Cognitive Computing and Internet of Things, Vol. 124, 2024, 123–132
https://doi.org/10.54941/ahfe1004710

Explainability of Industrial Decision Support System Using Digital Design Thinking With Scene2Model

Christian Muck[1], Julia Tschuden[2], Herwig Zeiner[2], and Wilfrid Utz[1]

[1]OMiLAB NPO, 10785 Berlin, Germany
[2]Joanneum Research Forschungsgesellschaft mbH, 8010 Graz, Austria

ABSTRACT

To ensure the acceptance of decisions made in complex cyber-physical environments, orchestrated between human and machine actors, not only the developers need to understand how a decision is reached, but also the decision-makers and stakeholders affected by the decisions. To this end this contribution discusses how high-level visualisations can be derived to support the explanation of decisions using OMiLAB's digital design thinking approach in an inverse manner. These visualisations will not be mere pictures, but diagrammatic models, containing additional information, which is understandable to machines, allowing to process them during an enrichment phase and interactively explain their involvement and impact to the users. The representation as conceptual models enables a) the cognitive perception by human actors, b) the machine interpretation for semantic lifting (focusing on elevating understandability) and c) further design iterations to adapt the system to become adequate and effective from a design but also operational perspective.

Keywords: Conceptual modelling, Decision support, Explainability, User-centered

INTRODUCTION

Today's technologies open up countless new improvements in information systems, which not only enable to optimize the system but also enable a better understanding of how the system works by its users. The problem of understanding how a system works is gaining increasing interest in the AI community, especially in the context of AI supported *Decision Support Systems (DSS)* (Bayer, Gimpel, & Markgraf, 2022). However, the right way to explain the decision of a DSS to a user depends on multiple factors, such as the algorithms used to make the decision, the way of representing the explanation and the users who receive the explanation. Explanations themselves should not only focus on explaining how the AI approach works, but also justify the suggestions made by the system. Furthermore, experts in the domain may find explanations insufficient, which lowers the chance of acceptance.

The method by which the explanations are presented to the users should not only be based on the intuition of the developer of the DSS, but should be based on established approaches (Miller, 2019). As such, explanations should be adapted to the target user. Such adaptation requires recognising their expertise level and their social context, such as beliefs and characteristics. This means that the explanations should be specifically tailored to the users and not to AI experts, e.g., to prefer the usage of causal links within decisions rather than to state probabilities (Miller, 2019). As users can come from various backgrounds, and be involved in both the design and explanation process, see Figure 1, their existing knowledge and explanation methods should be drawn upon when providing explanations to them.

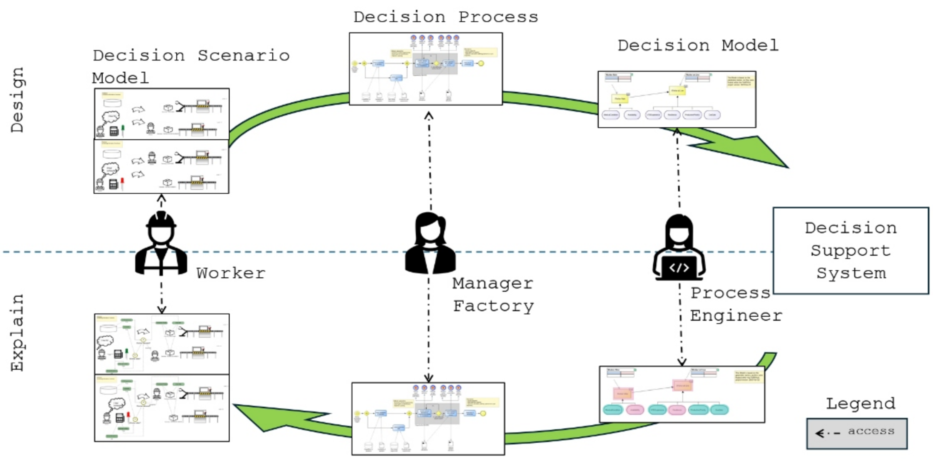

Figure 1: Target user definition and their involvement in the design and explanation process.

The research introduced in this paper is being carried out as part of the EU-funded FAIRWork project (cf. (Woitsch, Muck, Utz, & Zeiner, 2023)). In it we aim to develop the *Democratic AI-based Decision Support System (DAI-DSS)*, to improve decision-making in production environments. This system will not only optimise the production process, but will also propose decisions that are “fair” to users, considering their preferences and individual situations. This includes representing the results in a way that ensures that workers can understand and trust the decisions proposed by the DAI-DSS. This, by extension, promotes acceptance.

As the authors are partners in the FAIRWork project, this position paper aims to introduce how conceptual modelling can be used to visualize the information taken from the DAI-DSS to present it to the users. The information, when transformed into a conceptual model, can then be utilized by algorithms to further explain the made decisions to the users.

To achieve this, the models prepared during the design process, together with the decision output from the DSS itself, should be used to support the explanation of the decisions made to users, be they decision makers or workers.

The paper first presents related work that serves as a foundation for the presented idea. We then introduce the proposed approach to using conceptual models to capture and utilise decision knowledge. We start by looking at how we can use it in the decision support's design process and then how it can be used to explain concrete decision. We conclude this paper with our conclusion and an outlook for future work.

RELATED WORK

Decision-making in industry (Wendt & Manhart, 2020) is a complex process that is influenced by a variety of factors. Companies that have good data can make more informed decisions than others, and processes and technologies improve industrial decision-making in a number of ways. Software solutions help to collect and analyse data. Finally, people (as discussed in Nelles, Kuz, Mertens, & Schlick, 2016), play an important role in industrial decision-making as they interpret data, apply technology and make final decisions.

AI Supported Decision Making and Explainability

Decision-making is an important part of the daily work in industry. One of the first decision-making models with three steps, namely investigation, design and selection, was proposed by Simon (Simon, 1960). Mintzberg (Mintzberg, Raisinghani and Theoret, 1976) proposed a different model for decision making in companies. The overall decision making process is defined as a set of actions and dynamic factors. These factors identify actions and require a specific commitment to these actions. The overall process has seven central steps, which are non-sequential. Certain steps can be bypassed or interrupted, or include the ability to provide feedback.

Recently, a systematic review on artificial intelligence for decision support systems was given by Gupta (Gupta, Modgil, Bhattacharyya, & Bose, 2022). For decision making, many different AI based approaches are used. Another possibility would be time-aware knowledge graph-based approaches (e.g., Zeiner, Weiss, Unterberger, Maurer, & Jöbstl, 2019) or mathematical models for optimisation topics in a production line described in recent work (Bogner, Pferschy, Unterberger, & Zeiner, 2018).

As AI based approaches and mathematical models become an integral part of many application areas, such as decision process optimization, explainability and transparency becomes increasingly important for human decision makers and those who are affected by these decisions (Xu, Uszkorei, Du, Fan, Zhao, & Zhu, 2019). When users do not comprehend the decisions proposed by the AI based approaches or models, acceptance and trust in the system diminishes. Both are particiularly at risk of not being understood.

Explainable AI refers to the concept of creating AI based models whose decisions, underlying reasons and predictions can be interpreted by humans (Saarela & Jauhiainen, 2021). Therefore, using interpretable algorithms helps to build trust and accountability (Ashoori & Weisz, 2019), which leads to a more informed decision making. Deep learning algorithms operate as blackboxes and usually offer less explainability than their more traditional machine learning counterparts (Ashoori & Weisz, 2019). While they lack in explanatory simplicity, their complex and non-linear architecture allows for higher prediction accuracy. Current research efforts focus on making deep learning algorithms more transparent. Xu, Uszkorei, Du, Fan, Zhao,

& Zhu (2019) point out that there are two approaches to explainable AI: transparency design and post-hoc explanation. Transparency design aims to understand the model's structure, the inner workings of algorithms and the meaning of the components. Post-hoc explanation aims to provide analytical statements, visualisations or explanations.

Modelling

We propose to use conceptual models to support the explanation of the made decisions, as conceptual modelling, especially using domain concepts is already widespread used to describe system under study in a comprehendible way (Karagiannis, Buchmann, Burzynski, Reimer, & Walch, 2016; Frank, 2013). Using domain concepts in models improves the understanding by stakeholders familiar with the domain. Further, domain specific modelling methods often use visual or diagrammatic presentations, improving the understandability of the representation (Larkin & Simon, 1987; Harel & Rumpe, 2000) compared to pure textual descriptions by utilising spatial information to represent knowledge. Examples for diagrammatic conceptual models can be found in Figure 2. In such diagrams, concepts of a domain are represented by individual objects and their position on the model influences its interpretation (Mayr & Thalheim, 2020), hence the name conceptual modelling. Using spatial information alone is not enough in this context, as each object must have its own visualisation (Karagiannis & Kühn, 2002), whereby this visualisation must fit to the semantic of the object and must be interpretable by the users of the models (Moody, 2009).

As we already utilise conceptual modelling within FAIRWork to define decision knowledge and as input for the configuration of the DAI-DSS, it is sensible to further use this approach in the explanation of made decisions. We use conceptual modelling on different abstraction levels (Woitsch, Muck, Utz, & Zeiner, 2023), to support the configuration of the DAI-DSS to meet the requirements of specific decision scenarios. More information on how we do this can be found in section *Methodology*.

We use models on different abstraction levels, to tailor them to the need of the involved stakeholders, like experts or users (Karagiannis, Buchmann, & Utz, 2022). In the beginning we use conceptual models together with design thinking workshops, supported by the Scene2Model tool (Muck & Palkovits-Rauter, 2021). These are models on a high-abstraction level to support an easy communication of stakeholders of different backgrounds, whereby the diagrammatic models help to communicate and understand the situation. Such semantic rich and high-level representations are based on human interpretable pictures representing important concepts form the domain.

Afterwards we use more formal modelling methods, e.g., the *Decision Model and Notation (DMN)*, which allows to represent the decision knowledge in a computer interpretable way. Depending on the users such models on a higher-level or more formal can support the understanding of the decisions if the used concepts fit to the domain and the knowledge of the users (Muck & Utz, 2023).

As the understandability of an explanation depends on tailoring it to the consumers of the explanation (Bayer, Gimpel, & Markgraf, 2022), the available concepts in the modelling method must fit to the users to support an easy interpretation. To achieve this in the context of this paper, the modelling method itself must be adaptable so that the concepts are easy comprehendible by users (Karagiannis, 2015).

High-level representations, which are dynamically adapted, are already used in design thinking (cf. (Brown, 2008)) to communicate and explain complex ideas between people with various backgrounds. Having something visual on a high-abstraction level supports the understanding of the big picture and not everyone needs a complete understanding of all the details.

Design thinking itself is often used in physical workshops to foster co-creation between the participants. In these high-level representations are used to foster the communication and enable an exploration of possible solutions. But the representations, once created cannot only be used within the workshop, but also used to explain the idea to stakeholders, who did not participate in the workshop.

These workshops can be further supported by conceptual modelling, to create digital representation of the physical crated high-level representation, allowing for easier sharing and processing of the created idea (Muck & Palkovits-Rauter, 2021). As the digital representation is adaptable, they can be reused and enhanced to explain certain circumstances.

METHODOLOGY

As a general feature, the design process in our work follows a model-based approach. This involves externalising knowledge about a use case in the form of conceptual representations using domain-specific modelling languages that are appropriate and provide the necessary constructs for representation and processing. The decomposition is performed as a formalisation process - high-level scenarios are designed in a collaborative, interactive and agile setting involving expert stakeholders from different backgrounds.

The joint grouping, evaluation and assessment of these scenarios triggers either a scenario level interaction or a decomposition into process representations and technical architectures. In general, we use models with high abstraction levels within workshops, to explore and define the decision problem and its context. To achieve this, we use the Scene2Model tool (Muck & Palkovits-Rauter, 2021), which offers a semantically rich representation of scenarios for physical workshop environments as well as their digital models. This semantically rich representation are domain specific concept figures, which can easily be interpreted by humans. In order to make the physical components understandable to the machine, we established a metamodel of the figures that contains their semantic representation (Karagiannis & Kühn, 2002; Karagiannis, Buchmann, Burzynski, Reimer, & Walch, 2016). This alone, however, is not enough, as the representation must fit to the user's domain and context, to increase comprehensibility (Muck & Utz, 2023; Karagiannis, 2015).

The **methodology** is build on a loose coupling approach. This means that according to the requirements of the use case, the selected method or technique can be changed for each phase. The following steps were selected for our scenario.

Firstly, the preparation phase is concerned with identifying the concepts needed to represent the scenarios. This is driven by the domain-specific and project-specific requirements. In this context, domain-specific means the industrial sector, whereas project-based requirements are derived from the objectives of the user scenarios. The result of this phase is a visual design library that is appropriate and relevant for the on-site stakeholder workshops. In these on-site workshops, the stakeholders involved will use and understand the prepared visual design vocabulary.

Secondly, the scenario development is realized in a collaborative effort using the Scene2Model[1] approach. During Design Thinking workshops, conducted on-site or remotely, storyboards represent the abstract use case descriptions. The aim during this phase is to understand the case and analyse it from different perspectives (relevance, applicability) without limiting creativity through formalisation.

Lastly, based on the process design in step 2, the technical alignment of the technical services, in particular the decision services, is started. We consider this phase to be continuously evolving, depending on the dynamics of the available decision services, the dynamics of the application scenario and the lessons learned.

EXPLAINING CONCRETE DECISIONS USING CONCEPTUAL MODELLING

In the previous section we introduced how conceptual models can be used to encode knowledge, making it sharable with other stakeholders as well as understandable by machines. Such models can be used to explain the decision on a generic level, since they focus on how a decision is made, but they do not contain decision specific information. This is necessary so that the models can be used for designing the decision-making but limits their use in explaining actual decisions.

A domain specific example would be the definition of rules to determine which workers are allowed to work on which production lines. To let the system decide whether a specific worker is allowed to operate a specific line, the generic rules are not enough. Concrete data must be provided as input for the rules, and an overarching system is needed to execute them. This concrete knowledge, however, is not available in the design models, and only available to the DSS at run time. Therefore, the design models can be used to explain the generic decision, but they lack information to explain concrete situations to the affected workers, which is available in the decision support system.

Conceptual models, if fed with data from the real environment, can facilitate the monitoring of systems by lifting the gathered data towards the problem space and represent them on a higher abstraction level (Szvetits & Zdun, 2016). Through such an approach the state of information systems can be explained on a higher abstraction level, easing the comprehension by

[1] https://www.omilab.org/design-thinking

users. Additionally, such models can be used as input for further processing and for offering additional support in understanding and handling the system.

But not only existing models can be enhanced with data from a system to represent and analyse the current state, but also new models can be established based on data available in a system (van der Aalst et al., 2012). Such models can also be used to analyse the current situation and support the decision making in an organization.

For the approach introduced in this paper, we build upon these concepts and use conceptual modelling and their diagrammatic notation to support the explanation of made decisions to involved stakeholders, such as workers, managers and engineers. The diagrammatic representation of the models can only explain part of the decision, yet can be used as input for further processing, as the structure is processable by machines. This way, for example, a textual explanation could be created out of them, using generative AI.

To create such models, the concrete decision data from the decision support system, must be linked to meta information about the modelling methods, which should be used to create the models. In Figure 2 we visualize the conceptual structure of our idea. The DSS provides the information of concrete made decisions and real-time data and the models created in the design, provide a basic structure and the meta information.

The meta information from the available models and their meta model (containing meta information) is then used as a foundation to instantiate the explanation models. Meaning that not only the meta data from the design models can be used, but also the generic structure provided by the models.

Therefore, the data from the DSS must be semantically lifted, to fit to a modelling method. The concrete meta information that is added, depends on the user and which modelling method should be used to explain the decision. For example, for workers, who may not know formal modelling method, a high-level design thinking-based method is feasible. For decision makers, the DMN modelling method may be more appropriate.

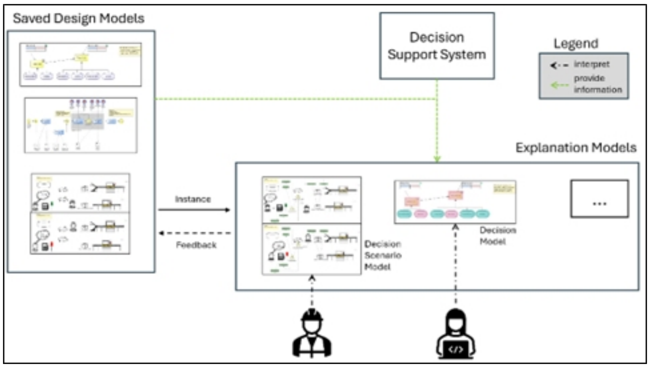

Figure 2: Conceptual overview of using conceptual models to explain decisions.

However, since the modelling method may not encompass everything needed for the explanation of the decision, modelling methods should be able to adapt to the explanation. High-level models, for example, need a way to represent the system's different decision parameters and may highlight the important ones for the concrete decision. A new concept could also be needed, which is not considered yet.

These adaptations then provide feedback to the modelling methods, used in the design step, to improve their expressability for future uses. Then all the needed concepts for the high-level design thinking models are available in the workshop and can be used from the start.

In the FAIRWork project, for example, we have used conceptual models to design and the define the decision making in the Democratic AI-based Decision Support System (DAI-DSS), which will be implemented within the project. Therefore, we have a **repository of design models** available, which can be used as input for the explanations. Additionally, we have **information about the concrete decision** available in the decision support system. These decision parameters include information such as the number of workers, the product to be created, available productions lines and many more.

If we want to explain the decision then with a model, we must **choose the used modelling method** and link the concrete information to the available meta information. For example, if we choose the high-level models, we can **semantically lift** each *WorkerID* available within the decision, to the concept of a worker used in the model. Each *ProductionLineID* can be linked to the production line concept. The attributes used, e.g., availability to the worker, can be mapped to an attribute concept and visually connected to the worker within the model or the attribute can be saved internally to the object of the worker.

The layout of the model can be defined through the design model, or through a layouting algorithm. As models are processable, they can also load information about the decision outcome and possible sub decisions and show it in the model.

The user can then explore the information within the models directly. Alternatively, the structured information within the models can be used to create a prompt for generative AI to generate an additional textual explanation. Another option is using other services to provide more information about the decisions made, e.g., showing which rules influenced the decision the most. All of those options should increase trust and comprehension and therefore the acceptance of the decisions.

CONCLUSION

The approach to identifying decision problems and their implementation using this overall approach is presented. The use of conceptual models to understand decision problems and specify decision processes helped to share the necessary knowledge and supported the design and prototyping of decision services. An important aspect of the approach is flexibility through model-based adaptation, which allows the decision support system to be customised so that different decision approaches can be used to address different decision problems. Depending on the application, these simpler models can be used to explain more complex AI-based models. Our approach therefore increases the user's trust and comprehension and, in turn, the acceptance of decisions made.

As the project is currently ongoing, we plan to test our approach on different complex models that provide decision support for the specified use cases, such as automated test building, machine maintenance, worker (re)allocation or workload balance.

ACKNOWLEDGMENT

This work has been supported by the *FAIRWork* project (www.fairwork-project.eu) and has been funded within the European Commission's Horizon Europe Programme under contract number 101069499. This paper expresses the opinions of the authors and not necessarily those of the European Commission. The European Commission is not liable for any use that may be made of the information contained in this paper.

REFERENCES

Ashoori, M., & Weisz, J. D. (2019). In AI We Trust? Factors That Influence Trustworthiness of AI-infused Decision-Making Processes. ArXiv, abs/1912.02675. Retrieved from https://api.semanticscholar.org/CorpusID:208637106.

Bayer, S., Gimpel, H., & Markgraf, M. (2022). The role of domain expertise in trusting and following explainable AI decision support systems. Journal of Decision Systems, 32, 110–138. doi: 10.1080/12460125.2021.1958505.

Bogner, K., Pferschy, U., Unterberger, R., & Zeiner, H. (2018). Optimised scheduling in human–robot collaboration–a use case in the assembly of printed circuit boards. International Journal of Production Research, 56(16), 5522–5540.

Brown, T. (2008, June). Design Thinking. Harvard Business Review, 84–92. Retrieved from http://www.ideo.com/images/uploads/thoughts/IDEO_HBR_Design_Thinking.pdf.

Gupta, S., Modgil, S., Bhattacharyya, S., & Bose, I. (2022). Artificial intelligence for decision support systems in the field of operations research: Review and future scope of research. Annals of Operations Research, 1–60.

Frank, U. (2013). Domain-Specific Modeling Languages: Requirements Analysis and Design Guidelines. In I. Reinhartz-Berger, A. Sturm, T. Clark, S. Cohen, & J. Bettin (Eds.), Domain Engineering: Product Lines, Languages, and Conceptual Models (pp. 133–157). Berlin, Heidelberg: Springer Berlin Heidelberg. doi: 10.1007/978-3-642-36654-3_6

Harel, D., & Rumpe, B. (2000). Modeling languages: Syntax, semantics and all that stuff, part i: The basic stuff.

Karagiannis, D. (2015). Agile modeling method engineering. Proceedings of the 19th Panhellenic Conference on Informatics, (pp. 5–10).

Karagiannis, D., & Kühn, H. (2002). Metamodelling platforms. EC-Web, 2455, p. 182.

Karagiannis, D., Buchmann, R. A., Burzynski, P., Reimer, U., & Walch, M. (2016). Fundamental Conceptual Modeling Languages in OMiLAB. In D. Karagiannis, H. C. Mayr, & J. Mylopoulos (Eds.), Domain-Specific Conceptual Modeling: Concepts, Methods and Tools (pp. 3–30). Cham: Springer International Publishing. doi: 10.1007/978-3-319-39417-6_1.

Larkin, J. H., & Simon, H. A. (1987). Why a diagram is (sometimes) worth ten thousand words. Cognitive science, 11, 65–100.

Mayr, H. C., & Thalheim, B. (2020). The triptych of conceptual modeling. Software and Systems Modeling, 1–18.

Miller, T. (2019). Explanation in artificial intelligence: Insights from the social sciences. Artificial Intelligence, 267, 1–38. doi: https://doi.org/10.1016/j.artint.2018.07.007.

Mintzberg, H., Raisinghani, D., & Theoret, A. (1976). The structure of "unstructured" decision processes. Administrative science quarterly, 246–275.

Miron, E.-T., Muck, C., & Karagiannis, D. (2019). Transforming Haptic Storyboards into Diagrammatic Models: The Scene2Model Tool. Proceedings of the 52nd Hawaii International Conference on System Sciences.

Moody, D. (2009). The "physics" of notations: toward a scientific basis for constructing visual notations in software engineering. IEEE Transactions on software engineering, 35, 756–779.

Muck, C., & Palkovits-Rauter, S. (2021). Digitising Tangible Design Thinking Artefacts to Support Information Exchange: The Storyboard-based Scene2Model Approach. In D. Karagiannis, M. Lee, & K. Hinkelmann (Eds.). Springer.

Muck, C., & Utz, W. (2023, March). A Recognition Service for Haptic Modelling in Scene2Model. In A. Martin, H.-G. Fill, A. Gerber, K. Hinkelmann, D. Lenat, R. Stolle, & F. van Harmelen (Ed.), Proceedings of the AAAI2023 Spring Symposium on Challenges Requiring the Combination of Machine Learning and Knowledge Engineering (AAAI-MAKE 2023), CEUR Workshop Proceedings. Retrieved from https://ceur-ws.org/Vol-3433/short4.pdf.

Nelles, J., Kuz, S., Mertens, A., & Schlick, C. M. (2016, March). Human-centered design of assistance systems for production planning and control: The role of the human in Industry 4.0. In 2016 IEEE International Conference on Industrial Technology (ICIT) (pp. 2099–2104). IEEE.

Saarela, M., & Jauhiainen, S. (2021). Comparison of feature importance measures as explanations for classification models. SN Applied Sciences, 3, 1–12.

Simon, H. A. (1960). The new science of management decision.

Szvetits, M., & Zdun, U. (2016). Systematic literature review of the objectives, techniques, kinds, and architectures of models at runtime. Software & Systems Modeling, 15, 31–69.

Wendt, T., & Manhart, S. (2020). Digital Decision Making als Entscheidung, nicht zu entscheiden: Zur Zukunft des Entscheidens in der Digitalisierung. Arbeit, 29(2), 143–160.

van der Aalst, W., Adriansyah, A., de Medeiros, A. K., Arcieri, F., Baier, T., Blickle, T.,... Wynn, M. (2012). Process Mining Manifesto. In F. Daniel, K. Barkaoui, & S. Dustdar (Ed.), Business Process Management Workshops (pp. 169–194). Berlin: Springer Berlin Heidelberg.

Woitsch, R., Muck, C., Utz, W., & Zeiner, H. (2023, September). Towards a democratic AI-based decision support system to improve decision making in complex ecosystems. Joint Proceedings of the BIR 2023 Workshops and Doctoral Consortium co-located with 22nd International Conference on Perspectives in Business Informatics Research (BIR 2023). CEUR Workshop Proceedings. Retrieved from https://ceur-ws.org/Vol-3514/paper94.pdf.

Xu, F., Uszkoreit, H., Du, Y., Fan, W., Zhao, D., & Zhu, J. (2019). Explainable AI: A Brief Survey on History, Research Areas, Approaches and Challenges. In J. Tang, M.-Y. Kan, D. Zhao, S. Li, & H. Zan (Ed.), Natural Language Processing and Chinese Computing (pp. 563–574). Cham: Springer International Publishing.

Zeiner, H., Weiss, W., Unterberger, R., Maurer, D., Jöbstl, R. (2019). Time-Aware Knowledge Graphs for Decision Making in the Building Industry. In: Freitas, P., Dargam, F., Moreno, J. (eds) Decision Support Systems IX: Main Developments and Future Trends. EmC-ICDSST 2019. Lecture Notes in Business Information Processing, vol. 348. Springer, Cham. https://doi.org/10.1007/978-3-030-18819-1_5