A TAXONOMY FOR UNCERTAINTY-AWARE EXPLAINABLE AI

Short Paper

Maximilian Förster, University of Ulm, Germany, maximilian.foerster@uni-ulm.de
Michael Hagn, University of Regensburg, Germany, michael.hagn@ur.de
Nico Hambauer, University of Regensburg, Germany, nico.hambauer@ur.de
Paula Jaki, TU Darmstadt, Germany, jaki@ise.tu-darmstadt.de
Andreas Obermeier, University of Ulm, Germany, andreas.obermeier@uni-ulm.de
Marc Pinski, TU Darmstadt, Germany, pinski@ise.tu-darmstadt.de
Andreas Schauer, University of Regensburg, Germany, andreas.schauer@ur.de
Alexander Schiller, University of Regensburg, Germany, alexander.schiller@ur.de
Alexander Benlian, TU Darmstadt, Germany, benlian@ise.tu-darmstadt.de
Bernd Heinrich, University of Regensburg, Germany, bernd.heinrich@ur.de
Ekaterina Jussupow, TU Darmstadt, Germany, ekaterina.jussupow@tu-darmstadt.de
Mathias Klier, University of Ulm, Germany, mathias.klier@uni-ulm.de
Mathias Kraus, University of Regensburg, Germany, mathias.kraus@ur.de
Daniel Schnurr, University of Regensburg, Germany, daniel.schnurr@ur.de

Abstract

Artificial Intelligence (AI) is increasingly used to augment human decision-making. However, especially in high-stakes domains, the integration of AI requires human oversight to ensure trustworthy use. To address this challenge, emerging research on Explainable AI (XAI) focuses on developing and investigating methods to generate explanations for AI outcomes. Yet, current approaches often yield limited explanations, neglecting the various sources of uncertainty that strongly influence AI-augmented decision-making. This paper presents a first step to establishing a foundation for future research in uncertainty-aware XAI. By applying the Extended Taxonomy Design Process, we aim to develop an integrated, hierarchical taxonomy to structure the key characteristics of uncertainty-aware XAI. Through this approach, we identify four primary sources of uncertainty: data uncertainty, AI model uncertainty, XAI method uncertainty, and human uncertainty. Furthermore, we propose a preliminary taxonomy as an initial foundational framework for the future design and evaluation of uncertainty-aware XAI.

Keywords: Explainable AI, Uncertainty-Aware Explanation, Taxonomy, AI-Augmented Decisions

1 Introduction

Artificial Intelligence (AI) is increasingly employed to augment human decision-making. AI-augmented decision-making differs from traditional decision support and fully automated AI systems by using AI as a sophisticated assistant with greater autonomy, while still relying on humans as the final decision-makers (Pathirannehelage et al., 2024). Enabling human oversight is of particular interest in high-stakes domains like healthcare, where the trustworthy and responsible application of AI models is required. This remains challenging due to the black-box nature of AI (Meske et al., 2022). The emergence of the field of Explainable AI (XAI) represents a promising avenue for addressing these challenges. XAI aims to provide an explanation of an AI model's outcome, enabling humans to better understand, trust, and effectively interact with AI models (Bauer et al., 2023a). Another crucial factor to consider when dealing with high-stakes domains is uncertainty. In line with Holton (2004), we define uncertainty as "a state of not knowing whether a proposition is true or false". However, in AI-augmented decision-making, AI predictions are never provably correct but only hypothetical and potentially suffering from including incorrect model assumptions and noisy or imprecise data (Hüllermeier & Waegeman, 2021). Due to such factors, AI predictions may be true or false, a state which is referred to as "uncertain" (Holton, 2004).

Current XAI approaches often neglect the various sources of uncertainty (e.g., resulting from incomplete data) inherent in AI-augmented decision-making. These uncertainties negatively affect the performance and trustworthiness of AI (Thompson et al., 2022). Ignoring or underestimating these uncertainties can create a false sense of trust and confidence in the AI model predictions, undermining the intended benefits of AI (Eiband et al., 2019). Thus, to ensure the trustworthy and responsible use of AI, it is essential to address not only the explainability of AI models' outcomes but also the various uncertainties that arise in AI-augmented decision-making and to communicate them effectively to users. Moreover, incorporating and conveying uncertainties within explanations can enhance AI-augmented decision-making by empowering users to make more informed decisions (Bauer et al., 2023a). Motivated by the need for research that effectively combines explainability and uncertainty awareness for human stakeholders (Gawlikowski et al., 2023), some early studies have started exploring how to convey certain aspects of uncertainty to users (cf. Section 3.3). However, to achieve meaningful uncertainty-aware XAI – i.e., explanations that integrate and communicate uncertainties in AI-augmented decision-making to users – it is necessary to understand the sources of uncertainty and their inherent characteristics to then develop appropriate methods that help humans handle uncertainty. Current research, however, lacks a clear characterization and integration of those sources in AI-augmented decision-making and their systematized foundation in a taxonomy. To address this gap, this research-in-progress paper presents preliminary results from our taxonomy design. Our work addresses the following research questions:

RQ1: *What are important characteristics of uncertainty-aware XAI for AI-augmented decision-making?*

RQ2: *How can these characteristics be systematized based on a comprehensive taxonomy?*

This work contributes to the existing body of literature in the following ways: First, our research presents a first step to establishing a foundation for uncertainty-aware XAI by elucidating and combining the multifaceted sources of uncertainty in AI-augmented decision-making. Second, by linking research on XAI and uncertainty, we enhance the understanding of the interplay among sources of uncertainty that are currently considered in isolation or in a fragmented manner. Finally, our preliminary taxonomy provides a foundation for future development of XAI methods that account for different sources of uncertainty, ultimately supporting high-stakes decision-making.

2 Methodology

We develop the taxonomy according to the Extended Taxonomy Design Process (ETDP) by Kundisch et al. (2022), which builds upon Nickerson et al. (2013). In Figure 1 we visualize the 18 Steps of the ETDP. With respect to problem identification, motivation, and the definition of objectives, our rationale for examining uncertainty in the context of AI-augmented decision-making and designing a taxonomy for researchers and practitioners has been outlined above (Steps 1-3 of the ETDP). The meta-characteristic for the taxonomy (Step 4) is uncertainty-aware XAI (cf. Section 1). Regarding ending conditions, we intend to follow the suggestions by Nickerson et al. (2013), including the requirements of unique dimensions and characteristics and no added dimensions in the final iteration. In Step 5, the evaluation goals regarding reliability, robustness, and completeness of the taxonomy are determined.

Regarding the iterative design of the taxonomy (Steps 6-10), five iterations are defined, of which we have conducted the first three (Figure 1; subsequent iterations are subject to change, and further ones may be added). For Iteration 1, we chose a conceptual-to-empirical approach following Kundisch et al. (2022) since the authors are active and knowledgeable regarding uncertainty in AI and XAI research due to their years of work experience in this field. Multiple author discussions, also based on literature overviews, were conducted over the course of several weeks and yielded an initial taxonomy that formed the basis for the following iterations.[1] To incorporate the analysis of objects under consideration (i.e., works on uncertainty in AI-augmented decision-making), an explorative literature search was conducted as Iteration 2. It aimed to identify relevant works without necessarily focusing on XAI. Such a broad scope was considered advantageous, as the literature was much sparser if a focus on XAI was mandatory, potentially leading to the omission of unaddressed but important issues. Overall, 143 papers were considered relevant. Multiple discussions incorporating the new knowledge gained from the analysis were conducted among the authors, which led to a further refinement of the taxonomy (Iteration 3). The results presented in this short paper are based on our findings after Iteration 3. We plan to conduct a structured comprehensive literature search on the intersection of XAI and uncertainty (Iteration 4) as well as interviews and focus groups with 12 external experts (Iteration 5). Once the ending conditions are confirmed, the taxonomy will be evaluated and reported (Steps 11-18).

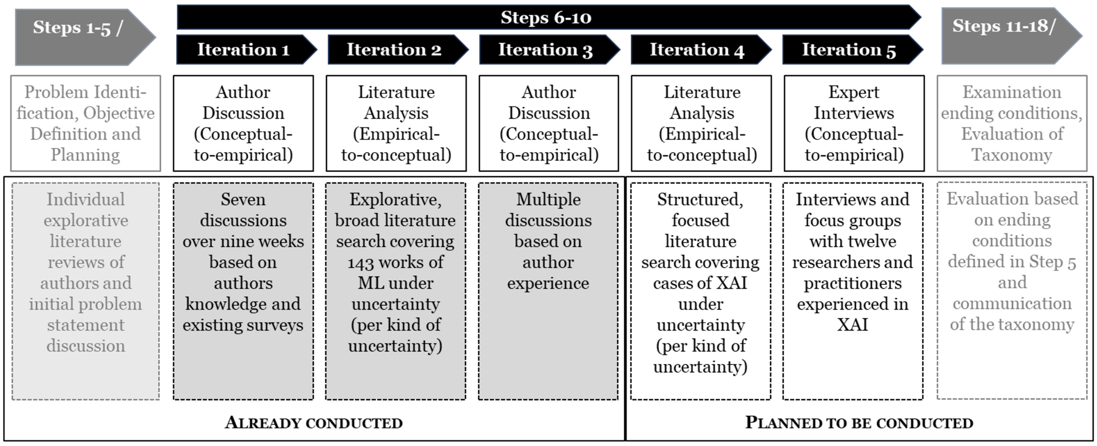

Figure 1. Application of the ETDP: Iterative Steps Focusing on Taxonomy Design (Steps 6-10).

3 Preliminary Results of the Explorative Literature Analysis

In this section, we present the results of our literature analysis conducted during Iteration 2 (see Figure 1). While existing literature offers taxonomies on uncertainty in decision-making in general (e.g., Lovell, 1995), on uncertainty in AI models (excluding XAI, e.g., Gawlikowski et al., 2023; DIN, 2024) and on XAI (excluding uncertainty; e.g., Schwalbe & Finzel, 2021), to the best of our knowledge, there is no comprehensive taxonomy or unified perspective on uncertainty-aware XAI. Based on our analysis of the fragmented literature and discussions among the authors of the taxonomy design, we identified four primary sources of uncertainty in AI-augmented decision-making: uncertainties arising from the data, the AI model, the XAI method, and the human. The results of our explorative literature search for works addressing the four sources of uncertainty (cf. Iteration 2) are described in the following.

3.1 Data Uncertainty

Data uncertainty resulting from potential data quality defects can take many different forms. It is well known that AI model performance decreases substantially under data uncertainty (Padmanabhan et al., 2022). Hence, impaired data quality represented by data uncertainty is commonly examined in AI contexts. However, data uncertainty can persist or may even intensify, e.g., if missing data values are imputed incorrectly, or the imputation introduces systematic bias. To address these issues, studies provide different perspectives and overviews on data uncertainty in AI. Gudivada et al. (2017), for instance, offer a broad overview, discussing different aspects of data uncertainty as challenges when applying AI on big data. Poor data quality as well as dataset quality in general can be an issue (Fouad

[1] Material documenting our methodological approach can be found in the Online Appendix at https://osf.io/cz9ja/.

et al., 2021). While these works provide valuable insights and overviews into data uncertainty in AI contexts, they do not address the relation of data uncertainty to other types of uncertainties, nor do they explore how such uncertainties interact with explanations. The relationship between data uncertainty and XAI has been examined only by Bertossi & Geerts (2020); however, as the authors state, their goal is "not a comprehensive survey" or taxonomy but "to provide a bit of insight into this general problem and to identify a couple of promising research directions" (Bertossi & Geerts 2020, p. 1).

3.2 AI Model Uncertainty

The second major source of uncertainty stems from AI models themselves, as they only approximate the real (functional) relationship between model input and output. Indeed, these models are associated with uncertainty, even when trained and tested on perfect data (i.e., in the absence of data uncertainty) (Kläs & Vollmer, 2018). More precisely, according to Zhou et al. (2022), AI model uncertainty refers "to a state that model cognition is restricted, which is due to the upper limit of the model fitting ability, the optimizing strategy, the parameters, the lack of knowledge" (Zhou et al., 2022, pp. 449–450). While only few existing methods for assessing AI model uncertainty provide explanations of uncertainty to human users (Gawlikowski et al., 2023), Guo et al. (2024) design a taxonomy that examines the relationship between AI model uncertainty and XAI. However, their work focuses on abstract mathematical formalizations and is not directly applicable to the broad majority of existing machine learning-based AI. In summary, research has yet to explore how AI model uncertainty interacts with other uncertainties or how it can be included in uncertainty-aware explanations.

3.3 XAI Method Uncertainty and Uncertainty-Aware Explanations

The development of XAI methods and their empirical evaluation for different user groups has received popular attention in recent years; however, the idea of integrating uncertainty in explanations is only emerging (Bobek & Nalepa, 2021). Accordingly, while various taxonomies on XAI exist (Meske et al., 2022), a systematic conceptualization and analysis of uncertainty-aware XAI are still missing. Existing studies on uncertainty-aware XAI consider only single sources of uncertainty and can be categorized according to whether they focus on I) data or AI model uncertainty or ii) XAI method uncertainty. With respect to data or AI model uncertainty, a few studies develop and investigate approaches that automatically generate first ideas for uncertainty-aware XAI along AI predictions, e.g., by displaying data uncertainty in feature attributions or by generating uncertainty-aware counterfactual explanations (Wang et al., 2021). With respect to XAI method uncertainty, e.g. Zhang et al. (2019) argue that explanations from XAI methods can be uncertain (e.g., caused by their stochastic elements). Initial studies, such as Marx et al. (2023), aim to develop approaches to quantify and reduce uncertainty of explanations. These initial, promising studies address selected aspects of uncertainty but do not consider a holistic perspective and potential interdependencies among uncertainties. Thus, there is a need for an integrated framework that conceptualizes uncertainty-aware XAI.

3.4 Human Uncertainty

While prior research has discussed data, AI model, and XAI method uncertainties to some extent, less attention has been given to human users as a source of uncertainty in AI-augmented decision-making. This includes how humans perceive and interpret various uncertainties and how their own uncertainty and traits interact with other uncertainty sources. A structured taxonomy of human factors impacting uncertainty in AI-augmented decision-making is still missing. Key studies emphasize skills like data interpretation and understanding the role of training data in predictions (Long & Magerko, 2020), with uncertainties arising from tasks, AI models, or data (Fügener et al., 2021). Risk attitudes and perceived uncertainty shape responses to AI, affecting trust and skepticism (Schecter et al., 2023). Data literacy, AI literacy, and expertise (Bauer et al., 2023b), and cognitive and metacognitive processes such as mental model development and responses to AI errors (Bansal et al., 2019) play crucial roles in how individuals rely on AI predictions. Further, users find AI more valuable in complex tasks, though trust remains stable despite unpredictable outcomes (Salimzadeh et al., 2024). AI advice may reduce human uncertainty but also risks overconfidence (Taudien et al., 2022). Disclosing AI model certainty scores aids users in calibrating their uncertainty and aligning mental models; but at the same time, can cause self-fulfilling prophecies (Bauer & Gill, 2024). While these studies provide valuable insights into specific aspects of making uncertainty more transparent to humans, they do not incorporate the other sources of uncertainty as discussed previously and they do not investigate the broader influence of uncertainty on human-AI interaction in order to develop a deeper understanding of human uncertainty in an AI-augmented decision-making process.

4 Preliminary Results of Taxonomy Development

The main result after Iteration 3 is a hierarchical taxonomy for uncertainty-aware XAI, as illustrated in Figure 2. As emphasized by Nickerson et al. (2013), a robust taxonomy ensures clear differentiation among objects under consideration. This involves incorporating several dimensions and characteristics to maximize distinctiveness and minimize overlaps. Our taxonomy comprises four top-level dimensions, seven dimensions, and 38 sub-dimensions. From the main steps of AI-augmented decision-making we derive the four top-level dimensions. Figure 2 shows the top-level dimensions in dark gray: data uncertainty, AI model uncertainty, uncertainty-aware explanations, and human.

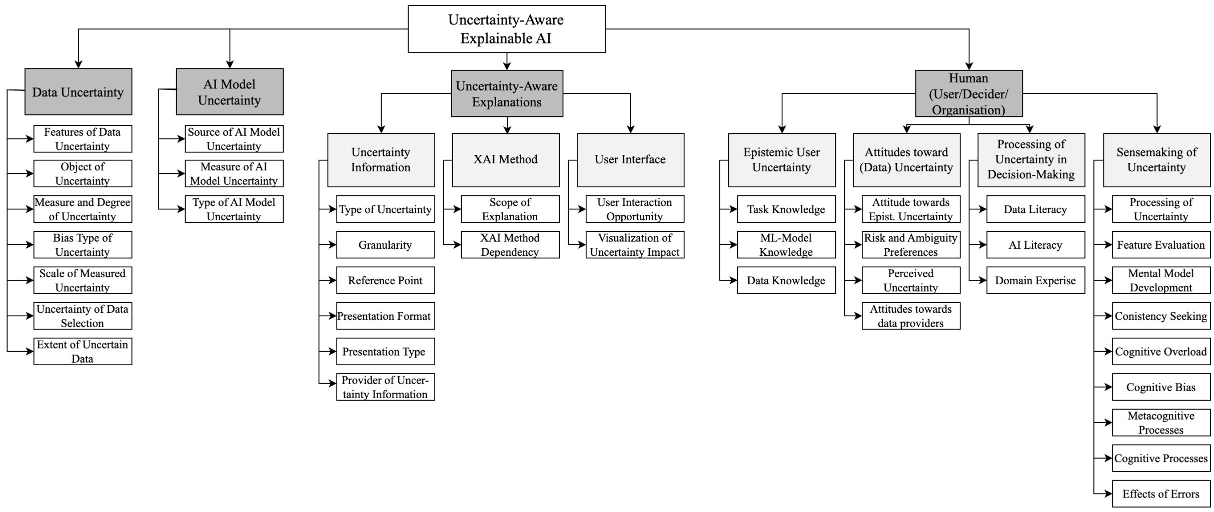

Figure 2. Preliminary Taxonomy for Uncertainty-Aware Explainable AI.

Data Uncertainty: For the first top-level dimension, we identify seven dimensions. First, the *features of data uncertainty*, such as accuracy, currency, completeness, consistency, and class imbalance. Second, the *object of uncertainty*, that is, whether uncertainty pertains to the attributes, attribute values, instances, labels, or other data elements (Fang, 2021). Third, the *measures and degrees of uncertainty* (e.g., share of missing data) (Gao & Wang, 2010). Fourth, understanding *bias types*, ranging from randomness to non-randomness (Fang, 2021), is crucial. Fifth, the *scale of measured uncertainty*, i.e., whether nominal, ordinal, or interval-scaled (Fouad et al., 2021). Sixth, *data selection* methods, such as fold selection or train-test-split, introduce their own effect on data uncertainty. Finally, the *extent of uncertain data* across various stages, including training, validation, and test/use phases, is important for assessing the effect of data uncertainty on the overall fidelity of AI predictions and their explanations.

AI Model Uncertainty: The second top-level dimension considers three dimensions of uncertainty in AI models. First, the *source of model uncertainty* identifies which of the various facets of the AI model and the steps of the AI pipeline induce uncertainty. This includes, for example, model assumptions, architecture and functioning, parameters, training procedures, and learning (Kläs & Vollmer, 2018). Second, the *measure of model uncertainty* indicates that model uncertainty can be assessed at the global (i.e., for the model as a whole) or at the local level (i.e., for a specific prediction) (Gawlikowski et al., 2023). Third, the *type of model uncertainty* refers to the nature of the induced model uncertainty as a systematic bias or a random (i.e., unsystematic) variance (Zhou et al., 2022).

Uncertainty-Aware Explanations: We derive three dimensions that resemble the main building blocks necessary for uncertainty-aware explanations, which take a central role as they intermediate between the AI and the human users. The first dimension of *uncertainty information* encompasses several subdimensions. First, the *type of uncertainty* denotes whether the uncertainty pertains to the data, model (and thus, prediction), explanation or human user (Schwab & Karlen, 2019). Further, the *information granularity* (aggregated or fine-granular), and possible *reference points* (exogenous or endogenous) can affect the human perception and processing of the uncertainty information. The *presentation format*, ranging from quantitative to qualitative or implicit presentation (Patro et al., 2019), in combination with the *presentation type* such as visual, textual, or numeric information (Bykov et al., 2020), offers different options to convey uncertainty information to the user. In addition, the *identity of the provider* of *uncertainty information* (first-party vs. third-party) is likely to affect how users perceive and react to uncertainty information, although this has not been analyzed in the literature so far. The second dimension on the properties of the *XAI method* includes subdimensions on the *scope of the explanation* (local or global) and the *model dependency of the XAI method* (model-agnostic or model-specific) (Brasse et al., 2023). Finally, the *user interface* as the third dimension is important for users' perception and processing of uncertainty information as it determines the *opportunities for user interaction* with the explanation (Chromik & Butz, 2021) and may *visualize the impact of uncertainty.*

Human: We derive four dimensions that significantly shape the process and outcomes of AI-augmented decision-making under uncertainty. First, *epistemic user uncertainty* characterizes the (partial) lack of knowledge of humans (Tversky & Kahneman, 1974) about their *metaknowledge of the task and AI model* or the underlying *data* (Fügener et al., 2021). Users may exhibit varying degrees of knowledge across these domains, e.g., due to different experience levels or domain expertise. Second, *attitudes toward data uncertainty* encompass *risk and ambiguity preferences*, *perceptions of uncertainty*, and *attitudes toward the data provider* (Schecter et al., 2023). Third, the *processing of uncertainty in decision-making* is shaped by a user's personal traits. Most notably, this includes *data literacy* (Schneider et al., 2023), *AI literacy* (Long & Magerko, 2020), and *domain expertise* (Bauer et al., 2023a), which have an important impact on how users perceive and interpret uncertainty information. Finally, *sensemaking of uncertainty* in specific decision situations involves various cognitive processes, such as *uncertainty processing*, *feature evaluation*, and *mental model development*. In addition, *consistency-seeking behaviors*, *cognitive biases, (meta)cognitive processes*, and the *impact of errors* are other critical factors that shape the outcomes of AI-augmented decision-making (Abdel-Karim et al., 2023).

5 Discussion: Contribution & Future Research Plan

Our paper presents a preliminary taxonomy for the systematization of uncertainty-aware XAI. Previous research considered different sources of uncertainty in isolation. We present a first step to synthesize the body of knowledge for an integrated view of uncertainty-aware XAI by designing a taxonomy. The next steps for our research are the further iterations of the EDTP. Iteration 4 will focus on a structured literature review examining the intersections of the four uncertainties with XAI. Iteration 5 will integrate external expert knowledge through interviews and focus groups. As our current preliminary taxonomy does not yet guarantee completeness, additional iterations will be conducted until ending criteria of the taxonomy design process are met. We will then conduct a comprehensive evaluation of the proposed taxonomy, use it to identify research trends and gaps, and develop a research agenda for uncertainty-aware XAI. Our taxonomy contributes to *theory* by facilitating a comprehensive understanding of uncertainty-aware XAI, enabling the development and evaluation of more reliable AI models and XAI methods, and providing a systematized foundation for the impact of uncertainty on AI-augmented decision-making. The employed taxonomy approach (adapted from Kundisch et al., 2022) provides a broader scope than a structured literature analysis alone. In particular, it is designed to uncover underlying relationships through an iterative design process, thereby enabling sensemaking and supporting future theory development. As a result, and as a limitation of our preliminary results, we did not adhere to the PRISMA reporting guidelines, which are specifically designed for systematic reviews. In *practice*, AI applications in high-stakes domains, such as healthcare, benefit from the integrated view of XAI and uncertainty-awareness to advance the responsible use of AI. For example, while existing XAI methods may enhance transparency and explainability in complex decision situations, these approaches do not yet account for the presence and implications of uncertainty. By making XAI methods uncertainty-aware provides human users with more transparent AI predictions, which is likely to promote confidence and trust in AI in high-stakes contexts.

References

Abdel-Karim, B., Pfeuffer, N., Carl, K., & Hinz, O. (2023). How AI-Based Systems Can Induce Reflections: The Case of AI-Augmented Diagnostic Work. *MIS Quarterly*, *47*(4), 1395-1424.

Bansal, G., Nushi, B., Kamar, E., Lasecki, W. S., Weld, D. S., & Horvitz, E. (2019). Beyond Accuracy: The Role of Mental Models in Human-AI Team Performance. *Proceedings of the AAAI Conference on Human Computation and Crowdsourcing*, *7*, 2–11.

Bauer, K., & Gill, A. (2024). Mirror, Mirror on the Wall: Algorithmic Assessments, Transparency, and Self-Fulfilling Prophecies. *Information Systems Research*, *35*(1), 226–248.

Bauer, K., Zahn, M. von, & Hinz, O. (2023a). Expl(AI)ned: The Impact of Explainable Artificial Intelligence on Users' Information Processing. *Information Systems Research*, *34*(4), 1582–1602.

Bauer, K., Zahn, M. von, & Hinz, O. (2023b). Please take over: XAI, delegation of authority, and domain knowledge. *SAFE Working Paper No. 394*.

Bertossi, L., & Geerts, F. (2020). Data Quality and Explainable AI. *Journal of Data and Information Quality*, *12*(2), 1–9.

Bobek, S., & Nalepa, G. J. (2021). Introducing Uncertainty into Explainable AI Methods. In *Proceedings of the ICCS 2021, Krakow, Poland* (pp. 444–457). Springer.

Brasse, J., Broder, H. R., Förster, M., Klier, M., & Sigler, I. (2023). Explainable artificial intelligence in information systems: A review of the status quo and future research directions. *Electronic Markets*, *33*(26).

Bykov, K., Höhne, M. M.-C., Müller, K.-R., Nakajima, S., & Kloft, M. (2020). *How Much Can I Trust You? -- Quantifying Uncertainties in Explaining Neural Networks.*

Chromik, M., & Butz, A. (2021). Human-XAI Interaction: A Review and Design Principles for Explanation User Interfaces. In *Lecture Notes in Computer Science. Human-Computer Interaction – INTERACT 2021* (Vol. 12933, pp. 619–640). Springer.

DIN. (2024). Artificial Intelligence - Uncertainty quantification in machine learning. DIN SPEC 92005.

Eiband, M., Buschek, D., Kremer, A., & Hussmann, H. (2019). The Impact of Placebic Explanations on Trust in Intelligent Systems. In *Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems* (pp. 1–6).

Fang, J. (2021). Feature screening for ultrahigh-dimensional survival data when failure indicators are missing at random. *Statistical Papers*, *62*(3), 1141–1166.

Fouad, K. M., Ismail, M. M., Azar, A. T., & Arafa, M. M. (2021). Advanced methods for missing values imputation based on similarity learning. *PeerJ. Computer Science*, *7*, e619.

Fügener, A., Grahl, J., Gupta, A., & Ketter, W. (2021). Cognitive Challenges in Human–Artificial Intelligence Collaboration: Investigating the Path Toward Productive Delegation. *Information Systems Research*, *33*(2), 678–696.

Gao, C., & Wang, J. (2010). Direct mining of discriminative patterns for classifying uncertain data. In *16th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*.

Gawlikowski, J., Tassi, C. R. N., Ali, M., Lee, J., Humt, M., Feng, J., Kruspe, A., Triebel, R., Jung, P., Roscher, R., Shahzad, M., Yang, W., Bamler, R., & Zhu, X. X. (2023). A survey of uncertainty in deep neural networks. *Artificial Intelligence Review*, *56*(S1), 1513–1589.

Gudivada, V., Apon, A., & Ding, J. (2017). Data quality considerations for big data and machine learning: Going beyond data cleaning and transformations. *International Journal on Advances in Software*, *10*(1), 1–20.

Guo, Z., Wan, Z., Zhang, Q., Zhao, X., Zhang, Q., Kaplan, L. M., Jøsang, A., Jeong, D. H., Chen, F., & Cho, J.-H. (2024). A survey on uncertainty reasoning and quantification in belief theory and its application to deep learning. *Information Fusion*, 101987.

Kläs, M., & Vollmer, A. M. (2018). Uncertainty in Machine Learning Applications: A Practice-Driven Classification of Uncertainty. In *Proceedings of the International Conference on Computer Safety, Reliability, and Security* (pp. 431–438). Springer.

Kundisch, D., Muntermann, J., Oberländer, A. M., Rau, D., Röglinger, M., Schoormann, T., & Szopinski, D. (2022). An Update for Taxonomy Designers. *Business & Information Systems Engineering*, *64*(4), 421–439.

Long, D., & Magerko, B. (2020). What is AI Literacy? Competencies and Design Considerations. In *2020 CHI Conference on Human Factors in Computing Systems,* Honolulu, USA.

Lovell, B. E. (1995). A taxonomy of types of uncertainty. Portland State University.

Marx, C., Park, Y., Hasson, H., Wang, Y., Ermon, S., & Huan, L. (2023). But Are You Sure? An Uncertainty-Aware Perspective on Explainable AI. In *Proceedings of The 26th International Conference on Artificial Intelligence and Statistics* (Vol. 206, pp. 7375–7391). PMLR.

Meske, C., Bunde, E., Schneider, J., & Gersch, M. (2022). Explainable Artificial Intelligence: Objectives, Stakeholders, and Future Research Opportunities. *Information Systems Management*, *39*(1), 53–63.

Nickerson, R. C., Varshney, U., & Muntermann, J. (2013). A method for taxonomy development and its application in information systems. *European Journal of Information Systems*, *22*(3), 336–359.

Padmanabhan, B., Sahoo, N., Burton-Jones, A., & others (2022). Machine learning in information systems research. *Management Information Systems Quarterly*, *46*(1), iii–xix.

Pathirannehelage, S. H., Shrestha, Y. R., & von Krogh, G. (2024). Design principles for artificial intelligence-augmented decision making: An action design research study. *European Journal of Information Systems*, 1-23.

Holton, G. A. (2004). Defining risk. *Financial Analysts Journal*, 60(6), 19-25.

Hüllermeier, E., & Waegeman, W. (2021). Aleatoric and epistemic uncertainty in machine learning: An introduction to concepts and methods. *Machine Learning*, 110(3), 457-506.

Patro, B., Lunayach, M., Patel, S., & Namboodiri, V. (2019). U-CAM: Visual Explanation Using Uncertainty Based Class Activation Maps. In *2019 IEEE/CVF International Conference on Computer Vision (ICCV)* (pp. 7443–7452). IEEE.

Salimzadeh, S., He, G., & Gadiraju, U. (2024). Dealing with Uncertainty: Understanding the Impact of Prognostic Versus Diagnostic Tasks on Trust and Reliance in Human-AI Decision Making. In *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems* (pp. 1–17).

Schecter, A., Bogert, E., & Lauharatanahirun, N. (2023). Algorithmic appreciation or aversion? The moderating effects of uncertainty on algorithmic decision making. In *Extended Abstracts of the 2023 CHI Conference on Human Factors in Computing Systems*.

Schwab, P., & Karlen, W. (2019). CXPlain: Causal Explanations for Model Interpretation under Uncertainty. In *Advances in Neural Information Processing Systems* (Vol. 32). Curran.

Schwalbe, G., & Finzel, B. (2024). A comprehensive taxonomy for explainable artificial intelligence: a systematic survey of surveys on methods and concepts. *Data Mining and Knowledge Discovery*, 38(5), 3043-3101.

Taudien, A., Fuegener, A., Gupta, A., & Ketter, W. (2022). Calibrating Users' Mental Models for Delegation to AI. In *Forty-Third International Conference on Information Systems*.

Thompson, A., Jagan, K., Sundar, A., Khatry, R., Donlevy, J., Thomas, S., & Harris, P. (2022). *Uncertainty evaluation for machine learning.*

Tversky, A., & Kahneman, D. (1974). Judgment under Uncertainty: Heuristics and Biases. *Science*, *185*(4157), 1124–1131.

Wang, D., Zhang, W., & Lim, B. Y. (2021). Show or suppress? Managing input uncertainty in machine learning model explanations. *Artificial Intelligence*, *294*, 103456.

Zhang, Y., Song, K., Sun, Y., Tan, S., & Udell, M. (2019). *"Why Should You Trust My Explanation?" Understanding Uncertainty in LIME Explanations.*

Zhou, X., Liu, H., Pourpanah, F., Zeng, T., & Wang, X. (2022). A survey on epistemic (model) uncertainty in supervised learning: Recent advances and applications. *Neurocomputing*, *489*, 449–465.