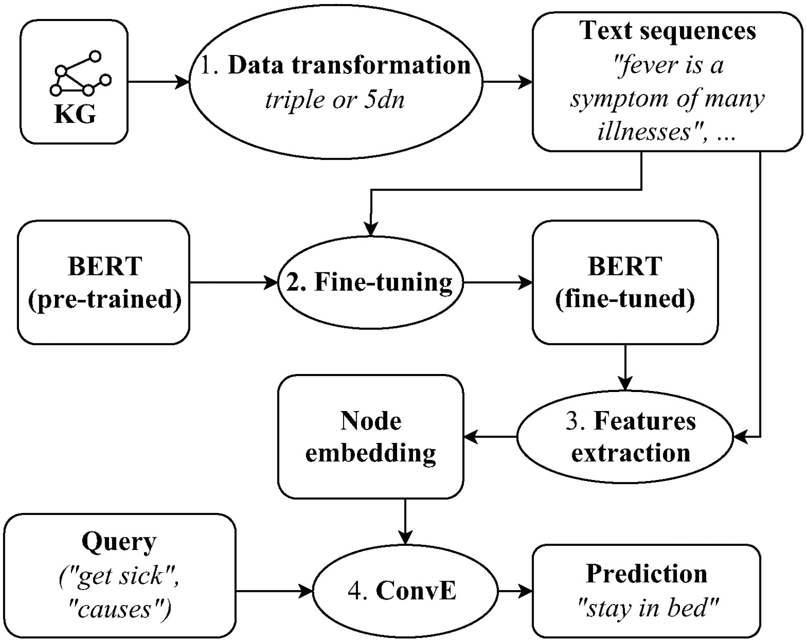

Figure 2.1.3 BERT-ConvE workflow.

Building upon existing work, this work also builds a BERT-based method (BERT-ConvE) to exploit transfer learning of BERT (fine-tuning) in combination with a convolutional network model ConvE [25]. See Figure 2.1.3 for an overview of the main components of the BERT-ConvE method, together with the general workflow. Experiments on ConceptNet [26] show that the proposed method outperforms strong baselines by 50% on knowledge graph completion tasks. The proposed method is suitable for sparse graphs as also demonstrated by empirical studies on ATOMIC [27] and sparsified-FB15k-237 datasets [28]. The next step is to apply the proposed method to the knowledge graph constructed in this work.

However, to our best knowledge, challenges in respect to the use of domain-specific languages, in specific the semiconductor industry, have not been addressed before. Moreover, context information plays a vital role in the correctness of the causal relation. The context information in the proposed use case indicates the settings in which the causal relation is present between two events (the cause and the effect). The prosed concept system addresses the prediction of causal relations, which is context-dependent. Thus, it introduces more sparsity to the resulted knowledge graph.

Our research addresses the causal knowledge graph completion challenge in industrial settings, considering the domain-specific language, structural information, and context information.

2.1.3 Reflections

We presented an AI-based knowledge management system for Risk Assessment and Root Cause Analysis in the Semiconductor Industry. The proposed system process various types of documents where causal relationships are being captured. As these documents were originally intended for interpretation by human experts and created by people of different backgrounds, these documents, in their current form, require additional information extraction algorithms.

Moreover, due to the manual creation of some of these documents and the nature of causal relation, some of these documents i.e., FMEAs, tend to contain a number of inconsistencies calling for consistency checks and automated means of quality improvements. As a response, we systematically defined: (i) a domain-specific model, consisting of a concept schema (i.e., types of cause/effects) and a relationship consistency schema (i.e., valid relations between the concepts), and (ii) a model of inconsistencies, consisting of mixed-up cells (including reverse direction), missing information, and merged cells.

Regarding the FA reports, the unstructured nature of their content (free text) requires a different approach than those utilized in FMEAs. Thus, the discovery of causal relations, among others, is handled with an ontology. The ontology, a formal set of descriptions of terms regarding the Failure Analysis work and the different links between them, is then utilized to map the reports. We hypothesize that this data structure will allow us to develop further AI-based algorithms such as language models.

Also, our research areas address causal knowledge representation in the semiconductor industry manufacturing and aim to study the effectiveness of incorporating context information regarding the hypotheses generation (i.e., predicting possible causal links in causal knowledge graph) methods. Moreover, it addresses the causal knowledge graph completion method in industrial settings. Hence, our research considers sparsity and noise introduced by automated information extraction, sparsity presented by the causal relations context information, and domain-specific language.

Finally, we devised a knowledge graph embedding method BERT-ConvE, that effectively exploits transfer learning and context-dependency of BERT in combination with a convolutional network model, ConvE. Experiments on knowledge graph completion task on publicly available knowledge graphs (ConceptNet, ATOMIC, sparsified FB15k-237) has shown that BERT-ConvE is suitable for sparse knowledge graphs where structural information is limited and textural information is informative for reasoning over the graph.

2.1.4 Conclusion

In conclusion, while human domain experts remain the key source of knowledge, the proposed system aims to mimic their ability to extract information from different data sources and extend knowledge between different scenarios to support the experts in their repetitive tasks. Moreover, although the proposed system concept may appear simple in design, the proposed use case (Risk Assessment and Root Cause Analysis in Semiconductor Industry) pushes the boundaries of many states of the art methods of artificial intelligence and natural language processing. Furthermore, our research areas highlight novel approaches to address causal domain knowledge information extraction, representation, and completion, leveraging a combination of advances in computer science, artificial intelligence, and natural language processing.

Acknowledgments

This work is conducted under the framework of the ECSEL AI4DI "Artificial Intelligence for Digitising Industry" project. The project has received funding from the ECSEL Joint Undertaking (JU) under grant agreement No 826060. The JU receives support from the European Union's Horizon 2020 research and innovation programme and Germany, Austria, Czech Republic, Italy, Latvia, Belgium, Lithuania, France, Greece, Finland, Norway.

References

[1] L. Ashley and G. Armitage, "Failure mode and effects analysis: an empirical comparison of failure mode scoring procedures," *Journal of patient safety*, vol. 6, no. 4, pp. 210–215, 2010, doi: 10.1097/pts.0b013e3181fc98d7.

[2] M. Scorsetti *et al.,* "Applying failure mode effects and criticality analysis in radiotherapy: lessons learned and perspectives of

enhancement," *Radiotherapy and oncology: journal of the European Society for Therapeutic Radiology and Oncology*, vol. 94, no. 3, pp. 367–374, 2010, doi: 10.1016/j.radonc.2009.12.040.

[3] N. Viscariello *et al.,* "A multi-institutional assessment of COVID-19-related risk in radiation oncology," *Radiotherapy and oncology : journal of the European Society for Therapeutic Radiology and Oncology*, vol. 153, pp. 296–302, 2020, doi: 10.1016/j.radonc.2020.10.013.

[4] N. McDonnell, "Transitivity and proportionality in causation," *Synthese*, vol. 195, no. 3, pp. 1211–1229, 2018, doi: 10.1007/s11229-016-1263-1.

[5] Erik HOLLNAGEL, "Evaluation of Expert Systems," in *Studies in Computer Science and Artificial Intelligence, Topics in Expert System Design*, Giovanni GUIDA and Carlo TASSO, Eds.: North-Holland, 1989, pp. 377–416. [Online]. Available: https://www.sciencedirect.com/science/article/pii/B9780444873217500193

[6] J.-L. Wu, L.-C. Yu, and P.-C. Chang, "Detecting causality from online psychiatric texts using inter-sentential language patterns," *BMC Medical Informatics and Decision Making*, vol. 12, no. 1, p. 72, 2012, doi: 10.1186/1472-6947-12-72.

[7] K. A. Rogers, "Hume on Necessary Causal Connections," *Philosophy*, vol. 66, no. 258, pp. 517–521, 1991, doi: 10.1017/S0031819100065165.

[8] L. A. Paul, N. Hall, and E. J. Hall, *Causation: A user's guide*: Oxford University Press, 2013.

[9] Julian Reiss, "Suppes' probabilistic theory of causality and causal inference in economics," *Journal of Economic Methodology*, vol. 23, no. 3, pp. 289–304, 2016, doi: 10.1080/1350178X.2016.1189127.

[10] K. Radinsky, S. Davidovich, and S. Markovitch, "Learning causality for news events prediction," in *Proceedings of the 21st international conference on World Wide Web*, Lyon, France: Association for Computing Machinery, 2012, pp. 909–918.

[11] J. Dunietz, L. Levin, and J. Carbonell, "Annotating Causal Language Using Corpus Lexicography of Constructions," in *Proceedings of The 9th Linguistic Annotation Workshop*, 2015, pp. 188–196. [Online]. Available: https://www.aclweb.org/anthology/W15-1622

[12] C. Mihăilă and S. Ananiadou, "Recognising discourse causality triggers in the biomedical domain," *Journal of bioinformatics and computational biology*, vol. 11, no. 06, p. 1343008, 2013.

[13] *Automatic extraction of causal relations from text using linguistically informed deep neural networks*, 2018.

[14] *Relation extraction: Perspective from convolutional neural networks*, 2015.

[15] *Causaltriad: toward pseudo causal relation discovery and hypotheses generation from medical text data*, 2018.

[16] I. Hendrickx *et al.,* "SemEval-2010 Task 8: Multi-Way Classification of Semantic Relations between Pairs of Nominals," in *Proceedings of the 5th International Workshop on Semantic Evaluation*, 2010, pp. 33–38. [Online]. Available: https://www.aclweb.org/anthology/S10-1006

[17] M. Honnibal, I. Montani, S. van Landeghem, and A. Boyd, *spaCy: Industrial-strength Natural Language Processing in Python*: Zenodo.

[18] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, "BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding," in *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)*, 2019, pp. 4171–4186. [Online]. Available: https://www.aclweb.org/anthology/N19-1423

[19] Thomas Wolf *et al., HuggingFace's Transformers: State-of-the-art Natural Language Processing*.

[20] D. Allemang and J. Hendler, *Semantic Web for the Working Ontologist: Effective Modeling in RDFS and OWL. Second Edition*, 2011.

[21] Z. Wei-ping, L. Ming-xin, and C. Huan, "Using MongoDB to implement textbook management system instead of MySQL," in *2011 IEEE 3rd International Conference on Communication Software and Networks*, 2011, pp. 303–305.

[22] D. Q. Nguyen, "A survey of embedding models of entities and relationships for knowledge graph completion," *arXiv: 1703.08098*, 2017.

[23] B. Schölkopf *et al.,* "Toward Causal Representation Learning," *Proceedings of the IEEE*, vol. 109, no. 5, pp. 612–634, 2021, doi: 10.1109/JPROC.2021.3058954.

[24] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, "BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding," in *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)*, 2019, pp. 4171–4186. [Online]. Available: https://www.aclweb.org/anthology/N19-1423

[25] T. Dettmers, P. Minervini, P. Stenetorp, and S. Riedel, "Convolutional 2d knowledge graph embeddings," in *Proceedings of the AAAI Conference on Artificial Intelligence*, 2018.

[26] R. Speer and C. Havasi, "ConceptNet 5: A large semantic network for relational knowledge," in *The People's Web Meets NLP*: Springer, 2013, pp. 161–176.
[27] *Atomic: An atlas of machine commonsense for if-then reasoning*, 2019.
[28] J. Pujara, E. Augustine, and L. Getoor, "Sparsity and Noise: Where Knowledge Graph Embeddings Fall Short," in *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing*, 2017, pp. 1751–1756. [Online]. Available: https://www.aclweb.org/anthology/D17-1184