Titel:

Human-Centered Design and Evaluation of Explanation User Interfaces – A Design Science Research Perspective

Inauguraldissertation zur Erlangung des akademischen Grades „Doktor der Wirtschaftswissenschaft“ des Fachbereichs Wirtschaftswissenschaft der Freien Universität Berlin

Vorgelegt von:	Enrico Bunde, M.Sc.
Geburtsort:	Brandenburg an der Havel
Jahr der Einreichung:	2023

Enrico Bunde

Human-Centered Design and Evaluation of Explanation User Interfaces – A Design Science Research Perspective

Gutachter

1. Univ.-Prof. Dr. Christian Meske, Ruhr-Universität Bochum, Germany
2. Univ.-Prof. Dr. Martin Gersch, Freie Universität Berlin, Germany
3. Univ.-Prof. Dr. Mathias Klier, Universität Ulm, Germany

Disputation

Berlin, 23.04.2024

Kurzfassung der Ergebnisse (Deutsch)

Künstliche Intelligenz (KI) ist nicht mehr aus unseren privaten und beruflichen Leben wegzudenken. In Form von persönlichen Assistenten unterstützen sie uns dabei Aufgaben zu erfüllen oder nehmen sie uns komplett ab. Implementiert in Entscheidungsunterstützungssystemen kann KI dazu beitragen, dass menschliche Entscheidungsträger:innen informierte Entscheidungen treffen können. Weiterhin kann KI vielseitige Dimensionen wie die kognitive Anstrengung oder Belastung und die Leistung positiv beeinflussen oder auch zu Zeitersparnissen führen. Spätestens die Veröffentlichung von ChatGPT hat gezeigt, wie mächtig aktuelle KI-Ansätze sein können und welche weitreichenden Konsequenzen sowie Implikationen KI-Systeme entstehen lassen. Doch viele dieser mächtigen KI-Ansätze teilen sich eine kritische Eigenschaft, welche zu großen Herausforderungen für die Gestaltung, Evaluierung und Nutzung von KI-Systemen führen kann. Dabei handelt es sich um das Blackbox-Problem. Viele neuartige und hochleistungsfähige KI-Ansätze basieren auf komplexen Architekturen und beinhalten viele nicht-lineare Berechnungen, wodurch diese an Intransparenz gewinnen und daher häufig als opak beschrieben werden. Durch diesen Umstand wissen wir, welche Daten in die KI als Eingabe eingegangen sind und bspw. im Falle einer Klassifikation erhalten wir eine Ausgabe in Form einer Klasse. Das Blackbox-Problem liegt zwischen der Ein- und Ausgabe. In Blackbox-KI-Ansätzen können wir nur schwer nachvollziehen, welche Merkmale der Eingabedaten besonders wichtig oder unwichtig für die Ausgabe waren. Wir können nicht einschätzen, ob die KI für uns sinnvolle oder unerwünschte Merkmale für die Entscheidungsfindung nutzt, welche bspw. zu Diskriminierung oder einer kognitiven Verzerrung führen können. In unkritischen Aufgaben wie bspw. der Sentiment-Klassifizierung von Rezensionen in positiv und negativ, sind die Auswirkungen des Blackbox-Problems eher gering. Wenn KI aber in eher kritischen Bereichen eingesetzt wird wie im Bereich der Gesundheitsversorgung, im Personal- oder Finanzwesen, steigen die potenziellen Risiken, welche aus dem Blackbox-Problem entstehen und können weitreichende Folgen haben.

Das Forschungsfeld erklärbare KI (EKI) widmet sich diesem Blackbox-Problem. In diesem Forschungsfeld werden Methoden entwickelt, welche Blackbox-KI erklärbar machen sollen, bspw. durch die Generierung von Erklärungen für individuelle KI-Ausgaben. Gleichermaßen wird daran geforscht, transparente KI-Modelle und Architekturen zu entwickeln, um das Blackbox-Problem erst gar nicht aufkommen zu lassen. Ein weiterer wichtiger Forschungsfokus im Bereich EKI besteht darin, zu untersuchen und verstehen, wie Erklärungen gestaltet und wahrgenommen werden. Forschung hat bereits gezeigt, dass es keine allgemeingültige Lösung gibt, welche für alle EKI-Systeme gleichermaßen funktionieren. Denn es gibt eine Vielzahl von Branchen und Anwendungsszenarien für EKI, wobei die Nutzer:innen sehr individuelle Bedürfnisse, Vorkenntnisse und Erwartungen an Erklärungen stellen können. Es gibt also eine große Anzahl von Aspekten, welche die Wahrnehmung von Erklärungen und die Interaktionserfahrung beeinflussen können. Diese Herausforderung wird dadurch verstärkt, dass Erklärungen in den unterschiedlichsten Formaten für Nutzer:innen gestaltet werden können, wie bspw. Erklärungen in natürlicher Sprache, die grafische Hervorhebung relevanter Eingabedaten oder auch die Visualisierung in Form von Diagrammen. In diesem Spannungsfeld gibt es viele Möglichkeiten, mittels wissenschaftlicher Erkenntnisgewinnung einen wertvollen Beitrag zu leisten. Denn das Forschungsfeld EKI wird in der Forschung häufig dadurch charakterisiert, dass mehr human-zentrierte Evaluierungen sowie dazugehöriges Designwissen fehlt.

Die Aspekte der humanzentrierten Gestaltung und Evaluierung wurden in dieser Dissertation erforscht. Als Fokus wurden die User Interfaces (UIs) von EKI-Systemen ausgewählt, da sie ein elementarer Bestandteil der Interaktionserfahrung darstellen und hier die Erklärungen den Nutzer:innen bereitgestellt werden. Diese Klasse der UIs wird Explanation UI (XUI) genannt. XUIs sind

jene UIs, welche Informationen über die Ausgabe eines EKI-Systems präsentieren und einen besonderen Fokus auf die Erklärungen legen. So präsentieren XUIs bspw. die Eingabedaten, die Ausgabe, eine oder mehrere Erklärungen und weitere Informationen, welche bei der Interpretation der Erklärungen unterstützen. Die Erforschung der human-zentrierten Gestaltung und Evaluierung von XUIs wurde in dieser Dissertation maßgeblich mittels des Design Science Research Paradigmas durchgeführt. In den individuellen Forschungsarbeiten, welche Teil dieser Dissertation sind, wurde die Gestaltung und Wahrnehmung von XUIs in unterschiedlichen Domänen untersucht. Dazu gehören Domänen wie die Gesundheitsbranche, Mobilitätsbranche, soziale Medien oder Personalmanagement. Neben den unterschiedlichen Branchen wurden viele verschiedene Akteure in die Gestaltung von human-zentrierten XUIs involviert. Hierzu gehören bspw. Domänenexpert:innen, Nutzer:innen oder von EKI-Ausgaben betroffene Personen. Durch die Involvierung von diversen Akteuren fließen vielseitige Erwartungshaltungen, Informationsbedürfnisse und Vorerfahrungen in den Gestaltungsprozess von XUIs ein. Im Rahmen des Gestaltungsprozesses wurden die Akteure auch in human-zentrierten Evaluierungen involviert. Dabei kamen unterschiedliche Forschungsmethoden zum Einsatz. Einerseits wurden qualitative Methoden verwendet und insbesondere semi-strukturierte Interviews durchgeführt. Andererseits wurden quantitative Methoden wie Onlineexperimente und Umfragen durchgeführt. Durch diese human-zentrierten Evaluierungen konnten viele tiefgehende Erkenntnisse gewonnen werden. Die gewonnenen Erkenntnisse beziehen sich auf die Problemidentifizierung mit existierenden vergleichbaren Systemen, Verbesserungspotenziale für die gestalteten XUIs, die Wahrnehmung sowie Effekte der gestalteten XUIs auf Nutzer:innen und die Wiederverwendbarkeit des entwickelten Designwissens.

Die individuellen Forschungsprojekte und das zuvor beschriebene Vorgehen führten dazu, dass ein umfangreiches Wissen hinsichtlich der human-zentrierten Gestaltung und Evaluierung entwickelt wurde. Die Forschungsergebnisse haben gezeigt, wie Erklärungen in XUIs bspw. die Vertrauenswürdigkeit, Nützlichkeit, Interaktivität, Nutzerbindung oder Zufriedenheit positiv beeinflussen können. Weiterhin wurde das entwickelte Designwissen in Form von Designprinzipien zusammen mit Praktiker:innen evaluiert. Dabei lag der Fokus auf der Wiederverwendbarkeit und die Evaluierungen haben gezeigt, dass sie als sehr hoch bewertet wird. Zusätzlich gaben die Nutzer:innen stets einen hohen Zuspruch hinsichtlich der Akzeptanz der Designprinzipien für eigene Projekte an oder sie Kollegen zu empfehlen. Die in den individuellen Projekten generierten Erkenntnisse und Wissensbeiträge wurden im Rahmen der kumulativen Dissertationsschrift zusammengeführt. Um dies zu erreichen, werden die übergeordneten Forschungsfragen präsentiert, welche sich in den individuellen Forschungsarbeiten widerspiegeln. Es werden die Wissensbasis und das Begründungswissen präsentiert, welche die Basis der durchgeführten und präsentierten Forschung darstellen. Die in den unterschiedlichen Projekten verwendeten Forschungsmethoden werden ebenfalls präsentiert. Anschließend werden alle im Rahmen der Dissertation gewonnen wissenschaftlichen Erkenntnisse mit einem Bezug zu human-zentrierter Gestaltung und Evaluierung von XUIs in einer Information Systems Design Theory (ISDT) zusammengefasst. Durch die entwickelte ISDT werden die erarbeiteten Erkenntnisse zusammengeführt und zugänglich gemacht. Hierdurch entstehen vielfältige Wissensbeiträge mit Relevanz für die Forschung und Praxis. Dazu gehören bspw. die in der ISDT enthaltenen Designprinzipien, welche Wissen dazu bereitstellen, wie sich konkrete Designkonfigurationen von XUIs implementieren lassen. Hinzu kommen die empirischen Erkenntnisse, welche einen Einblick hinsichtlich der Effekte von unterschiedlichen Designkonfigurationen auf die Interaktionserfahrung für Nutzer:innen mit XUIs gewähren.

Summary of the Results (English)

Artificial intelligence (AI) has become integral to our private and professional lives. In the form of personal assistants, they support us in completing tasks or take them off our hands thoroughly. Implemented in decision support systems, AI can help human decision-makers make informed decisions. Furthermore, AI can positively influence various dimensions such as cognitive effort or load and performance or even lead to time savings. The publication of ChatGPT has shown how powerful current AI approaches can be and the far-reaching consequences and implications that AI systems can have. However, many of these powerful AI approaches share a critical property that can lead to significant challenges in designing, evaluating, and using AI systems. This is the black box problem. Many novel and high-performance AI approaches are based on complex architectures and involve many non-linear calculations, making them less transparent and, therefore, often described as opaque. Due to this fact, we know which data has been entered into the AI as input, and, for example, in the case of a classification, we receive an output in the form of a class. The black box problem lies between the input and output. In black box AI approaches, it is difficult for us to understand which features of the input data were particularly important or unimportant to the output. We cannot assess whether the AI uses features that are useful or undesirable to us for decision-making, which could, for example, lead to discrimination or bias. In non-critical tasks such as the sentiment classification of reviews into positive and negative, the effects of the black box problem are relatively small. However, when AI is used in more critical areas such as healthcare, human resources, or finance, the potential risks arising from the black box problem increase and can have far-reaching consequences.

Explainable AI (XAI) research is dedicated to this black box problem. In this research field, methods are being developed to make black box AI explainable, for example, by generating explanations for individual AI outputs. At the same time, research is being carried out to develop transparent AI models and architectures in order to prevent the black box problem from arising in the first place. Another important research focus in XAI is examining and understanding how explanations should be designed and the associated effects. Research has shown that no universal solution works equally for all XAI systems. There are a variety of industries and application scenarios for XAI, whereby users can have very individual needs, previous knowledge, and expectations of explanations. There are, therefore, a large number of aspects that can influence the perception of explanations and the interaction experience. This challenge is compounded by the fact that explanations can be designed for users in various formats, such as explanations in natural language, graphical highlighting of relevant input data, or visualization in the form of diagrams. In this area of tension, there are many opportunities to make a valuable contribution by acquiring scientific knowledge. The XAI research field is often characterized by the lack of more human-centered evaluations and the associated design knowledge.

The aspects of human-centered design and evaluation were explored in this dissertation. The user interfaces (UIs) of XAI systems were chosen as a focus because they represent an elementary part of the interaction experience, and this is where the explanations are provided to the users. This class of UIs is called Explanation UI (XUI). XUIs are those UIs that present information about the output of an XAI system and place a particular focus on the explanations. For example, XUIs present the input data, the output, one or more explanations, and other information that supports the interpretation of the explanations. The research into the human-centered design and evaluation of XUIs in this dissertation was largely carried out using the Design Science Research paradigm. In the individual research work that is part of this dissertation, the design and perception of XUIs in different domains were examined. These include domains such as the healthcare industry, mobility industry, social media, and human resources management. In addition to the different industries, many different actors were involved in the design of human-centered XUIs. This includes, for example, domain experts, users, or people

affected by XAI systems. By involving various actors, diverse expectations, information needs, and previous experiences flow into the design process of XUIs. The actors were also involved in human-centered evaluations as part of the design process. Different research methods were used. On the one hand, qualitative methods were used, and, in particular, semi-structured interviews were carried out. On the other hand, quantitative methods such as online experiments and surveys were carried out. Many in-depth insights were gained through these human-centered evaluations. The insights gained relate to problem identification with existing comparable systems, the potential for improvement for the designed XUIs, the perception and effects of the designed XUIs on users, and the reusability of the developed design knowledge.

The individual research projects and the approach described above led to the development of extensive knowledge regarding human-centered design and evaluation. The research results have shown how explanations in XUIs can positively influence trustworthiness, usefulness, interactivity, user loyalty, or satisfaction. Furthermore, the design knowledge developed in design principles was evaluated by practitioners. The focus was on reusability, and the results have shown that it is rated very high. In addition, users consistently reported high acceptance of the design principles for their projects or recommended them to colleagues. The insights and knowledge contributions generated in the individual projects were brought together as part of the cumulative dissertation. In order to achieve this, the overarching research questions are presented, which are reflected in the individual research work. The knowledge base and the justification knowledge are presented, representing the basis of the research carried out and presented. The research methods used in the different projects are also presented. Subsequently, all scientific findings gained as part of the dissertation regarding human-centered design and evaluation of XUIs are summarized in an Information Systems Design Theory (ISDT). Through the developed ISDT, the knowledge gained is brought together and made accessible. This creates a wide range of relevant knowledge contributions to research and practice. These include, for example, the design principles contained in the ISDT, which provide knowledge on how concrete design configurations of XUIs can be implemented. In addition, empirical findings provide insight into the effects of different design configurations on the interaction experience for users with XUIs.

Table of Contents

1 Introduction ... 11

1.1 Research Motivation ... 11

1.2 Research Objective and Research Questions ... 14

1.3 Structure of the Cumulative Dissertation and Included Articles ... 17

2 Research Background and Knowledge Base ... 21

2.1 The Relevance of Human-Centered (X)AI ... 21

2.2 Human-Centered Design in XAI and XUI Research ... 26

2.3 Human-Centered Evaluation in XAI and XUI Research ... 28

3 Research Design ... 34

3.1 Design Science Research in Information Systems Research ... 34

3.2 Design Science Research Approach ... 36

3.3 Further Research Methods Used ... 39

4 Summary and Consolidation of the Research Contributions ... 42

4.1 Summary of the Research Contributions of the Individual Articles ... 42

4.2 Design Theories in Design Science Research ... 45

4.3 Towards an Information Systems Design Theory for Human-Centered XUIs ... 46

4.3.1 Defining the Purpose and Scope ... 46

4.3.2 Establishing the Constructs ... 47

4.3.3 Principles of Form and Function ... 48

4.3.4 Artifact Mutability ... 53

4.3.5 Testable Propositions ... 54

4.3.6 Justificatory Knowledge ... 56

4.3.7 Principles of Implementation ... 56

4.3.8 Expository Instantiations ... 58

4.3.9 The Information Systems Design Theory for Human-Centered XUIs ... 58

5 Discussion ... 63

5.1 On the Human-Centered Design of XUIs ... 63

5.2 Theoretical Contributions and Implications ... 68

5.3 Practical Contributions and Design Implications ... 69

5.4 Limitations and Future Research ... 70

6 Conclusion ... 71

7 References ... 73

8 Anhang ... 85

8.1 Erklärung §10 (3) Promotionsordnung ... 85

8.2 Articles Published in the Cumulative Dissertation ... 86

List of Figures

Figure 1. An illustration of the XAI concept (based on Gunning & Aha, 2019, p. 48). 12
Figure 2. Example of LIME for tabular data (reproduced). 22
Figure 3. Example of LIME for text data (reproduced). 22
Figure 4. Example of SHAP for text data (reproduced). 23
Figure 5. Example of Grad-CAM for image data - explanation for the class dog (left) and for the cat (right) (reproduced). 23
Figure 6. Example for SHAP and GradientExplainer on image data to visualize the 7th layer of a VGG16 (reproduced). 24
Figure 7. The second Copernican Revolution shifted from algorithms and AI to humans being the center of attention (adapted from Schneiderman, 2020a, p. 112). 26
Figure 8. Human-centered design in the field of HCI (based on Nazar et al., 2021, p. 153321). 27
Figure 9. Overarching approach in human-centered evaluation (Vilone & Longo, 2021). 31
Figure 10. Articles of the cumulative dissertation positioned in the Information Systems Research Framework (based on Hevner et al., 2004, p. 80; Hevner, 2007, p. 88). 35
Figure 11. Design knowledge production and consumption within the cumulative dissertation (based on Drechsler & Hevner, 2018; p. 89, Hevner, 2021; p. 74-75). 37
Figure 12. Relationship of the individual research projects included in the cumulative dissertation and the research questions. 42

List of Tables

Table 1. Outline of the cumulative dissertation. 19
Table 2. List of publications included in the cumulative dissertation. 20
Table 3. Human-centered design and evaluation studies focus on XAI methods, systems, and XUIs. . 33
Table 4. Overview of research methods, evaluation and data collection methods, and justificatory knowledge. 41
Table 5. Design principle of performance communication. 48
Table 6. Design principle of human-centered selection of XAI methods and presentation format. 49
Table 7. Design principle of case- and output-related comparison. 50
Table 8. Design principle of contextual information. 51
Table 9. Design principle of assistance with the interpretation. 51
Table 10. Design principle of cosmetic customization. 52
Table 11. Design principle of functional customization. 52
Table 12. Design principle of XAI method customization. 52
Table 13. An ISDT for human-centered XUIs (based on Gregor & Jones, 2007, p. 322). 63

Abbreviations

AI	Artificial Intelligence
BSR	Behavioral Science Research
cf.	Latin: confer/conferatur; English: compare
CNF	Conference
DSR	Design Science Research
e.g.	Latin: exempli gratia; English: for example
HCI	Human-Computer Interaction
HICSS	Hawaii International Conference on Systems Sciences
ICIS	International Conference on Information Systems
i.e.	Latin: id est; English: that is (to specify something)
IFS	Information Systems Frontiers
IS	Information Systems
ISR	Information Systems Research
ISM	Information Systems Management
IT	Information Technology
JNL	Journal
PhD.	Doctor of Philosophy
VHB JQ 3	Verband der Hochschullehrer für Betriebswirtschaft Jourqual 3[1]
XAI	Explainable Artificial Intelligence

[1] The VHB JQ 3 ranking is a quality rating of relevant ISR journals and conferences by the members of the Association of University Lecturers for Business Administration (German: VHB).

1 Introduction

1.1 Research Motivation

In the present time, we witness how AI, like ChatGPT, evolves into a technology that attracts users at an unprecedented rate (Dwivedi et al., 2023). Such state-of-the-art AI models and the resulting products are capable of disrupting entire industries and businesses (Yigitcanlar et al., 2022), changing our workplaces and work routines (Calisto et al., 2021; Rai et al., 2019), penetrating our private lives in the form of AI assistants (Mirchi et al., 2020), supporting us in various decision-making scenarios (Sarro et al., 2020), or even becoming our artificial friends (Einola & Khoreva, 2023). AI can collaborate with humans (Patel et al., 2019) and outperform human experts (Tschandl et al., 2019). Therefore, it is not surprising that research on AI is multidisciplinary, and the scope of application and emerging research opportunities seem almost endless (e.g., Agerfalk et al., 2022; Padmanabhan et al., 2022; Samtani et al., 2023). Domains where AI is researched include, for example, healthcare (Shad et al., 2021), finance (Weber et al., 2023), law (Haque et al., 2023), education (Li & Gu, 2023), or cyber security (Adadi & Berrada, 2018). At the same time, AI has found its way into ISR and offers exciting new research opportunities as leaps in development occur almost daily (Padmanabhan et al., 2022; Samtani et al., 2023).

With great power comes great responsibility, accompanied by many challenges surrounding AI. The challenges include, for example, the governance and regulation of AI systems (Buiten, 2019; Schneider et al., 2022), the potential deskilling or displacement of professionals that use AI systems in their work environment (Benbya et al., 2021), dangers such as the automation bias (Sujan et al., 2019), or the management of AI systems (Berente et al., 2021). The core challenge at the center of attention in this cumulative dissertation is the black box problem. While many state-of-the-art AI models achieve unprecedented performance, they often lack transparency, also called the black box problem (Adadi & Berrada, 2018). Highly performant AI approaches like deep learning neural networks are ascribed to being opaque as they consist of many millions or even a billion of parameters and non-linear functions (Angelov et al., 2021). In this context, the research field of XAI aims to make the behavior of black box AI systems more intelligible for human users by providing explanations (Gunning et al., 2019). Following Arrieta et al. (2020, p. 85), XAI can be defined as follows: “Given an audience, an explainable Artificial Intelligence produces details or reasons to make its functioning clear or easy to understand.” The relevance of XAI systems is crucial in susceptible application areas such as healthcare and places where humans are at stake, such as law, finances, or transport (Adadi & Berrada, 2018; Angelov et al., 2021; Gunning & Aha, 2019).

A broad selection of XAI methods exists, from which designers and developers can choose suitable methods to integrate into novel AI-based information systems. New and novel XAI methods are constantly introduced, mainly from computer science (Arrieta et al., 2020). Some XAI methods and AI models, categorized as globally explainable, can explain the logic of AI models and their reasoning process, leading to different possible outcomes, such as decision trees or rule lists (Adadi & Berrada, 2018). Other well-established XAI methods are post-hoc explainability methods and, more specifically, local explanations, which segment the solution space and generate explanations for less complex solution subspaces, such as Shapley explanations or saliency maps (Adadi & Berrada, 2018; Arrieta et al., 2020). Many further XAI methods exist and various ways to present explanations, for example, in the form of text and dialogue explanations, visualizations, feature relevance explanations, or example-based explanations (Arrieta et al., 2020; Leichtmann et al., 2023; van der Waa et al., 2021). The increasing relevance of XAI due to the wide dissemination of AI systems and the potentially far-reaching consequences of AI usage in different areas gives rise to a plethora of interdisciplinary research opportunities (Langer et al., 2021; Rahwan et al., 2019; Taylor & Taylor, 2021). Knowledge of how humans select or evaluate explanations, employ cognitive biases, and social expectations to explanation processes can be identified across disciplines such as philosophy, psychology, cognitive science, or computer science (Miller, 2019).

The ISR community has the opportunity to conduct behavioral- and design-oriented research (Meske et al., 2022). Moreover, the research interest from the ISR community is quite diverse and includes, for example, the management of (X)AI (Berente et al., 2021; Schneider et al., 2022), the design of (X)AI (Kane et al., 2021; Meske et al., 2022), or the collaboration of humans with (X)AI (Fügener et al., 2022). However, despite the active research efforts in the field of XAI, it is characterized by a lack of design-oriented studies (e.g., Leichtmann et al., 2023; Wells & Bednarz, 2021). Additionally, the call for more user studies or evaluations comes up repeatedly (e.g., Barda et al., 2020; van der Waa et al., 2021). This cumulative dissertation addresses both aspects by focusing on the human-centered design and evaluation of user interfaces in XAI systems, further referred to as explanation user interfaces (XUI). A well-designed XUI facilitates humans' understanding and acceptance of AI systems (Gunning et al., 2019). From the perspective of humans, they represent the entry point for meaningful interaction with AI systems (Song et al., 2020), making them an integral component of Human-Computer Interaction (HCI) and AI-human interaction (AI-HI) (Sundar, 2020). XUIs usually provide information about the AI model, including information regarding the AI model's performance, used data, contextual information, and one or more XAI features (Barda et al., 2020; Leichtmann et al., 2023). Generally, XUIs are part of a more extensive decision support system, and the provided information should support users in solving a problem or accomplishing a task (Gunning & Aha, 2019; Gunning et al., 2019). Figure 1 illustrates the XAI concept and the differences between opaque AI and XAI. The upper half of the figure demonstrates how opaque AI provides an output, but users cannot comprehend why the output resulted or which reasons have led to this specific outcome. The lower half of the figure demonstrates that XAI provides additional information via the XUI, and users can now comprehend why an output resulted.

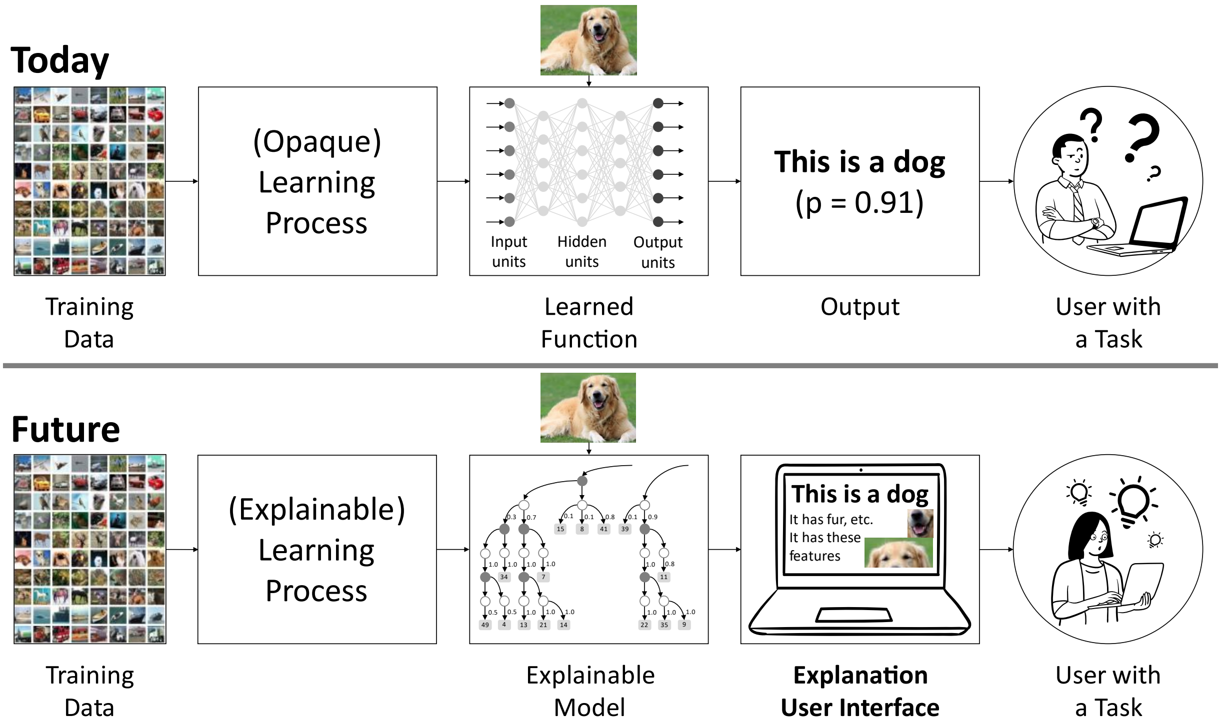


Figure 1. An illustration of the XAI concept (based on Gunning & Aha, 2019, p. 48).

Since XAI can be used in many different areas, XUIs must satisfy a diverse user base's expectations, needs, and requirements (Langer et al., 2021). The user base of XAI can be categorized into different stakeholder groups, such as regulators, developers, users or managers of AI, and individuals affected by AI-based decisions (Meske et al., 2022). Stakeholders have varying backgrounds, prior experiences with (X)AI, and motivations to use XAI (Arrieta et al., 2020). While regulators may need XAI to ensure that regulations are satisfied (Meske et al., 2022), developers can use XAI to improve the performance of the underlying AI (Adadi & Berrada, 2018), and affected individuals by AI-based decisions may want to understand an outcome (Minh et al., 2022). For example, when an XAI-based system has decision-making authority over credit applications, explanations could be used by a loan officer who communicates the outcome to the client to justify the decision and give reasons (Strich et al., 2021). Since the design of XUIs plays a significant role in building trust (Schoenherr et al., 2023) and acceptance (Gunning et al., 2019) and significantly influences the entire interaction experience, it is highly relevant to involve the relevant stakeholders in the design process (Yuan et al., 2023). The different stakeholder groups' needs and expectations can be considered using case-specific aspects. In addition, the design of XUIs and their functionalities are influenced by the choice of XAI methods, their implementation, and other XUI functionalities, which can affect the interaction experience for users in various ways (Gunning & Aha, 2019). Different explanation types can have different effects, as shown by van der Waa et al. (2020), who found that rule-based explanations positively affected system understanding, which was not the case for example-based explanations. The interaction experience is further influenced by a wide range of UI elements independent of XAI, such as interacting with the content by zooming, scrolling, or generating new content (Sundar et al., 2015). Consequently, many dimensions must be considered if XUIs are to be designed appealingly and XAI's strengths and added values are to be fully exploited. Due to all these nuanced aspects that can influence the design of and, therefore, interaction with explanations, no one-size-fits-all approach exists (Sokol & Flach, 2020).

In recent years, research on XAI has emphasized a human-centered approach when designing XAI systems, which is well-suited to involve different stakeholder groups in the design and development process (Nazar et al., 2021). Such a human-centered approach promotes human factors like self-efficacy or creativity, clarifies responsibility, and facilitates social participation (Schneiderman, 2020a). The overarching goal of a human-centered approach is to achieve reliable, safe, and trustworthy XAI systems (Schneiderman, 2020b). The concept of human-centeredness can be taken even further. From a socio-technical perspective, individuals and entire organizations should be considered (Herrmann & Pfeiffer, 2022). Consequently, when developing XAI systems with a human-centered approach, it is necessary to consider that (X)AI technologies are part of a more extensive system that includes humans, which should not be replaced but instead augmented (Jarrahi, 2018; Riedl, 2019).

This cumulative dissertation explores the human-centered design and evaluation of XUIs through a design-oriented approach, namely the DSR paradigm. For this purpose, XUIs with different design configurations in several domains were developed and evaluated in a human-centered way. Cyclic DSR frameworks and processes (e.g., Peffers et al., 2007; Kuechler & Vaishnavi, 2012) enable a structured research approach and the involvement of relevant stakeholder groups from relevant application domains (Hevner et al., 2004; Gregor & Hevner, 2013). For example, in the first phases of such iterative DSR processes, human-centered requirements for the XUI can be developed based on semi-structured interviews (e.g., Bunde et al., 2022; Meske & Bunde, 2023). Later in the iterative DSR process, when demonstrations and evaluations are necessary, developed artifacts can be evaluated qualitatively or quantitatively human-centered, for example, through semi-structured interviews and focus groups or surveys and laboratory experiments (e.g., Bunde, 2023; Bunde et al., 2022; Bunde et al., 2023; Meske & Bunde, 2023; Meske et al., 2020). In the different research projects, which are brought together in this cumulative dissertation, relevant stakeholders were involved in the evaluation, which has led to valuable opportunities and highly relevant insights, such as the optimization of requirements and design knowledge based on the feedback from human-centered qualitative evaluations or the effects of various XUI-related elements on the HCI and AI-HI measured through human-centered quantitative evaluations.

In summary, the research in this cumulative dissertation focuses on the human-centered design of XUIs and the interaction experience of relevant stakeholders with XUIs through human-centered evaluations. A design theory based on Gregor and Jones (2007) is developed to combine the design knowledge generated in the various research projects. For this purpose, a selection of the individually developed design principles is reconceptualized based on the scheme of Gregor et al. (2020). These reconceptualized design principles represent the prescriptive design knowledge that can support researchers and practitioners in developing comparable artifacts (Chandra Kruse et al., 2015). Through the human-centered evaluations in the various research projects, empirical findings are available regarding the effect different design configurations in XUIs can have on the user's interaction experience. These empirical findings are translated into propositions. The two components of design principles and propositions stand out in particular. The knowledge that supports the human-centered design of XUIs is made accessible through reusable design principles, considering concrete human-centered defined requirements. In addition, there are empirical findings from the human-centered evaluations, which show how certain design principles and design features can meet requirements and have positive effects such as increasing trustworthiness, perceived interactivity, or task performance (e.g., Bunde, 2023; Bunde et al., 2023; Meske & Bunde, 2023).

After presenting the research motivation for this cumulative dissertation, the research objective is defined in the following subsection, and concrete research questions are established.

1.2 Research Objective and Research Questions

Research on XAI is very diverse within the ISR community and ranges, for example, from the acceptance of XAI to the design and evaluation of XAI systems (Agerfalk et al., 2022; Meske et al., 2022). As already described in the research motivation, the human-centered design and evaluation of XUIs is the focus of this cumulative dissertation. In research on XAI, there are diverse research contributions. These contributions include, for example, literature overviews and conceptual research projects (e.g., Adadi & Berrada, 2018; Haque et al., 2023; Minh et al., 2022), the development of technical artifacts, such as XAI systems or XUIs (e.g., Barda et al., 2020; Herm et al., 2022; Leichtmann et al., 2023), or the investigation of the effects of XAI on users (e.g., Naisehe et al., 2023; Senoner et al., 2022; van der Waa et al., 2021). Such research contributions form the knowledge base for the DSR project, included in this cumulative dissertation. Despite all this research on XAI, there is comparatively little DSR-oriented research and, thus, scientifically developed and evaluated design knowledge in the status quo on XAI. There is also other research outside of DSR on the topic of human-centered design of XAI systems and XUIs, where design patterns and knowledge are introduced (e.g., Leichtmann et al., 2023; Schoonderwoerd et al., 2021; van der Waa et al., 2021). However, there is a significant difference between the latter design patterns and principles in rather classical software development and the design knowledge developed by DSR research.

What is unique about design knowledge in the context of DSR is that it creates a relationship between the problem and solution space, for example, by identifying a problem in the problem space and developing and evaluating a solution to solve the identified problem in the solution space (vom Brocke & Maedche, 2019; vom Brocke et al., 2020). Furthermore, design knowledge representation is fascinating because it can be about artifacts, design principles, or design theories (Gregor & Jones, 2007; Gregor et al., 2020; Hevner et al., 2004). Design knowledge has another exciting property because design knowledge can be both produced and consumed (Gregor & Hevner, 2013). A distinction is made between different knowledge bases. The prescriptive knowledge base (lambda), which describes man-made artifacts, includes, for example, constructs, models, methods, instantiations, and design theories (Gregor & Hevner, 2013). Thus, as a researcher, one would search, for example, in the prescriptive knowledge base for existing artifacts and design theories developed for a similar problem to learn from them. In addition, there are the descriptive knowledge bases (omega), which can be used, for example, to identify justificatory knowledge that matches the research goals (vom Brocke et al., 2020; Gregor & Hevner, 2013; Hevner et al., 2004). Just as the consumption of knowledge is essential for DSR projects, a scientific contribution is also produced in the form of design knowledge, which in turn can generate knowledge for both the descriptive and prescriptive knowledge base (Hevner, 2021). Knowledge for the descriptive knowledge base could be created through human-centered evaluations of instantiated artifacts, for example, laboratory experiments focusing on AI-HI. In the case of the prescriptive knowledge base, a contribution is made through developing design principles for human-centered XUIs or through the design theory introduced in this cumulative dissertation.

The design theory developed here ultimately represents an abstract, coherent body of prescriptive knowledge, representing the principles of form and function, methods, and justificatory knowledge for designing human-centered XUIs. According to Gregor (2006), such a design theory can be classified as a type five theory: design and action. Such theories provide explicit prescriptions for constructing artifacts, a desirable mature design knowledge contribution to DSR projects (Hevner et al., 2004; Gregor & Hevner, 2013). In the DSR community, there are different voices on the subject of design theory and its importance for the scientific contribution of DSR projects. However, design theories and the ISDT introduced by Gregor and Jones (2007) are a desirable form of design knowledge representation with a high research impact since such design theories are published in outlets such as the Journal of the Association for Information Systems, the European Journal of Information Systems, MIS Quarterly, or the Information Systems Journal (e.g., Avdiji et al., 2020; Coenen et al., 2018; Giessmann & Legner, 2016; Kane et al., 2021). Based on the editorial by Ivari (2020), different concepts, i.e., types of design theories, which have different goals or can even overlap in some cases, can be distinguished. The ISDT for human-centered XUIs developed here, according to Gregor and Jones (2007), could be described as a combination of types 1 and 3. For example, type 1 design theory focuses on the theoretical origin of the meta-design for human-centered XUIs (Ivari, 2007; 2020), and type 3 on the relationship between the instantiated human-centered XUI and its effectiveness in solving a defined problem (Ivari, 2020; Venable, 2006). The overarching research objective of this cumulative dissertation is to introduce comprehensive prescriptive design knowledge in the form of an ISDT for human-centered XUIs. This research has been guided by the following overarching research question (RQ1):

RQ1:	How should a design theory be constructed to provide explicit prescriptions for constructing explanation user interfaces while taking a human-centered perspective in supporting the achievement of overarching goals of explainable artificial intelligence and positively influencing the interaction experience?

An ISDT based on Gregor and Jones (2007) is developed in this synopsis to answer the research question. This design theory consists of up to eight components: six mandatory and two optional. The research projects in this cumulative dissertation, represented by individual publications, form the basis for the developed ISDT for human-centered XUIs. The design theory will be further grounded in the status quo of interdisciplinary research on human-centered design and evaluations of XAI. What makes such a design theory and the associated prescriptive design knowledge so relevant is the abstraction that leads to generalizability and thus enables the reusability of design knowledge in other domains (Gregor et al., 2020; Hevner et al., 2004; Ivari et al., 2021). Two aspects or components of the ISDT for human-centered XUIs are particularly relevant for the cumulative dissertation as scientific contributions. On the one hand, the design principles can support the design of human-centered XUIs.

Design principles also represent prescriptive design knowledge. The design principles developed in the cumulative dissertation represent a reconceptualized version of design principles from the various research projects that were brought together in the cumulative dissertation.

Design principles are essential for communicating prescriptions for constructing artifacts and, therefore, for communicating design knowledge (Hevner et al., 2004). Design principles are one well-established form of conceptualizing design knowledge as a DSR contribution type (Gregor & Hevner, 2013). From the DSR perspective, design principles can lead to DSR knowledge contributions to improvements, for example, by developing new solutions for known problems (Gregor & Hevner, 2013). They aim to provide knowledge for creating artifacts in different contexts where the artifacts belong to the same type or class (Ivari et al., 2021). The relevance of design principles is further justified by various reasons, for example, following Chandra Kruse et al. (2015, p. 4040): (i) They can be used not only to capture but also to communicate essential design knowledge; (ii) They enable the abstraction away from single instantiations and settings, which leads to generalized prescriptive knowledge; and (iii) They are a vital part of the development of more comprehensive design knowledge or design theory (Gregor & Hevner, 2013). Since design principles are prescriptive statements that convey information on what and how to build an artifact to achieve predefined design goals (Chandra Kruse et al., 2015), it is also essential to consider the individual who must act upon these prescriptive statements. These implementers apply abstract specifications as design principles to a concrete instance domain (Gregor et al., 2020). By explicitly defining the targeted implementers, which are either scholars who can adapt existing design principles or practitioners who want to design an instance of the artifact that belongs to the class of artifacts covered by the design principles, a dialogue between research and practice is stimulated, which can ultimately improve the relevance (Te'eni et al., 2017). The framework of minimum reusability evaluation of Ivari et al. (2021) was applied to evaluate the proposed design principles and communicate the associated design knowledge comprehensibly. The reusability evaluation involves potential implementers from the target community of practitioners. Therefore, a sub-goal pursued in all research projects was the development of design principles for human-centered XUIs. This goal was represented by the first sub-research question (RQ1.1):

RQ1.1:	What principles of form and function can be established by design principles that guide the design of human-centered explanation user interfaces?

Evaluating instantiated artifacts is essential in DSR projects as it provides feedback and a better understanding of the addressed problem to ultimately improve the artifact quality (Hevner et al., 2004). In evaluating the instantiated prescriptive design knowledge in the form of artifacts, the goal of generating evidence for its usefulness is pursued (Gregor & Hevner, 2013). The evaluation goals can differ depending on the use case, application context, and targeted stakeholders. Generally, the evaluation should measure how a design solution solves the identified and addressed problem (vom Brocke & Maedche, 2019). Consequently, the evaluation of design solutions and design theories is essential because of the feedback generated for further development and to ensure the rigor of the DSR project (Venable et al., 2016). The object of evaluation concerning the instantiated design knowledge in this cumulative dissertation is the XUI. The relevant design characteristics of XUIs can differ substantially, and a variety of notions of explainability can be influenced by these design characteristics, including comprehensibility, interestingness, mental fit, satisfaction, understandability, and many more (Arrieta et al., 2020; Vilone & Longo, 2021). The research articles in the cumulative dissertation identified appropriate notions of explainability in the associated knowledge base and operationalized them in human-centered evaluations.

Therefore, another aspect of significant interest for the research within the cumulative dissertation was the perceived interaction experience with the designed XUIs for relevant end user groups. This interaction experience was examined using different research methods. This included qualitative research methods, such as evaluating instantiated artifacts using semi-structured interviews and comparing different design configurations in XUIs in independent groups through online experiments. Such procedures are among the human-centered evaluation approaches (Vilone & Longo, 2021). For example, the interviews enabled optimization potentials to be identified with real end users from the examined domain, which improved the design. By conducting controlled laboratory and online experiments, the effects of design configurations in XUIs could be measured, for example, with a positive influence on dimensions such as trustworthiness, the perceived information content, the perceived interactivity, the impact on task performance, or satisfaction with the explanations. All these empirical findings regarding the interaction experience and the perception of the XUIs by actual end users from the examined and relevant domain represent an essential gain in knowledge and flow into the propositions of the ISDT for human-centered XUIs. This is the second relevant research focus, an essential part of the ISDT. The second subordinate research question specifies this subordinate research goal:

RQ1.2:	What propositions can be formulated to summarize the effects of the investigated design configurations for explanation user interfaces as measured in the human-centered evaluations?

Finally, the research foci of RQ1.1 and RQ1.2 are explored to a different extent in the research projects that are part of this cumulative dissertation. Design principles for different human-centered XUIs and design configurations were developed in different application domains, such as hate speech, fake news, or skin lesion detection. The instantiated design principles in the form of XUIs in different degrees of maturity were evaluated in a human-centered manner both qualitatively, for example, through semi-structured interviews, or quantitatively through surveys and online experiments.

In the evaluation, a distinction must also be made that, on the one hand, the reusability of design principles based on Ivari et al. (2021) was evaluated. Since the design principles focus on XUIs, stakeholders in evaluating the design principles were, in particular, software developers, user interface designers, or user experience designers. On the other hand, the instantiated artifacts were evaluated with different stakeholder groups from an end user perspective. For example, social media moderators, physicians, domain experts, or affected individuals were involved in the quantitative human-centered evaluations. Through these two subordinate research goals and thereby generated research results, a relevant knowledge contribution has been made regarding the design of human-centered XUIs and the effect of different design configurations on end users. In this cumulative dissertation, these knowledge contributions are reconceptualized, and together with the remaining components of an ISDT, a design theory for human-centered XUIs is developed.

After the research project of this cumulative dissertation was initially motivated in the first subsection, the research object and the research questions were presented in this subsection. The following subsection describes the cumulative dissertation's structure and the articles that are part of this dissertation.

1.3 Structure of the Cumulative Dissertation and Included Articles

Two aspects influenced the structure of this synopsis. On the one hand, the basic structure of DSR publications in top ISR journals was used as a template. On the other hand, much of the content and its structure were influenced by the established status quo of DSR research, including guidelines for structuring, communicating, and presenting DSR (e.g., Hevner et al., 2004; Gregor & Hevner, 2013; Gregor et al., 2020; Ivari et al., 2021).

The structure of this synopsis can be described as follows: In subsection 1.1, the first motivation for researching human-centered design and evaluation of XUIs was presented. Subsequently, the central and subordinate research goals and questions were presented in subsection 1.2. In subsection 1.3, the structure of the entire synopsis is described, and an overview of the relevant research contributions is given. The research papers are then referenced throughout the synopsis using a unique ID.

In the following Section 2, the research background and the relevant knowledge base for the cumulative dissertation are processed and presented. The content was organized into three different subsections. In subsection 2.1, the relevance of human-centered (X)AI is presented. On the one hand, the basics of XAI are described as the relevance of human-centeredness. Subsection 2.2 presents the status quo of human-centered design research in XAI and XUI research. It will show how XAI systems and XUIs can benefit from human-centeredness, with examples and added values. In the last subsection, 2.3, the subject of the human-centered evaluation of XAI and XUI research is specifically addressed. Here, general XAI evaluation approaches and human-centered approaches are presented.

In the third section, the entire research design is presented. In the first subsection, 3.1, the superordinate research procedure according to DSR is motivated, justified, and described. In a second subsection, 3.2, the literature explains and justifies the DSR methods used in the different research projects. In the third subsection, 3.3, the topic of design theory and its importance in the context of DSR is described. Furthermore, it is explained that an ISDT is being developed and why this decision was made this way. In the last subsection, 3.4, all qualitative and quantitative research methods are presented again in tabular form. The methods are briefly explained in a short description and linked to the individual research projects in which they were used and how.

Section 4 summarizes the individual research projects in the cumulative dissertation. The individual research projects are briefly described, and the core findings are shown. In addition, based on these different research projects, the scientific findings generated, and the status quo of research on human-centered XAI, the ISDT is being developed as a superordinate scientific contribution to communicating prescriptive design knowledge for human-centered XUIs. The complete eight components of an ISDT, according to Gregor and Jones (2007), are derived, combined, and described. Then, the entire ISDT for the design of human-centered XUIs is summarized again.

Section 5 critically reflects on and discusses the cumulative dissertation, i.e., the individual research projects and the resulting synopsis. The first subsection, 5.1, discusses the different research contributions of the individual research projects and the added value of the ISDT developed in the synopsis for both science and practice. This is then deepened by discussing the theoretical contributions and implications in subsection 5.2 and the practical side in subsection 5.3. The discussion section ends with subsection 5.4, in which the limitations of the individual research projects and the cumulative dissertation are discussed. In addition, the potential for future research arising from the limitations is also described.

Lastly, with Section 6, the synopsis and cumulative dissertation are concluded. A summary of the structure of the cumulative dissertation, including the sections and subsections, is presented in Table 1.

	Section	**Subsection**
Synopsis	1. Introduction	1.1 Research Motivation 1.2 Research Objective and Research Questions 1.3 Structure of the Cumulative Dissertation and Included Articles

	2. Research Background and Knowledge Base	2.1 Relevance of Human-Centered (X)AI 2.2 Human-Centered Design in XAI and XUI Research 2.3 Human-Centered Evaluation in XAI and XUI Research
	3. Research Design	3.1 Design Science Research in Information Systems Research 3.2 Design Science Research Approach 3.3 Further Research Methods Used
	4. Summary and Consolidation of the Research Contributions	4.1 Summary of the Research Contributions of the Individual Articles 4.2 Design Theories in Design Science Research 4.3 Developing an Information Systems Design Theory for Human-Centered XUIs 4.3.1 Defining the Purpose and Scope 4.3.2 Establishing the Constructs 4.3.3 Principles of Form and Function 4.3.4 Artifact Mutability 4.3.5 Testable Propositions 4.3.6 Justificatory Knowledge 4.3.7 Principles of Implementation 4.3.8 Expository Instantiations 4.3.9 The Information Systems Design Theory for Human-Centered XUIs
	5. Discussion	5.1 On the Human-Centered Design of XUIs 5.2 Theoretical Contributions and Implications 5.3 Practical Contributions and Design Implications 5.4 Limitations and Future Research
	6. Conclusion	
	Informed and justified through the individual articles included in the cumulative dissertation supplemented by the status quo of relevant research streams.	

Table 1. Outline of the cumulative dissertation.

The dissertation project presented here was structured cumulatively, resulting in several publications. The six articles relevant to the cumulative dissertation were written between 2020 and 2023 and were published in journals or conferences listed in the VHB JQ3 ranking. Only the last article is currently under review. The first article (A1), the basis for the cumulative dissertation project, was published in the C-ranked journal Information Systems Management (ISM). A second article (A2), a short paper, was published in the proceedings of the International Conference on Information Systems (ICIS), an A-ranked conference. Another paper (A3) was published in the Hawaii International Conference on Systems Science (HICSS) proceedings, a C-ranked conference. The research project presented in the article (A4) was published in the B-ranked journal Information Systems Frontiers (ISF). The penultimate article (A5), part of the cumulative dissertation, was published in the proceedings of the International Conference on Design Science Research in Information Systems and Technology (DESRIST), a C-ranked conference. The last article (A6) was submitted to the C-ranked journal Human-Computer Interaction (HCI) as an individual research contribution and is currently under review. Apart from the individual project in A6, the other research projects were carried out with various interdisciplinary researchers. They came from various universities and research institutions, such as the Freie Universität Berlin, Ruhr-Universität Bochum, the University of Liechtenstein, the University of Vienna, the Karlsruhe Institute of Technology, and the German Research Center for Artificial Intelligence. The publications listed in Table 2 below ultimately represent the results of these collaborations.

The table gives an overview of the individual research projects in the cumulative dissertation. These are six articles, all of which are listed in the table and referenced in further text based on their ID. Thus, the table provides an overview of the unique ID for each item. Furthermore, the title of the research project is listed along with the publication date. In addition, the authors are listed, and the name of the journal or conference is given. Referring to the latter, the VHB JQ3 ranking is also given. In the last column, the credit points that can be counted and achieved for each research credit are shown, as well as their total. This cumulative dissertation's formal requirements regarding the credit points to be achieved (required: 2.0) of the Freie Universität Berlin, School of Business and Economics have been achieved (sum: 2.28). The contribution statements for the papers can be found in Appendix 8.2.

ID	Title and (Year)	Author(s)	Journal/ Conference (VHB JQ3-Ranking)	Credit Points
A1	Explainable Artificial Intelligence: Objectives, Stakeholders, and Future Research Opportunities (2022)	Christian Meske Enrico Bunde Martin Gersch Johannes Schneider	Information Systems Management (ISM; VHB: C)	0.25
A2	Improving Customers' Decision-Making on Blackboxed Multimodal Platforms – A Design Science Approach (2020)	Christian Meske Enrico Bunde Jan Fabian Ehmke	Proceedings of the International Conference on Information Systems (ICIS; VHB: A) [Short Paper]	0.00
A3	Fake or Credible? Towards Designing Services to Support Users' Credibility Assessment of News Content (2022)	Enrico Bunde Niklas Kühl Christian Meske	Proceedings of the Hawaii International Conference on Systems Sciences (HICSS; VHB: C)	0.33
A4	Design Principles for User Interfaces in AI-Based Decision Support Systems: The Case of Explainable Hate Speech Detection (2023)	Christian Meske Enrico Bunde	Information Systems Frontiers (ISF; VHB: B)	0.50
A5	Giving DIAnA more TIME – Guidance for the Design of XAI-Based Medical Decision Support Systems (2023)	Enrico Bunde Daniel Eisenhardt Daniel Sonntag Hans-Jürgen Profitlich Christian Meske	Proceedings of the International Conference on Design Science Research in Information Systems and Technology (DESRIST; VHB: C)	0.20
A6	Conceptualizing and Designing Customization Features for Explanation User Interfaces (2023)	Enrico Bunde	Human-Computer Interaction (HCI; VHB: C; Submitted and Under Review)	1.00
				Sum: 2.28
				Required: 2.00

Table 2. List of publications included in the cumulative dissertation.

This subsection completes the first section. The following second section gives an overview of the research background and the descriptive and prescriptive knowledge base, which were relevant to the cumulative dissertation.

2 Research Background and Knowledge Base

2.1 The Relevance of Human-Centered (X)AI

With the steadily increasing spread of AI in different areas of life (Maslej et al., 2023), the relevance and interest in XAI (Haque et al., 2023; Minh et al., 2022) are also increasing. Because even if AI creates many positive possibilities and potentials, it also brings many challenges (Adadi & Berrada, 2018; Arrieta et al., 2020; Haque et al., 2023). These can be challenges of all kinds, and questions arise regarding the regulation of AI, the power to act, agency, accountability, management of AI, transparency, or explainability (e.g., Gunning et al., 2019; Minh et al., 2022; Rudin, 2019). For example, it was found that a faulty AI in a clinical setting could mislead a spectrum of clinical workers, including experts (Tschandl et al., 2019). Further research has shown how decision support systems can also lead to users falling victim to what is known as automation bias and, for example, no longer critically questioning and reflecting on AI recommendations but rather tending to accept them (Goddard et al., 2012). Other research has also shown that although XAI has positive goals, such as making the user understand the AI decision, it can make users more likely to follow the recommendation, even if the AI is incorrect (van der Waa et al., 2021). By involving relevant stakeholder groups in the development of XUIs and using human-centered methods, XUIs can be designed in such a way that they lead to positive effects for users, such as improving task performance (Leichtmann et al., 2023), making the AI recommendation more comprehensible to users (Gunning et al., 2019), or designing XUIs that are useful and easy to use (Meske & Bunde, 2023).

Assuming that AI should augment humans and not replace them, examining the interaction between humans and AI, or XAI, is essential. In particular, the XUIs that are the focus of the cumulative dissertation represent a vital interaction point for users. Explanations are presented for users, and there are many different options for generating, designing, or offering explanations to the user (e.g., Adadi & Berrada, 2018; Ribeiro et al., 2016; van der Waa et al., 2021). A positive or negative experience with the UI can be an essential reason for users to decide for or against an application or enjoy its usage (Barda et al., 2020; Gunning & Aha, 2019). In the context of XUIs, it is crucial to take a human-centered perspective and involve users in different phases of the design and evaluation of XUIs (Gu et al., 2023; Schneidermann 2020a; 2020b). There are a multitude of explanations and ways they can be presented. For example, explanations can be communicated in text format via dialog (Miller, 2019), relevant image areas are highlighted or marked in color (Sokol & Flach, 2020), a table of the most relevant features (Ribeiro et al., 2016), or explanations based on examples (Leichtmann et al., 2023). All these different explanations lead to different interaction experiences and offer different information content. The perception of the explanation is influenced by several aspects on the user side, such as their algorithm aversion (Dietvorst et al., 2015), preferences concerning the presentation format of the explanations (Langer et al., 2021), the scope of the explanation (Barda et al., 2020), the interactivity (Bunde, 2023), or the ease of use (Meske & Bunde, 2023). Appropriate design elements can positively influence the user's interaction experience (e.g., Gunning et al., 2019; Haque et al., 2023; Minh et al., 2022).

To demonstrate some presentation formats in the context of XAI, some examples of modern XAI methods follow. The publicly available Python libraries were used to reproduce the examples. Figure 2 shows an example of the XAI method LIME (Ribeiro et al., 2016). This is an explanation of a classification task based on tabular data. More specifically, different data characteristics classify fungi as edible or poisonous. The explanation shows that the prediction probability is 1.00 or 100% for the poisonous class, shown on the far left. A bar chart in the middle shows each class's different data attributes, values, and relevance. Apart from one data characteristic (gill-size=broad), all other data characteristics stand for the class poisonous. The data characteristics are displayed again on the far right. This is a local explanation that explains a specific classification task.

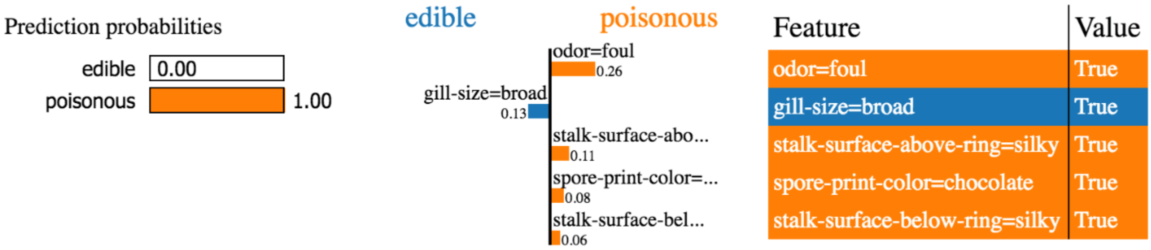


Figure 2. Example of LIME for tabular data (reproduced[2]).

The explanation presented in Figure 3 was generated with the Python library Eli5 and is based on the previously explained XAI method LIME. This is an example of text classification. It was recognized as a class of the 'sciMed' news article with a probability of 0.996, or 99.6%. Furthermore, this explanation can be seen as an established approach by which explanations are designed in text classification. The words are marked in color depending on their relevance and the class to which they belong. In the example, these are the green words for the 'sciMed' class and the red-marked words for other classes. This type of heatmap for texts is established. It can lead to positive effects, as shown in A4 of the cumulative dissertation, for example, that such explanations significantly influence perceived cognitive effort, trustworthiness, perceived informativeness, and mental model (process).

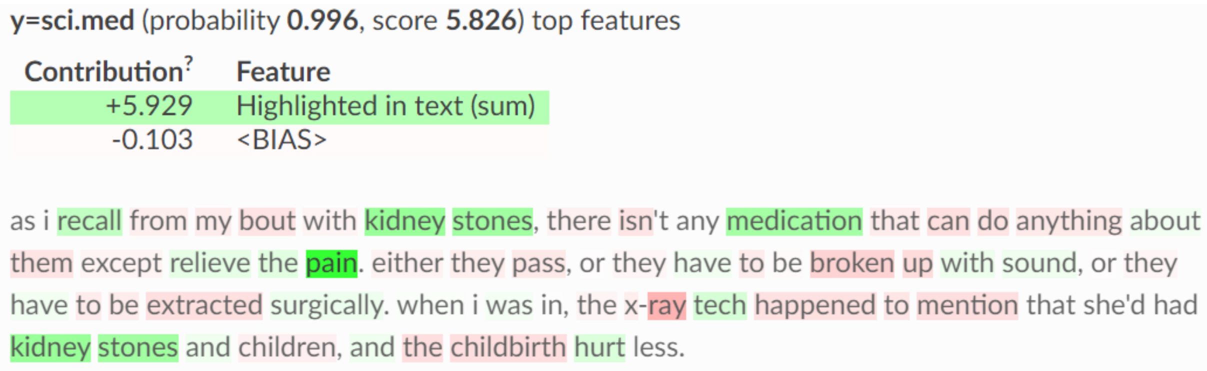


Figure 3. Example of LIME for text data (reproduced[3]).

Figure 4 shows one last example of a text classification explanation with a different structure or presentation format for the different information. This explanation was generated using SHAP and explains a sentiment classification. The different classes are shown above, and the class 'sadness' is active, so an explanation is shown for this class. Below is the sentence to be classified, and the words are also marked in color. The red color represents class 'sadness,' and the word in blue represents another class. The relevance of each data feature, here the words, can be seen in the middle of the chart. The Python library SHAP allows the generation of these explanations in an interactive format.

[2] https://github.com/marcotcr/lime/tree/master
[3] https://eli5.readthedocs.io/en/latest/tutorials/black-box-text-classifiers.html

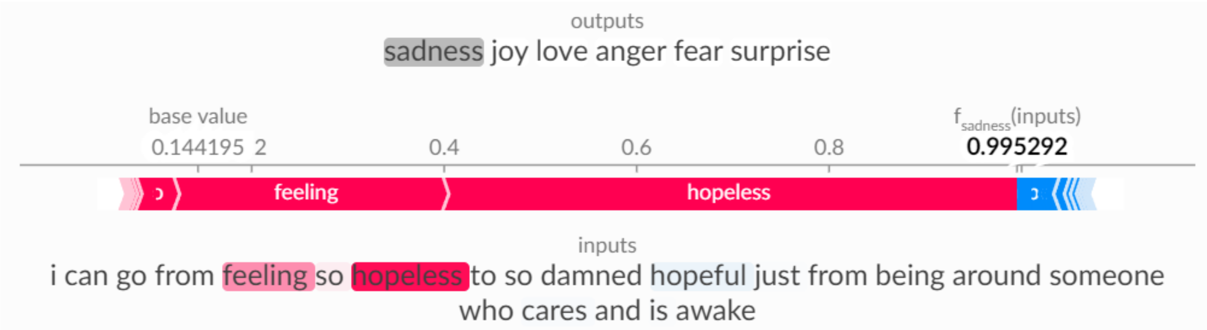


Figure 4. Example of SHAP for text data (reproduced[4]).

In the following Figure 5, there is an example explanation for a computer vision classification task. A convolutional neural network classifies animals in images. On the left picture is the explanation for the class 'dog', and on the right is the explanation for the class 'cat'. The explanation was generated using GradCAM. GradCAM stands for Gradient Class Activation Mapping and is a method for visualizing activation areas in deep neural networks. It allows for identifying the most critical regions of an image that contribute to the classification by the network. Backpropagating the classification gradient calculates a weighting coefficient for each activation map. These coefficients are then used to create a heatmap-like visualization highlighting the regions of interest in the input image. More interpretable than traditional methods, GradCAM provides insight into the decision-making of deep learning models, making it valuable for tasks such as medical diagnosis and image understanding. The Python library Elie5 was used to generate the GradCAM examples.

Figure 5. Example of Grad-CAM for image data - explanation for the class dog (left) and for the cat (right) (reproduced[5]).

A final example shows explanations for another computer vision classification task in Figure 6. One can see the two input images and the explanations for the two most likely classifications. The image caption represents the explanation that stands for the most probable class. Highly relevant regions are highlighted. The explanation was generated using the 'GradientExplainer' of the SHAP library. The 'GradientExplainer' is a SHAP library technique for interpreting machine learning algorithms. It is based on the Shapley value concept from game theory and allows the contributions of individual features to the prediction of a model to be quantified. The 'GradientExplainer' uses gradient calculations to

---

[4] https://shap.readthedocs.io/en/latest/example_notebooks/text_examples/sentiment_analysis/Emotion%20classification%20multiclass%20example.html

[5] https://eli5.readthedocs.io/en/latest/tutorials/keras-image-classifiers.html

determine the change in the model's prediction for each feature. By integrating the gradients over all combinations of features, the Shapley values are calculated, representing each feature's average importance for the prediction. This allows for a transparent and reliable explanation of the model decisions and makes it easier to understand and verify machine learning algorithms.

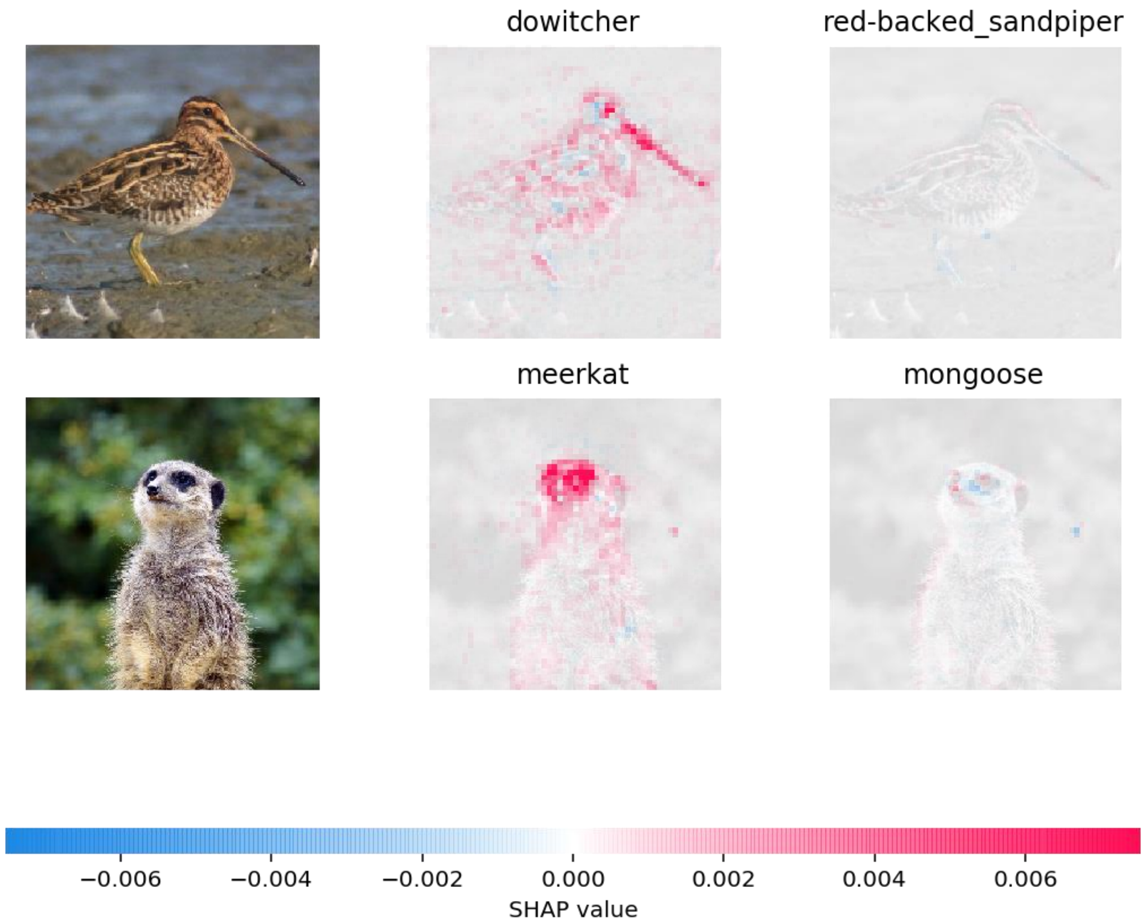


Figure 6. Example for SHAP and GradientExplainer on image data to visualize the 7th layer of a VGG16 (reproduced[6]).

Such XAI methods have been used and studied in versatile and highly relevant application domains like health (Barda et al., 2020), finance (Strich et al., 2021), and manufacturing (Senoner et al., 2022). Moreover, such explanations are beneficial for users and lead to versatile positive outcomes, including increased acceptance (Kocielnik et al., 2019), efficient debugging of machine learning models (Kulesza et al., 2015), or compensating for technological errors (Jussupow et al., 2021). Researchers, mainly from computer science, are constantly introducing novel XAI methods, and a wide range of challenges also accompany this. For example, explanations can be designed in different presentation formats, including charts, tables, highlighting regions of images in image classification tasks, highlighting words in text classification tasks, explanations in natural language, or as dialogue (Barda et al., 2020; Gunning et al., 2019; Miller, 2019; Sokol & Flach, 2020; van der Waa et al., 2021). The presentation format alone can influence the explanations' perception and effectiveness or usefulness (Minh et al., 2022; Naiseh et al., 2023). Another aspect that influences the design of explanations is the targeted audience since different stakeholders may be interested in XAI for varying reasons (Ali et al., 2023). Machine learning

[6] https://shap.readthedocs.io/en/latest/example_notebooks/image_examples/image_classification/Explain%20an%20Intermediate%20Layer%20of%20VGG16%20on%20ImageNet.html

engineers might use XAI to debug AI models or optimize performance (Kulesza et al., 2015). Users can be interested in XAI to understand why AI models came to a specific output and whether they can trust this output (Naiseh et al., 2023). Since AI is used in various application domains with different levels of criticality, versatile stakeholders can be interested in XAI, including regulators, companies, or humans affected by AI models' output (Langer et al., 2021). Consequently, when designing explanations, versatile individual expectations, needs, and experiences should be considered (Meske et al., 2022). Due to all these nuanced aspects that can influence the design of explanations, no one-size-fits-all approach exists (Sokol & Flach, 2020). Research on XAI has recognized this need for a more human-centered perspective with a focus on the individual needs of stakeholders for the development and evaluation of XAI (Langer et al., 2021).

Despite the ongoing research efforts from many different research disciplines, the field of XAI is characterized by a lack of human-centered evaluation studies (van der Waa et al., 2021; Wells, & Bednarz, 2021). This is highly problematic since humans have to work with XAI systems, and how they define, select, or evaluate explanations is influenced by individual-employed cognitive biases and expectations of the explanation process (Miller, 2019). Consequently, recent research has focused on the stakeholders interested in XAI, such as regulators, developers, users, managers, or individuals affected by AI-based decisions (Meske et al., 2022). The design's objectives also depend on the goals and aims of the involved stakeholder groups. For example, domain experts may want to trust the model and gain scientific knowledge, or managers may want to assess regulatory compliance (Arrieta et al., 2020). What further emphasizes the relevance of a human-centered approach in developing and evaluating XAI systems is the large number of characteristics, notions, and goals pursued by XAI that can all influence the perception of XAI systems (Arrieta et al., 2020; Minh et al., 2022; Vilone & Longo, 2021). Moreover, a human-centered approach is essential and valuable since it could take the role of XAI beyond explaining a particular AI system and support users in establishing appropriate trust (Gunning et al., 2019). Consequently, a human-centered approach to XAI fits well into the cumulative dissertation with its DSR approach since it focuses on "[…] uncovering what, when, and how to explain to human end users, by iteratively involving the users in the development process (e.g., through interviews, hypothetical scenarios, focus groups, and questionnaires)." (Schoonderwoerd et al., 2021, p. 2).

Such a human-centered approach to XAI can lead to XAI systems that consider users' individual needs and improve usability (Sovrano & Vitali, 2022). It is also a beneficial approach to designing XAI systems to support everyday lay users in interpreting their explanations and outputs (Fiok et al., 2022). When talking about human-centered XAI, it is essential to note that this perspective understands intelligent systems as a part of a more extensive system, which consists of different human stakeholders, and considers social responsibility aspects, including fairness, accountability, transparency, or explainability (Riedl, 2019). Following a human-centered approach when designing (X)AI systems aims to achieve a high level of human control and automation to increase human performance, understand when complete human control or full automation is necessary, and avoid potential dangers of excessive human control or automation (Schneiderman, 2020b). Similar approaches like human-in-the-loop learning are well-established in XAI, where this approach aims, for example, to enable users to interact with explanations, provide feedback, and ultimately optimize performance (Gunning et al., 2019). However, this approach puts AI at the center of attention, and human-centeredness puts, on the contrary, humans at the center of attention (Schneiderman, 2020a; Schneiderman, 2020b). Figure 7 illustrates this shift. Schneiderman (2020a) describes this shift in human-centered AI as the second Copernican Revolution, and I argue that this also holds for a human-centered perspective on XAI. Ideally, human-centered approaches in designing (X)AI reframe the traditional and technology-centric approach to an approach emphasizing the relevance of involved stakeholders and humans (Schoenherr et al., 2023).

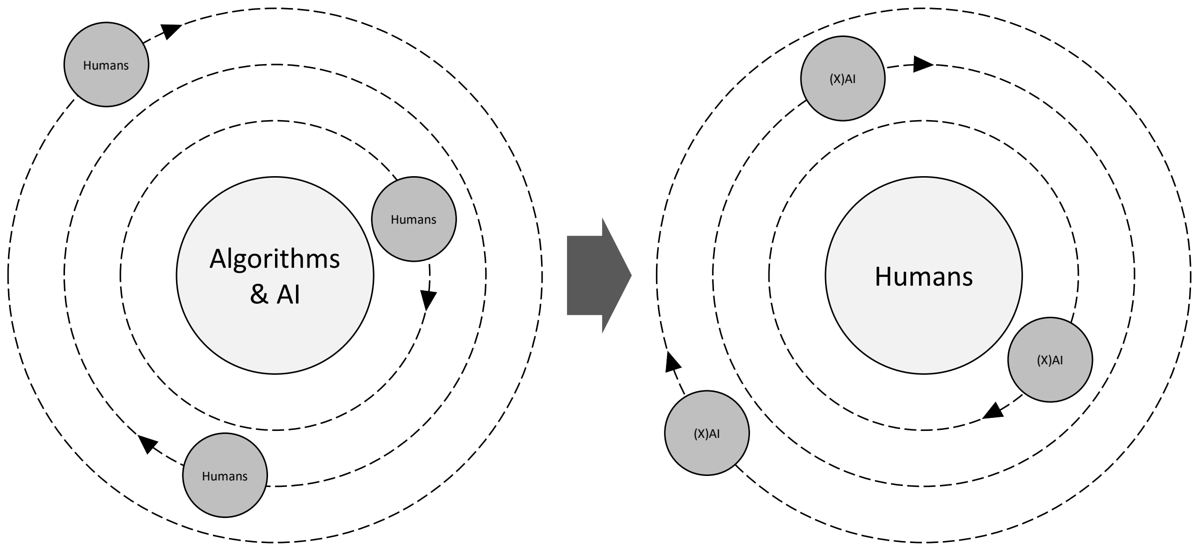


Figure 7. The second Copernican Revolution shifted from algorithms and AI to humans being the center of attention (adapted from Schneiderman, 2020a, p. 112).

This subsection aimed to highlight the basics of XAI and the importance of human-centeredness in this context. Due to the diverse XAI design configurations and their effects on the interaction experience for individual users, it becomes more relevant to involve humans in the design and evaluation process in a human-centered manner to satisfy their needs adequately. The following subsection presents the basics of human-centered design and its relevance for the design of XAI and XUI.

2.2 Human-Centered Design in XAI and XUI Research

The human-centered approach to developing information systems has a long history in the research fields of HCI and ISR (Gasson, 2003). In the discipline of ISR, the interest in human-centered approaches started to grow in the mid-1990s and achieved notable momentum in the 2000s (Ivari & Ivari, 2011). The added value of human-centered information systems is manifold. For example, they can benefit end users and contribute to the success of organizations and businesses (Zhang et al., 2005). On an individual level, human-centered design can also focus on the well-being of users and positively influence the interaction experience with digital artifacts by making them more meaningful, purposeful, and sustainable (Shen et al., 2022). From the designers' perspective, human-centeredness requires empathy so that they can understand and interpret not only the problems that the users may face but also their perspectives when using the system to be designed (Barrett et al., 2015). This requirement is a substantial strength of human-centeredness since this approach seeks to involve the relevant stakeholders in the development process and considers individual requirements (Maguire, 2001). At the same time, a human-centered approach can lead to designed XAI systems that enhance human intelligence instead of replacing humans (Li & Gu, 2023). Moreover, human-centered design is an approach to "[…] communicate, interact, empathize and stimulate the people involved, obtaining an understanding of their needs, desires, and experiences which often transcends that which the people themselves actually realize." (Giacomin, 2014, p. 610).

However, despite being an established and valuable approach, human-centeredness needs to be addressed more in current research focusing on designing XAI systems and XUIs (Nazar et al., 2021). Following a human-centered approach to designing (X)AI systems, goals such as supporting human self-efficacy, encouraging creativity, clarifying responsibility, and facilitating social participation are pursued (Schneiderman, 2020a). Moreover, focusing on human-centered design could increase potential benefits for users and society in different areas, including business, education, healthcare, environmental preservation, and community safety (Schneiderman, 2020c). Here, one of the grand challenges of the human-centered approach lies in generating comprehensive design knowledge and knowledge regarding the evaluation and governance of AI systems (Garibay et al., 2023). It is also essential to consider the potential pitfall of defining relevant stakeholders as too narrow of groups, which could, in the worst case, lead to neglecting further potential stakeholders and their individual needs (Norman, 2005).

However, I aim to avoid such pitfalls by having human-centeredness in mind and following the DSR approach to develop generalizable design knowledge (vom Brocke et al., 2020). Consequently, the human-centered design perspective is operationalized to incorporate the perspectives of relevant stakeholders to design usable XUIs (Maguire, 2001). Figure 8 provides an adapted overview of relevant aspects of human-centered design from the perspective of HCI, which also illustrates intersections with DSR. Therefore, the overarching goal is to improve the visual design of UIs, focusing on the interaction of users and cognitive abilities (Nazar et al., 2021). Figure 8 shows how there are also relevant and established phases in HCD that can be run through, including analysis, requirement gathering, design, and evaluation. These phases overlap with cyclic DSR frameworks, such as design, development, and evaluation, such as Peffers et al. (2007). Thus, the HCD perspective can be well integrated into a DSR project. By involving relevant stakeholders, their experiences and expectations can flow into the design process in different phases.

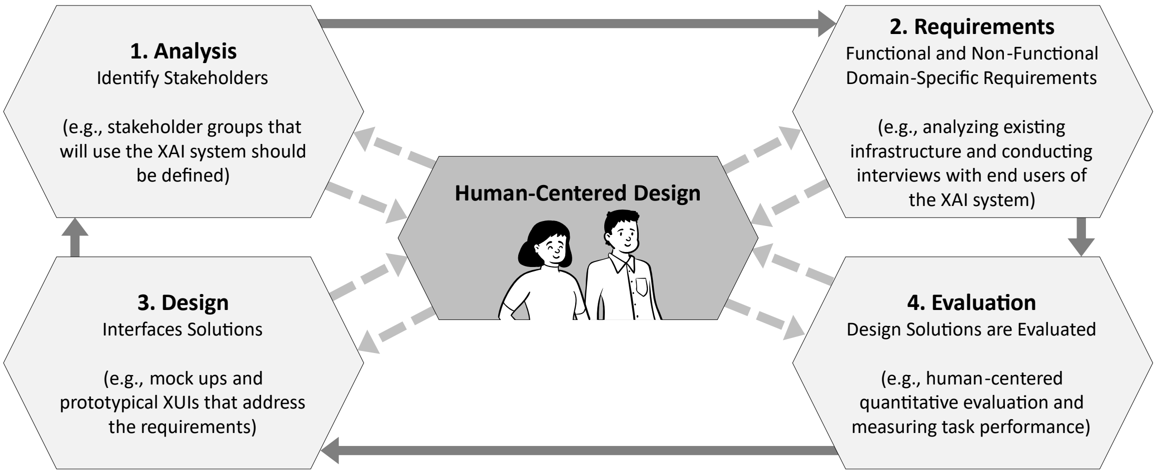


Figure 8. Human-centered design in the field of HCI (based on Nazar et al., 2021, p. 153321).

The design of XUIs is highly relevant because it is a critical component for any information system's success and especially involves unique design challenges for XAI systems (Dudley & Kristensson, 2018). One of the fundamental challenges lies in the presentation of explanations in a comprehensible manner since they come in different formats and styles, including explanations presented in text form, by visual means, or by using examples to explain the rationale for the output of AI models (Arrieta et al., 2020). A human-centered design that considers the individual needs and expectations of UIs can foster a feasible interaction experience for users (Bradley et al., 2022). Likewise, a human-centered design approach can lead to systems that explain their decisions and foster human comprehension (Leichtmann et al., 2023). The core information presented in XUIs is usually about the AI model and integrating one or more XAI methods to provide meaningful information to the user (Chromik & Butz, 2021).

However, many versatile functionalities and design features can be incorporated in XUIs, including features to control the AI decision-making process through detection threshold adjustments, meaningful icons to communicate information concisely, comparisons with other classes in the case of a classification task, or contextual information for the current task (Kocielnik et al., 2019; Leichtmann et al., 2023; Meske & Bunde, 2023; Schoonderwoerd et al., 2021). Since the evaluation criteria can differ significantly depending on the application domain, use case, and involved stakeholders, the human-centered approach enables designers to consider individual human factors when developing XAI systems as well as XUIs (Arrieta et al., 2020; Langer et al., 2021; Schoonderwoerd et al., 2021). Well-designed XUIs can aid users in understanding the black box AI system, and it affects outcomes like the ability of users to establish appropriate levels of trust or debug an underlying AI model (Adadi & Berrada, 2018; Haque et al., 2023). Although XUIs can reduce the information processing effort, the preferred design options within XUIs can differ significantly depending on the stakeholder group (Barda et al., 2020). Therefore, the human-centered approach for designing XUIs is favorable since skipping the groundwork regarding the necessity of explainability features or the necessary design options may result in a misspent effort (Chen et al., 2022). Combining the interdisciplinary knowledge of fields such as HCI or social sciences and incorporating them into the human-centered design of XUIs, more powerful and usable systems can result in a broad range of areas and problems (Chen et al., 2018). The future success of modern AI systems could also greatly benefit from a human-centered approach, primarily when interdisciplinary professionals work together to test and optimize state-of-the-art (X)AI systems, ultimately leading to avoiding extreme algorithmic bias (Xu, 2019).

This subsection dealt with human-centered design and its importance in the ISR community and for XAI and XUI research. The relevance of involving relevant stakeholder groups during the design is a desirable approach, as shown here. The following and last subsection presents human-centeredness in evaluating XAI and XUI research.

2.3 Human-Centered Evaluation in XAI and XUI Research

The influence of explanations for support in decision-making processes can be diverse (Haque et al., 2023). For example, it can be examined to what extent explanations can explain an AI-supported recommendation so that users can understand the system and its output (van der Waa et al., 2021) or build trust (Ali et al., 2023). Different XAI methods can be compared, and differences in user perception can be examined (van der Waa et al., 2021). Different interaction experiences can be explored by instantiating different XAI design configurations (Gunning et al., 2019; Gunning & Aha, 2019). As explained above, the demands and expectations of explanations and the explanation process can be very subjective (Miller, 2019). This supports the two relevant aspects of the human-centered evaluation of XAI and XUIs. On the one hand, feedback regarding optimization potential can be obtained to improve the artifact (Nazar et al., 2021). On the other hand, it can be examined how people perceive XAI systems or XUIs in different contexts, such as in the field (Senoner et al., 2022) and in laboratory experiments or online experiments (Bauer et al., 2023; Leichtmann et al., 2023). Human-centered evaluations fit perfectly into the evaluation phases of DSR frameworks such as Peffers et al. (2007), Kuechler and Vaishnavi (2012), or to evaluate the reusability of design principles (Ivari et al., 2021). On a higher level of abstraction, the human-centered evaluation of the developed XUIs also makes it possible to ensure that the problems previously identified in the DSR process, or the design requirements developed have been satisfied (Hevner et al., 2004; Gregor & Hevner, 2013; Nazar et al., 2021 Xu, 2019). Thus, the involvement of people and, in particular, a diverse sample of the targeted user base is beneficial for the development of XUIs and was successfully implemented in the individual DSR projects included in this cumulative dissertation.

Despite the active research on XAI, the relevance, limitations, and lack of user evaluations for state-of-the-art XAI are emphasized throughout the XAI literature (e.g., Anjomshoae et al., 2019; Dosilovic et al., 2018; Gilpin et al., 2018; Haque et al., 2023; van der Waa et al., 2021; Wang et al., 2019; Wells & Bednarz, 2021). Two overarching categories of evaluations in XAI can be differentiated: the objective evaluation, which is rather formal, objective, or technical, and the human-centered evaluation, which is often more subjective and organized as a qualitative or quantitative evaluation study (Vilone & Longo, 2021). A comprehensive literature review has shown a bias toward algorithm-centered evaluation, further underscoring the lack of human-centered evaluations and the potential insights they could generate (Sperrle et al., 2021). There are many different reasons why human-centered evaluations of XAI and XUIs are so relevant. For example, versatile stakeholder groups are interested in AI-generated explanations and have varying information needs, prior knowledge, or backgrounds (Arrieta et al., 2020; Meske et al., 2022). Aggravating the situation is that human preferences are not always rational, and the psychology of humans is not entirely measurable (Selbst et al., 2019; Stark, 2018). Human-centered evaluations can also shed light on subjects like the acceptance or satisfaction of imperfect AI recommendations, where XUIs are essential to consider (Kocielnik et al., 2019). In addition, the need for explanations and their presentation format may differ in various application scenarios depending on the criticality of the AI use, its possible consequences, and the used (X)AI methods (Adadi & Berrada, 2018; Vilone & Longo, 2021). For example, an XUI for medical domain experts sometimes requires different UI elements than an XUI for other stakeholder groups, such as other users or people affected by the decision, to enable a positive interaction experience or to satisfy individual needs and expectations (Barda et al., 2020; Bunde et al., 2023).

A human-centered approach to evaluating XAI systems is emerging to investigate how far individual demands are met and satisfied (Barda et al., 2020; Schoonderwoerd et al., 2021; Vilone & Longo, 2021). Human-centered evaluations complement objective evaluations like measuring the accuracy of AI models, and the contextual needs of the users and affected stakeholders can be better understood, ultimately leading to an improved experience with deployed XAI systems (Beede et al., 2020). Moreover, the opportunity arises to leverage human intelligence as an essential feedback mechanism, and it can empower humans to test and understand the XAI systems interactively (Cai et al., 2019). Human-centered evaluations are particularly relevant for XUIs deployed in real-world settings (Guidotti, 2021). However, human-centered evaluations are complex since many different human and technical factors can be considered and could potentially influence the interaction experience of users with XUIs (Sperrle et al., 2021). Therefore, the evaluation is one of the grand challenges of building human-centered XAI systems (Garibay et al., 2023). For a human-centered evaluation of XAI systems, the XUI is an essential component since users with tasks interact with the XUI (Gunning & Aha, 2019). Here, it is crucial to evaluate the final XUI design and, more critically, XAI and stakeholder perceptions during all phases of the human-centered design process (Eshan et al., 2022). In this way, it can be ensured from a human-centered but also DSR perspective that, for example, the problem under investigation has been correctly understood and defined or that the derived design requirements correspond to the user's ideas (Nazar et al., 2021). Since the users of AI systems are the primary stakeholders who adopt, use, or must work with the outputs of such systems, a human-centered approach gains relevance (Yuan et al., 2023).

In current research on human-centered evaluation of XAI and XUIs, there are projects such as field studies and, much more often, surveys, laboratory experiments, or online experiments in a wide variety of research disciplines (Arrieta et al., 2020; Vilone & Longo, 2021). These include, for example, the domains of commerce (Zimmermann et al., 2023), health (Barda et al., 2020), or health (Beede et al., 2020). Depending on the research discipline, human-centered evaluations reach a very different depth (Nazar et al., 2021). In addition, very different aspects are examined, such as the influence of specific XAI design configurations (Kocielnik et al., 2019) or the comparison between different explanation types (van der Waa et al., 2021). Concerning the degree of maturity of the object of evaluation in the context of a human-centered XAI evaluation, the following can be stated: Research projects, which often come from computer science and introduce new types of XAI methods, often use human-centered evaluations to evaluate their XAI methods and to achieve added value compared to the status quo. These are often very controlled experiments and surveys. Human-centric evaluations that integrate interactive XUIs go one step further. These XUIs are usually evaluated in a realistic application context with relevant stakeholders. For example, this can be the case in the health sector, such as diabetes self-management (van der Waa et al., 2021). XUIs can appear very realistic due to interactive elements and the possibility of manipulating data and are, therefore, often used for surveys or online experiments to investigate the perception of XUIs and interaction between the user and the XUI (Gunning & Aha, 2019; Kulesza et al., 2015). For example, interactive XUIs can be used to compare different design configurations or explanation types (van der Waa et al., 2021). The XAI systems are the artifacts with the highest level of maturity, which could be developed, for example, as part of a DSR project. Such systems are often designed and developed human-centered in a specific application context (Yuan et al., 2023). This includes, for example, the health sector, in which, for example, an XAI system was developed, and both the design and evaluation were carried out in a human-centered manner (Barda et al., 2020). In this way, the relevant individual requirements of the complex user base could be identified and satisfied. Furthermore, a more in-depth investigation was possible into how such an XAI system can change or influence the everyday clinical work of the stakeholders involved. The human-centered approach will gain relevance, as many different dimensions can be examined in a human-centered manner, such as calibrated trust in (X)AI (Ali et al., 2023), the change in work environments and processes (Fink et al., 2021), or comparison of continuously newly developed XAI methods or explanations (Haque et al., 2023).

The most diverse aspects regarding the perception of or interaction with XUIs can be examined. These include, for example, how different design configurations lead to an XAI system being perceived by users as useful and trustworthy (Minh et al., 2022; Vilone & Longo, 2021), how different types of explanations influence the understanding of users (van der Waa et al., 2021), or what effects explanations can have in real-world settings (Senoner et al., 2022). In the same way, however, it should also be examined whether people could recognize manipulated and thus incorrect explanations or AI advice (Kocielnik et al., 2019) or whether explanations might lead to users being more likely to believe the system, even if the AI-based system generates incorrect recommendations (van der Waa et al., 2021). Consequently, there is an almost inexhaustible potential for investigating users' perceptions and interactions with XUIs. It examines how explanations have to be designed to lead to user satisfaction, what role interactive elements have in XUIs, how explanations influence mental models, or how people cognitively process different explanations (Ali et al., 2023; Angelov et al., 2021; Haque et al., 2023; Miller, 2019). To be able to examine these and many other aspects, relevant stakeholder groups must be involved. This can be achieved through human-centered evaluations, and the status quo of XAI research already offers some insights in this context.

Figure 9 shows what a human-centered approach to evaluating XAI systems, or XUIs, can look like. The relevant users interact with the artifact and usually perform a task, such as a classification task. During this part of the experiment, data can already be collected, such as user decisions and their performance during the task to be solved, for example, by calculating accuracy. Furthermore, interactions with the XUI can be monitored and analyzed, for example, to determine whether certain functions are used or how often. Subsequently, qualitative interviews can be conducted to discuss the system's performance and interaction experience and determine the potential for optimization. Alternatively, quantitative methods can be used, for example, to conduct a survey and measure psychological constructs such as trust or perceived usefulness, mainly using Likert scales. There is also the possibility of combining the qualitative and quantitative approaches.

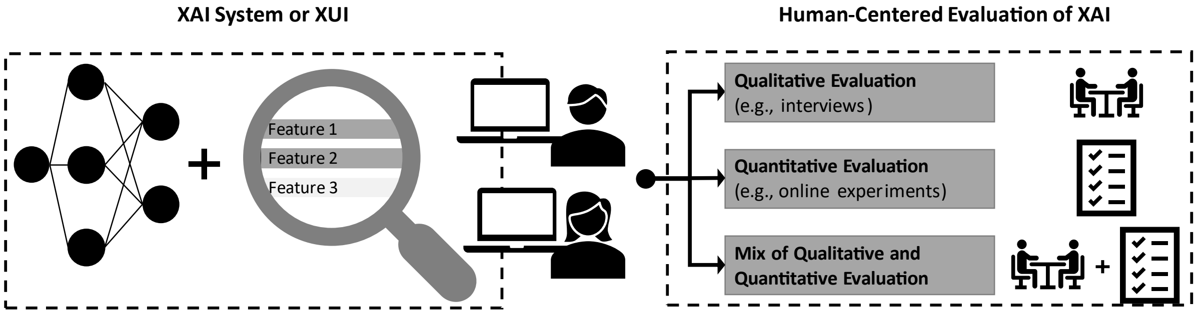


Figure 9. Overarching approach in human-centered evaluation (Vilone & Longo, 2021).

Table 3 below presents an overview of dimensions examined in human-centered evaluations. Exemplary studies were identified for XAI methods, XAI systems, and XUIs. Therefore, the overview is representative and attempts to typify a body of research on human-centered XAI evaluations (vom Brocke et al., 2015). The overview also aims to show active research in the context of human-centered evaluations, which can still be substantially expanded. Some dimensions, such as task performance, understandability, or trust, are frequently measured and examined. On the other hand, numerous dimensions are rarely examined, such as acceptance, actionability, or controllability. Moreover, many other dimensions can become relevant depending on the focus of the investigation and the context. For example, Ribeiro et al. (2018) introduced a new XAI method and evaluated it using a human-centered approach. They found that anchors enable users to predict the behavior of an AI model with less effort and high precision. In another research project, Leichtmann et al. (2023) developed an XAI system to help mushroom pickers distinguish between edible and poisonous mushrooms. Their human-centered evaluation found that users provided with explanations outperformed those without access to explanations.

Moreover, explanations have led to better-calibrated trust levels. With a focus on XUIs, van der Waa et al. (2021) conducted a human-centered evaluation as a quantitative experiment. Test subjects with diabetes were involved and used an XUI to plan their food intake. Different explanations were compared, and one particular finding was that subjects followed recommendations with explanations more often, even if they were incorrect. However, since there are a significant number of characteristics and notions of XAI, this overview emphasizes the need for more human-centered evaluations and research, which is highly important to advance the field of XAI (Sperrle et al., 2021; Vilone & Longo, 2021; Wells & Bednarz, 2021). Thus, the relevance of further human-centered evaluations and experiments should be emphasized once again since these findings can generate valuable insights for research and practice.

The tabular summary follows on the next page, which contains relevant evaluation dimensions of XAI that were evaluated in the associated studies in an essentially human-centered manner. This table ends this subsection. After the developed and relevant knowledge base, presented in Section 2, Section 3 describes the research design of the individual research projects that are part of the cumulative dissertation and the further research methods used.

Type of Artifact	Evaluation Dimensions of XAI: Acceptance	Actionability	Behavioral Intention	Causality	Cognitive Effort or Effort	Cognitive Load	Cognitive Relief	Comprehensibility	Confidence	Controllability	Customization	Ease of Use	Effectiveness	Information Amount	Informativeness	Interactivity	Interestingness	Internality	Interpretability	Learning	Mental Model or Mental Fit	Persuasion	Quality	Reaction Time	Satisfaction	Simplicity or Simplification	Stability	Task Performance	Trust or Trustworthiness	Understanding or Understandability	Usability or Usefulness	Reference
XAI Method					●																●							●				Ribeiro et al. (2018)
						●			●																			●				Kaur et al. (2020)
									●															●				●				Huysmans et al. (2011)
																			●													Bau et al. (2017)
		●		●			●	●							●	●	●					●				●				●		Vilone and Longo (2022)
																									●							Weitz et al. (2021)
XAI System																							●					●				Senoner et al. (2022)
																								●	●					●		Knapic et al. (2021)
																															●	Spinner et al. (2020)

																											●		●		Kulesza et al. (2015)
																			●								●	●			Leichtmann et al. (2023)
								●			●	●												●				●	●	●	Zacharias et al. (2022)
									●								●									●					Ha et al. (2022)
	●		●																												Calisto et al. (2021)
								●																●					●		Dikmen and Burns (2020)
													●						●		●						●		●		van der Waa et al. (2021)
				●	●						●			●						●								●		●	Meske and Bunde (2023)
								●															●				●				Alufaisan et al. (2021)
																													●		Kim et al. (2014)
																														●	Harbers et al. (2010)
																											●	●	●		Lim et al. (2009)
			●								●																	●		●	Bunde (2021)
XUI										●				●	●	●								●			●				Bunde (2023)
Σ	1	1	2	2	2	1	1	1	5	1	1	3	1	1	3	2	2	1	1	2	2	2	1	3	5	1	1	10	5	8	5	23

Table 3. Human-centered design and evaluation studies focus on XAI methods, systems, and XUIs.

3 Research Design

3.1 Design Science Research in Information Systems Research

This cumulative dissertation follows a design-oriented approach with a long European tradition (Winter, 2008). The DSR approach is the dominant ISR paradigm in German-speaking countries (Österle et al., 2011). However, DSR is also considered a fundamental research approach in ISR internationally and has become well-accepted (March & Storey, 2008; Peffers et al., 2018). While the ISR discipline constantly evolves, DSR has gained momentum recently (Jeyaraj & Zadeh, 2020). In addition, DSR has gained popularity with doctoral students, who use DSR to create knowledge by designing novel or innovative artifacts and investigating their use or performance (Cater-Steel et al., 2019). DSR is fundamentally a problem-solving paradigm rooted in engineering and the sciences of the artificial (Hevner et al., 2004; Simon, 1996). Through DSR projects, highly relevant contributions are represented through design knowledge that is valuable for research and practice (Gregor & Jones, 2007). This makes DSR an excellent research paradigm to address challenges in the context of digital innovation and on a societal level (Becker et al., 2015; vom Brocke & Maedche, 2019; Rai, 2017).

Design knowledge contains information and insights about the means-end relationship between the problem and solution space (vom Brocke et al., 2020; Venable, 2006). Consequently, design knowledge conveys knowledge about innovative design solutions to real-world problems (vom Brocke & Maedche, 2019; vom Brocke et al., 2020). Design principles are one representation form used to specify design knowledge in an accessible manner (Gregor et al., 2020). Another kind of design knowledge can be reflected through instantiated artifacts such as software products (Gregor & Hevner, 2013). Design theories are another desirable representation of design knowledge, summarizing the essential findings and learnings of DSR projects (Gregor & Hevner, 2013; Gregor & Jones, 2007). There is a broad discussion on what constitutes an appropriate scientific contribution of DSR, where one perspective is that they form a continuum with at least dimensions: "[…] from very novel artifacts to rigorous theory development and form early visions of technology impact to studies of technology impact on users, organizations, and society." (Baskerville et al., 2018, p. 369). By following the DSR approach, further ISR methods, and established guidelines for the execution of high-quality DSR projects, I aim to contribute prescriptive, purposeful, relevant, and valuable knowledge for the design of XUIs and empirical insights about the Human-XAI Interaction (Baskerville & Pries-Heje, 2019; Hevner, 2007).

The following Figure 10 provides a high-level summarization of the articles included in the cumulative dissertation. The articles are positioned within the Information Systems Research Framework for DSR, which supports "[…] understanding, executing, and evaluating IS research combining behavioral-science and design-science paradigms." (Hevner et al., 2004, p. 79). The foundation for the cumulative dissertation was established in research project A1. In this article, a literature analysis on XAI and explainability was carried out specifically in the ISR discipline. In doing so, the overarching objectives of XAI were conceptualized, various stakeholder groups were described, and quality features for personalized explanations were identified. In addition, behavior-oriented and design-oriented future research possibilities were shown. Consequently, A1 represents the basis of the entire cumulative dissertation.

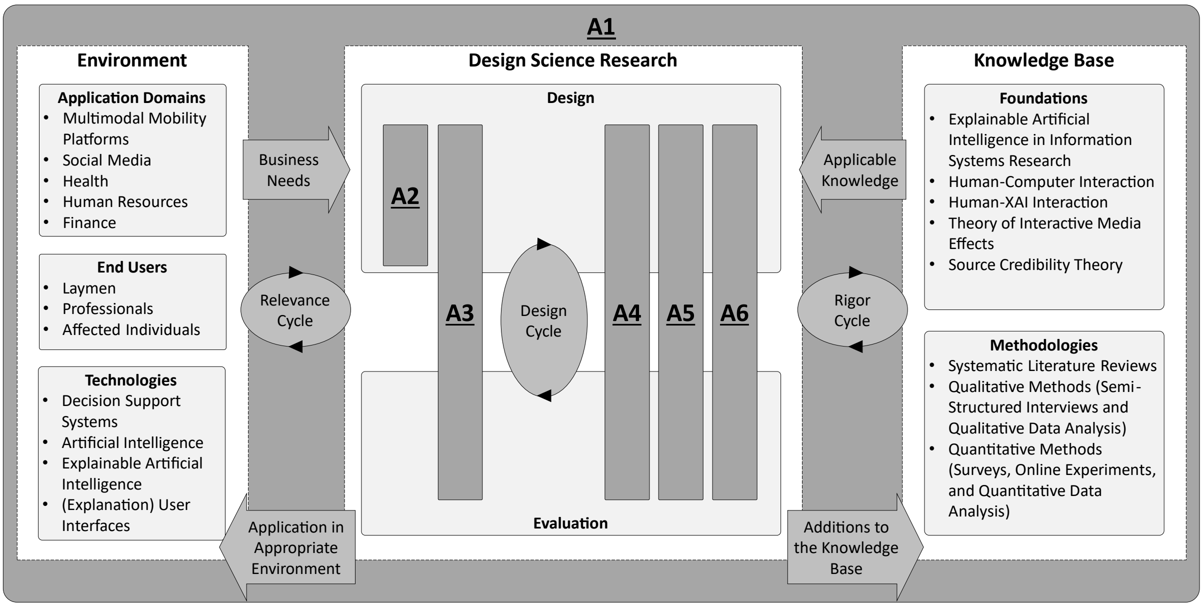


Figure 10. Articles of the cumulative dissertation positioned in the Information Systems Research Framework (based on Hevner et al., 2004, p. 80; Hevner, 2007, p. 88).

Overall, the various projects in the cumulative dissertation examined very different application domains. However, the knowledge base varied only slightly, as it primarily focused on the status quo of research, with my focus on the human-centered design and evaluation of XAI systems. However, parts of the project-related knowledge base may differ, such as the application domains and the existing solutions that must be considered. In concrete terms, the DSR project presented in A2 was carried out in the application context of travel planning. Black-box multimodal mobility platforms were examined, as were how their XUIs should be designed in a human-centric manner. Here, the status quo of research on the design of multimodal mobility platforms, technical feasibility from the perspective of operations research, or the topic of XUI design was included. Added to this was the justificatory knowledge of heuristic decision-making, which influenced the developed design principles based on the findings from interviews with mobility experts and regular end users, as well as the knowledge base.

The DSR projects presented in A3 and A4 were carried out in the application domain of social media. A3 was about the design of AI-based services that should support the credibility assessment of online news. Various challenges were identified in the knowledge base concerning presenting relevant information from such a service to end users. Together with the targeted end users, the prototype was qualitatively evaluated in a human-centered manner. Article A4 presents a more extensive DSR project for XAI-based hate speech detection. The goal was to design an XUI that supports professional social media moderators in their work to moderate social platforms. The knowledge base and our design considered existing solutions and design knowledge. A wide range of information and feedback was obtained through many qualitative and quantitative human-centered evaluations. This way, the XUI could be optimized over three design cycles, its usefulness could be proven, and the significant influence of explanations could be measured. Regular end users, professional social media moderators, and software developers were involved in the evaluation episodes.

In the DSR project presented in A5, the design of XUIs for medical decision support systems was investigated. An existing system, which was developed to detect skin lesions, was examined using an archaeological approach. What is unique about this system is that it should be usable by both doctors and patients. Therefore, methods from HCI and usability research were used, and the existing system was analyzed from a human-centered perspective by combining the think-aloud method with semi-structured interviews to understand the system's perception from the user's perspective. The participants included patients, medical students, and medical specialists. Based on the qualitative data analysis, design requirements and principles intended for an optimized system were developed. The knowledge base here was heavily influenced by the Theory of Interactive Media Effects and the topic of HXAII. The design principles were evaluated regarding their reusability with user interface and user experience designers.

In the last article, A6, customization for XUIs was conceptualized, designed, and human-centered evaluated. In doing so, the knowledge base around customization in HCI research, the Theory of Interactive Media Effects, and HXAII research were used to a large extent. In two design cycles, design configurations of customization were operationalized, implemented, and quantitatively evaluated in a human-centered manner. In the first experiment, it was measured to what extent the customization features in XUIs influence the perceived quality of the explanation. In a second design cycle, the interaction with an XUI with customization features was examined more deeply, and a structural equation model was developed to analyze the effects.

In summary, the topics of human-centered design and evaluation of XUIs in different application domains were examined. Different stakeholder groups, such as laymen, professionals, or affected individuals, were involved. The projects were informed by a shared knowledge base, from which they consumed knowledge but also for which they produced and returned design knowledge. However, the knowledge base then varied slightly depending on the application context and domain. All of the knowledge gained from the human-centered investigations into the design and evaluation of XUIs is brought together in this cumulative dissertation in a superordinate design theory, an Information Systems Design Theory (ISDT) based on Gregor and Jones (2007).

In this first subsection of the research design, the use of the DSR paradigm was motivated and justified. Furthermore, the individual articles in the cumulative dissertation were positioned in the ISR framework based on Hevner et al. (2004) and Hevner (2007). As a result, a high-level overview of the individual research projects was presented. On the other hand, the relevant components—environment, DSR, and knowledge base—were broken down for the individual projects. The following and second subsection provides more information about the methodological DSR perspective, how design knowledge was developed.

3.2 Design Science Research Approach

After the last subsection introduced the DSR paradigm as a methodological approach for the cumulative dissertation and the individual research projects were positioned in the ISR framework for DSR by Hevner et al. (2004) and Hevner (2007), this section aims to present in more detail the DSR methods. For this purpose, it should first be explained how design knowledge was generated through the individual research projects and how it was brought together in the context of the cumulative dissertation, resulting in a superordinate design theory, i.e., the ISDT for human-centered XUIs.

As described, the foundation for the cumulative dissertation was built in A1, where the phenomenon of XAI in the discipline of ISR was explored. Established methods for executing rigorous systematic literature reviews were followed, including Webster and Watson (2002) and vom Brocke et al. (2015). A set of XAI-specific objectives, stakeholders, and future research opportunities focused on ISR were derived in the context of this research project. In addition, the rich history of research on explainability in the context of knowledge-based systems, expert systems, and intelligent agents that date back to the 1980s and 1990s was also covered. It was in A1 that further promising potential RQs and research opportunities from behavioral science and DSR perspectives were identified. Consequently, A1 was the primary motivation to use a DSR approach for the cumulative dissertation and the starting point for the conducted research during the cumulative dissertation and all resulting research articles published.

Figure 11 summarizes the articles included in the cumulative dissertation. While A1 built the foundation of the dissertation project, the research projects that followed intensely focused on developing prescriptive design knowledge for XUIs in different application contexts, instantiating the design knowledge, and evaluating it with relevant stakeholders. An exception is A2, a short paper published in the International Conference on Information Systems proceedings, where only prescriptive design knowledge was proposed but not instantiated or evaluated. The remaining articles, A3, A4, A5, and A6, are typical DSR projects where different DSR-specific methods have been used, for example, the ISR Framework for DSR (Hevner et al., 2004), cyclic DSR frameworks like Peffers et al. (2007) or Kuechler and Vaishnavi (2012), evaluation strategies based on Venable et al. (2016), and reusability evaluations for design principles (Ivari et al., 2021). Therefore, the contribution of this cumulative dissertation is in line with the overarching key contributions required in DSR (Baskerville et al., 2018, p. 361): (i) designing novel IT artifacts in an application context with measurable improvements and (ii) establishing prescriptive knowledge to extend and generalize the knowledge contribution of the DSR project.

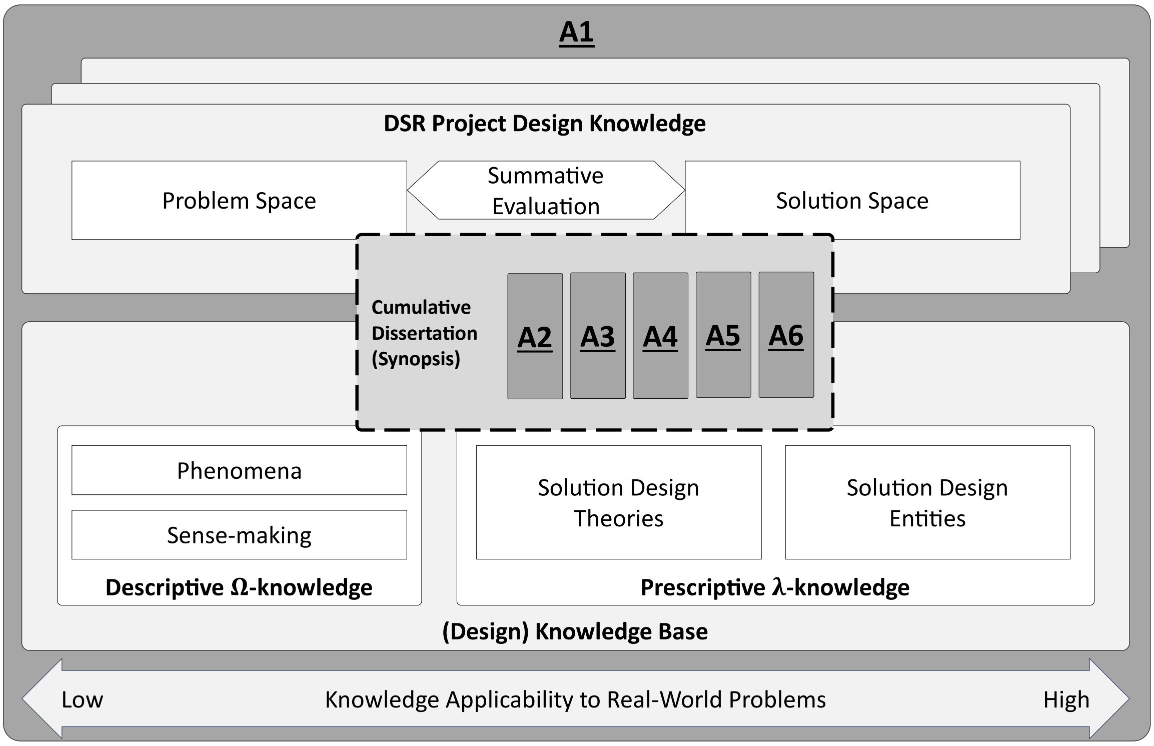


Figure 11. Design knowledge production and consumption within the cumulative dissertation (based on Drechsler & Hevner, 2018; p. 89, Hevner, 2021; p. 74-75).

In the following, the DSR-specific methods and procedures are discussed. Since design knowledge contributions can come in different forms, the status quo of DSR was followed by using well-established frameworks, guidelines, and recommendations. One of the essential DSR frameworks that influenced all DSR projects is the ISR Framework introduced by Hevner et al. (2004). While it underlies all articles, it heavily influenced the research design of A2 and A3. Here, the framework was used to understand, explain, position, and conduct both DSR projects. In A2, relevant stakeholders were involved, including end users and experts. This human-centered approach made it possible to investigate the problem space with a focus on the experiences, needs, and expectations of the stakeholders while investigating the problem space. In A3, a human-centered evaluation of the instantiated prescriptive design knowledge was conducted, and relevant stakeholders were involved. Ultimately, the ISR framework developed by Hevner et al. (2004) was also used to position and classify the different research projects in this cumulative dissertation.

Another crucial methodological building block from the DSR paradigm is established cyclic frameworks or iterative processes. Such cyclic frameworks are widespread within the DSR research of the ISR community and are therefore published in top IS journals (e.g., Diederich et al., 2020; Meth et al., 2015; Morana et al., 2019). Cyclic frameworks of Peffers et al. (2007) and Kuechler and Vaishnavi (2012) have been used in individual research projects. For example, in A4, the cyclic framework of Peffers et al. (2007) was used. The different steps have been completed sequentially and in three consecutive design cycles. In Articles A5 and A6, the Kuechler and Vaishnavi (2012) cyclic framework was followed. In A5, for example, a design cycle was run through, with the evaluation of the design principles being organized in an iterative process that began with qualitative semi-structured interviews and, via optimization based on the knowledge gained, led to a quantitative survey to assess the reusability of the proposed design principles developed in the DSR project. Kuechler and Vaishnavi (2012) followed in A6, with two completed design cycles in this research project.

Design knowledge was constantly developed as part of the various research projects in the cumulative dissertation, with A1 being an exception. In the remaining articles, design principles were developed and in later research, more precisely in A5 and A6, and the cumulative dissertation itself they were formalized according to the scheme for design principles of Gregor et al. (2020). Furthermore, the prescriptive design knowledge for XUIs was instantiated and evaluated in different prototypical XUIs with different maturity levels. Two exceptions should be briefly mentioned here. A2 is an exception because the prescriptive design knowledge was not instantiated due to the nature of short papers. A5 is a second exception because, in this project, an existing artifact was examined and evaluated using a human-centered approach. Then, an extensive set of challenges, requirements, and design principles was derived and developed. Thus, the instantiation of prescriptive design knowledge was not the objective in A5 either. In the remaining articles, A3, A4, and A6, the prescriptive design knowledge was always instantiated and evaluated qualitatively or quantitatively with a human-centered approach.

Another essential part of DSR research and projects is the evaluation phase. The instantiated prescriptive design knowledge resulted in prototypical interactive XUIs, mainly developed as web applications to make them easily accessible to potential users for evaluation. A significant advantage of this approach was, for example, that no installation was necessary to be able to use the instantiated XUI. In articles A3, A4, and A6, the prescriptive design knowledge was instantiated in XUIs. In A3, for example, a design cycle was run through, and an initial human-centered, qualitative evaluation of the prototype XUI was carried out. Three design cycles were run through in A4, whereby the first prototypical XUI became increasingly mature, while the DSR project ran through the three design cycles. Both human-centered qualitative and quantitative research methods were used in the evaluations in A4. In the last article, A6, two design cycles were run through, whereby two XUIs were developed for different task types based on the previously developed prescriptive design knowledge. The XUIs were quantitatively evaluated in both design cycles in a human-centered manner. Based on this summary, it can be seen that the evaluation is a vital part of the DSR research because it should be determined to what extent the innovative solution developed solves the previously identified problem or achieves defined goals. There are also established frameworks in DSR research to make evaluations rigorous and scientific. For example, in articles A3, A4, and A6, the DSR evaluation strategy was based on Venable et al. (2016). The framework of Ivari et al. (2021) can be used in both a human-centered qualitative and quantitative evaluation to evaluate the reusability of the developed design principles, which was done in A4, A5, and A6. In A4, a human-centered quantitative evaluation of the design principles was conducted. In A5, a mix of human-centered qualitative and quantitative methods was used for the evaluation, and in A6, a human-centered qualitative evaluation was conducted.

This subsection presented how the different research contributions, part of the cumulative dissertation, consume and produce knowledge from the overarching knowledge base or the knowledge base, respectively. Furthermore, the individual research projects were briefly summarized, focusing on the DSR methods, which were supplemented with further detailed information. The following and last subsection presents an overview of all the methods used, i.e., all scientific methods used in the individual research projects that are part of the cumulative dissertation.

3.3 Further Research Methods Used

During the execution of the individual DSR projects, several established research methods from ISR and HCI were used. The chosen research methods were used in different stages of the DSR projects. It is necessary to use rigorous methods to construct a design solution and its evaluation to achieve high research rigor (Arnott & Graham, 2012). In addition, different methods can be used to investigate the problem at hand and to develop a firm knowledge base (Hevner et al., 2004; Gregor & Hevner, 2013; Webster & Watson, 2002).

The rigorous evaluation of design solutions is one of the significant tasks in DSR (March & Storey, 2008). It is a crucial aspect that aims to provide evidence for the usefulness of the artifact, and many different criteria can be evaluated, including criteria such as validity, utility, quality, and efficacy (Gregor & Hevner, 2013). Consequently, the evaluation aims to ensure that the design solution meets the identified requirements and solves the real-world problem (Beck et al., 2013; Gregor & Jones, 2007). Many different evaluation methods can be used within DSR projects, including observational, analytical, experimental, testing, or descriptive (Hevner et al., 2004). Accordingly, different research methods were used, depending on the phase of the DSR project and the associated goals. Therefore, this subsection provides an overview of the different research methods used. The human-centered qualitative and quantitative research methods used during the evaluations play a significant role here but are supplemented by other research methods.

The systematic literature review is one of the most frequently used research methods used as the primary research approach for A1 and has built the foundation for individual DSR projects. The research method of systematic literature reviews was also crucial for A3, where challenges and contributions in the context of fake news detection were identified, which were the basis for the subsequently developed prescriptive design knowledge. A systematic search and analysis of relevant literature is critical to establishing a firm foundation for advancing knowledge (Webster & Watson, 2002). It is also one of the most widely applied research methods since all researchers must analyze the literature relevant to their research endeavor (Okoli, 2015). Moreover, systematic literature reviews have been recognized due to their impact on the discipline of ISR (Schryen et al., 2017). However, selecting suitable and relevant literature is a non-trivial task, making it essential to follow established research methods to achieve a rigorous systematic literature review (Wolfswinkel et al., 2013). Established methods were used in the individual DSR projects, mainly vom Brocke et al. (2015), supplemented by Webster and Watson (2002). Systematic literature reviews are also established in DSR projects, often to initiate the project. It is used either as a stand-alone methodology or in combination with qualitative methods such as interviews (e.g., Chanson et al., 2019; Diederich et al., 2020; Meth et al., 2015; Morana et al., 2019; Toreini et al., 2022). Consequently, the systematic literature review was an essential method for A1 and A3 but was also elementary for every research project that is part of the cumulative dissertation.

Another research method of high importance for individual DSR projects stems from the qualitative research field. A broad range of human-centered qualitative research methods exist that can be used within DSR projects, for example, interviews, questionnaire studies, artifact studies, lab-based design studies, or focus group studies (Goldkuhl, 2019). Qualitative research methods such as interviews have been used for a long time as a data collection method in ISR (Benbasat et al., 1987). In the case of DSR, interviews are also well-established and have been used during the early phases of DSR projects to investigate the problem space (e.g., Diederich et al., 2020; Lins et al., 2019; Morana et al., 2019; Tuunanen et al., 2023) or for later phases like the evaluation of instantiated artifacts (e.g., Chanson et al., 2019; Diederich et al., 2020; Morana et al., 2019; Seidel et al., 2015). To achieve high transparency for the conducted semi-structured interviews, data collection information was provided "[…] about where, when, how, and from whom data was collected, and how data was analyzed […]" (Sarker et al., 2013, p. xiii). The semi-structured interviews were organized based on individually developed interview guides that suited the DSR project and involved stakeholders. Moreover, internal applicability checks with colleagues were conducted to ensure that the questions covered the relevant thematical aspects and met the planned time frame (Rosemann & Vessey, 2008). For the recruitment of relevant participants in the semi-structured interviews conducted, the professional network was used with contacts in areas like social media (A3 and A4), health (A5), or software development (A5). Additional participants were identified based on the snowball sampling approach (Patton, 2002). Therefore, the first participant was asked to suggest other individuals with similar knowledge and experience (Bagayogo et al., 2014). The snowball sampling and semi-structured interviews stopped after theoretical and data saturation were reached. For the qualitative data analysis, thematic analysis was used by following the guidelines of Braun and Clarke (2006). Based on the thematic analysis, challenges, design requirements, or potential for optimizing the artifacts themselves or design principles were uncovered. Consequently, important stakeholder groups were involved through human-centered qualitative evaluations and generating essential insights regarding the problem and solution space.

In addition to qualitative research methods, human-centered quantitative methods for evaluating prescriptive design knowledge and instantiated artifacts were used. Between-subjects were conducted to investigate the XUIs in a controlled experiment and environment, an established design evaluation method (Hevner et al., 2004). Such experiments in the form of field or laboratory experiments have a long history in ISR (Galliers & Land, 1987). The experiments that are part of this cumulative dissertation were conducted as online experiments, which is an approach gaining relevance and is highly suitable for studying human behavior (Fink, 2022). In the individual research projects part of this cumulative dissertation, human behavior refers to the perception of different design configurations in XUIs, the comparison of black box AI and XAI, or the interaction of users with the XUI. In addition, online experiments provide desirable advantages, such as the ability to recruit a broad range of participants and the accommodation of large numbers of participants (Karahanna et al., 2018). With colleagues and a small sample size of the targeted stakeholders, pilots and pre-studies have been conducted to ensure that the experiment design is appropriate for the individual research project. Such pre-studies are relevant and effective for uncovering potential study design problems and advancing their structure (Waters, 2011). In early research projects, the platform CloudResearch, formerly TurkPrime, was combined with Amazon Mechanical Turk. CloudResearch is an Internet-based platform that enables researchers to save time and resources while increasing their data quality, and it provides versatile and helpful features to conduct complex behavioral studies via Amazon Mechanical Turk (Litman et al., 2017). The platform Prolific was used in later online experiments to be independent of two platforms, i.e., CloudResearch and Amazon Mechanical Turk. Prolific also offers good support, extensive features, and other quality features. In addition, the participants on Prolific generate high data quality, for example, in terms of attention, comprehension, honesty, and reliability (Peer et al., 2022). To analyze the quantitative data sets and appropriate statistical methods were used. These include comparative tests such as the Mann-Whitney U test (A4), t-test (A6), descriptive statistics (A4, A5, and A6), or structural equation models (A6).

Table 4 concisely summarizes the research methods, evaluation and data collection methods, and the justificatory knowledge for the individual article.

Article	Research Method	Evaluation and Data Collection Method	Justificatory Knowledge
A1	• Literature Review	• Systematic Literature Review	• Explainability and XAI in ISR
A2	• Literature Review • ISR Framework for DSR	• Systematic Literature Review • Qualitative Evaluation with Semi-Structured Interviews	• XAI in ISR • Design of Multi-Modal Mobility Platforms • Heuristic Decision-Making
A3	• Literature Review • ISR Framework for DSR • Framework for Evaluation in DSR	• Systematic Literature Review • Qualitative Evaluation with Semi-Structured Interviews	• XAI in ISR • (X)AI-based Fake News Detection in Social Media • Source Credibility Theory
A4	• Literature Review • Cyclic Framework for DSR • Framework for Evaluation in DSR • Framework for Reusability Evaluation of Design Principles	• Systematic Literature Review • Qualitative Evaluation with Semi-Structured Interviews • Quantitative Evaluation with Online Surveys and Online Experiments with a Between-Subjects Design	• XAI in ISR • Transfer Learning • (X)AI-based Hate Speech Detection in Social Media • Knowledge from HCI
A5	• Literature Review • Cyclic Framework for DSR • Framework for Reusability Evaluation of Design Principles	• Systematic Literature Review • Think-Aloud Method • Semi-Structured Interviews • Qualitative Evaluation with Semi-Structured Interviews • Quantitative Evaluation with Online Surveys	• XAI in ISR • (X)AI-based Clinical Decision Support Systems • Theory of Interactive Media Effects • Knowledge from HCI
A6	• Literature Review • Cyclic Framework for DSR • Framework for Evaluation in DSR • Framework for Reusability Evaluation of Design Principles	• Systematic Literature Review • Quantitative Evaluation with Online Experiments with a Between-Subjects Design	• XAI in ISR • Customization Features in (X)UIs • Theory of Interactive Media Effects • Human-XAI Interaction • Knowledge from HCI

Table 4. Overview of research methods, evaluation and data collection methods, and justificatory knowledge.

Now that the DSR paradigm has been motivated as a superordinate research procedure, the overlaps with human-centeredness have been highlighted, and the DSR methods and other research methods have been dealt with more deeply. Section 4 will summarize the research contributions and be used to develop the ISDT for human-centered XUIs. It represents the core design knowledge contribution and summary of all articles in this cumulative dissertation.

4 Summary and Consolidation of the Research Contributions

4.1 Summary of the Research Contributions of the Individual Articles

After DSR was described as the overarching research design in the last section and the research methods of the different research projects were described, the scientific contributions resulting from the research projects are to be summarized in this section. For this purpose, an overview is first given of which research questions were processed by the individual research projects, which is visualized in Figure 12. Article A1 represents the basis for the identified research questions and was thus the starting point for the higher-level RQ1. Article A2 is a short paper that addresses RQ1.1. The research questions RQ1.1 and RQ1.2 are addressed by Article A3. The extensive DSR project presented in A4 contributes to answering all research questions, i.e., RQ1, RQ1.1, and RQ1.2. Article A5 addresses the research questions RQ1 and RQ1.1. The last article, A6, addresses all research questions, RQ1, RQ1.1, and RQ1.2. Consequently, the different research projects contribute individually to answering the overarching RQ1 and developing a design theory in the form of an ISDT.

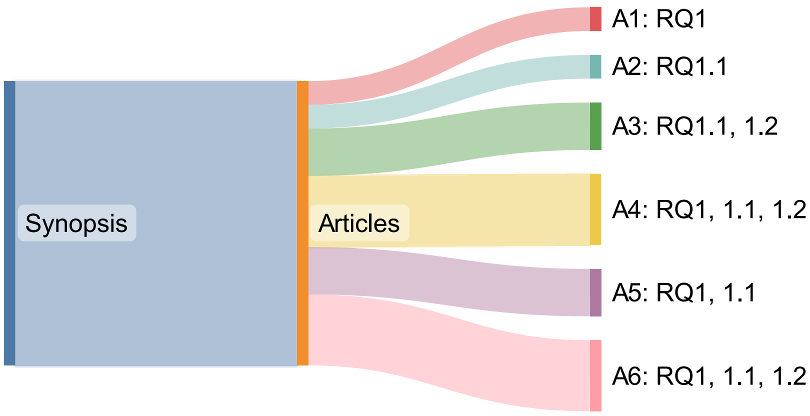


Figure 12. Relationship of the individual research projects included in the cumulative dissertation and the research questions.

A total of six articles flow into this cumulative dissertation. The focus of the first article, **A1**, was to examine the topic of XAI in the context of ISR. As described before, the research project was based on a systematic literature search and represents the foundation of the cumulative dissertation. The core focus of the article was to examine what role the ISR community can play in the context of XAI research. For example, the potential risks of black box AI were described to clarify XAI's relevance and added value from an ISR research perspective. The historical research on explainability in the context of the 1980s and 1990s decision support systems was also reviewed. Based on the developed knowledge base, which resulted from the systematic literature research, XAI-related terms were defined, overriding reasons for using XAI were described, overarching stakeholder groups were described, and quality criteria for personalized explanations were established. The article ended by showing the diverse future research possibilities for the ISR community from a behavioristic and design-oriented perspective.

The second article, **A2**, a short paper, presents a research project on the design of multimodal mobility platforms and their UIs. These platforms derive information from a vast space of solutions that may lead to a different presentation of the best-suited options. Since travelers cannot always comprehend how the recommended options are composed, they perceived it as opaque. A human-centered qualitative approach in semi-structured interviews was chosen to understand how users perceive current solutions. Recruiting diverse users for the interviews, including operations research experts, was possible. Based on a previously developed interview guide, the interviews were conducted, and the transcribed data were analyzed using the thematic analysis method. Valuable insights were gained from this. On the one hand, challenges were identified that users encounter while using such platforms, such as the individual trade-offs between preferences (e.g., time and costs) that are not considered, and travelers with special needs may not find the filter options they require. Meta-requirements were first developed to overcome the complex challenges and were addressed by design principles.

The relevant issue of fake news detection was examined in Article **A3**. A focus was placed on social media platforms since fake news has been proven to spread rapidly in these environments. Therefore, this article examined the design of services that support users in assessing the credibility of news content in online environments. The systematic literature research has emphasized a lack of design knowledge, especially in the context of fake news warnings developed in this article. Furthermore, relevant challenges were identified through the systematic literature review. These challenges were oriented towards the danger of fake news from the user perspective, but challenges regarding the AI-based detection of fake news were also considered. In the next step, the identified challenges were addressed by design requirements grounded in the knowledge base and source credibility theory. Based on this knowledge base, design principles for XUIs were defined for services that support users in assessing the credibility of news content in online environments and associated design features. The design features were instantiated in the prototype XUI, which was human-centered and qualitatively evaluated with 13 relevant stakeholders using semi-structured interviews. The transcribed data were analyzed using the thematic analysis method. The analysis aimed to understand how the targeted users perceive such XUIs and the proposed design elements when interacting with them. In addition, the aim was to identify optimization potential for future research. The UI was perceived as useful, and the participants described their preference for lightweight UI design over complex decision support systems. When automated fake news detection is integrated, they demand explanations for the output. Participants were also willing to rate information sources and perceived the ease of comprehension through visualized rating scales as positive.

A larger DSR project was presented in article **A4**, in which three design cycles were run through to design and evaluate human-centered XUIs for AI-based decision support systems in the context of hate speech detection. The knowledge base was developed through a systematic literature search. It resulted in lack of design knowledge for XAI-based hate speech detection systems and human-centered evaluation of such systems. This research gap was addressed by designing an XUI for such a system and involving relevant stakeholders in the form of professional social media moderators in the various human-centered qualitative and quantitative evaluations. For the development of the XUI, generic design requirements from established literature were initially used, which helped to achieve human decision-maker goals. These generic design requirements were translated into context- and project-specific design requirements for the DSR project. Subsequently, design principles were developed and anchored in the previously developed knowledge base, which was very diverse and included disciplines such as ISR, psychology, computer science, and HCI. The design principles were addressed with design features and then transferred to the XUI. Different human-centered evaluations took place in the three design cycles.

In the first human-centered qualitative evaluation, 11 professional moderators from social platforms were involved, and the aim was to evaluate the design in terms of perception and usefulness from the user's perspective. In addition, the optimization potential for the XUI was identified. A second human-centered quantitative evaluation was performed with 190 participants with experience moderating social platforms. On the one hand, constructs such as perceived ease of use, usefulness, or intention to use were measured. On the other hand, text fields were integrated to receive participant feedback, which should again be used to optimize the XUIs. The data from the first human-centered evaluation and the data produced by the text fields were each analyzed with thematic analysis. In the third design cycle, the reusability of the design principles was quantitatively evaluated with 80 software developers in a human-centered manner and was rated very positively. In addition, a between-subject design experiment was carried out with the final XUI in the third design cycle. A group of participants was confronted with an AI version, i.e., a version without XAI features. The other group interacted with the XUI, with XAI features. A total of 360 subjects took part. The statistical evaluation showed that the XUI with XAI features was rated as having led to a significantly lower perceived cognitive effort and a significantly higher perceived informativeness, mental model, and trustworthiness. Finally, the design knowledge was summarized in an explanatory design theory, which summarizes a design solution's general requirements and components.

In the penultimate article, **A5**, an archaeological DSR approach was used to analyze an existing UI of an XAI-based medical decision support system. The examined system can be used for the AI-based classification of skin lesions. The system should be usable for both patients and doctors. A systematic literature search showed that for systems of this type, which are used in a medical context, multi-stakeholder needs, requirements, and expectations are not regularly considered. Thus, a human-centered archaeological in-situ analysis of the existing system was planned and carried out. Twelve participants were involved, with six representing the patients' and the physician's perspectives. The think-aloud method was used in this analysis, while the participants had to solve some tasks with the system. This was followed by a semi-structured interview, which focused on the perception of the XUI. The resulting data were also analyzed in this project using thematic analysis. Challenges that users experienced while interacting with the XUI were identified, and based on them, an extensive set of design requirements has been developed. The design requirements were then addressed with design principles anchored in the research status quo and through the Theory of Interactive Media Effects. The developed design principles were initially evaluated concerning their reusability with four experienced software developers through a human-centered qualitative evaluation in semi-structured interviews. Based on the knowledge gained, the design principles were slightly optimized and then quantitatively evaluated in a human-centered manner with 66 experienced software developers. The design principles were positively perceived in the human-centered qualitative and quantitative evaluation. Extensive design knowledge was thus developed, considering, and involving relevant stakeholder groups' perspectives.

Article **A6** presents the final research project of this cumulative dissertation. In this DSR project, the concept of customization for the design of XUIs was conceptualized, transferred to prescriptive design knowledge, instantiated, and evaluated from a human-centered perspective. Customization is well-established in various areas, including HCI research, for example. Interestingly, in addition to customization, the concept of personalization has already found its way into the XAI field. However, customization has not yet been introduced or investigated in XUIs. The core contribution of this article is twofold. On the one hand, the conceptualization of customization for the design of XUIs and the resulting prescriptive design knowledge represent one contribution. The human-centered quantitative evaluations also enabled the investigation of the interaction between users and the XUI, particularly with the customization features.

Furthermore, the customization features' influence on the explanation's perceived quality was proven. The explanation quality was measured using the constructs of perceived interestingness, perceived informativeness, perceived interactivity, and satisfaction with the explanation. Through a between-subjects experiment design with two groups of 90 people each, one group used an XUI with customization features, and the second group used an XUI without customization features. Significantly higher values for the XUI with customization features were measured for all these constructs. Interestingly, the participants who used the XUI with customization features also achieved higher task performance in the classification task in the human resources context. In the second design cycle, the design principles were expanded and again quantitatively evaluated in a human-centered manner. The second online experiment examined user engagement and satisfaction when using XUIs with customization features. An experiment with 215 test subjects was carried out for this purpose. It examined how the dimensions of perceived interactivity, perceived interestingness, and perceived customization predict user engagement, which was confirmed for perceived interactivity and customization but had to be rejected for perceived interestingness. Furthermore, the statistical analysis of the data showed that user engagement positively predicts user satisfaction.

In contrast to the first design cycle, in which the group with customization features clicked on the cosmetic customization features significantly more often than on functional customization, the evaluation in the second design cycle revealed a different picture. No significant difference between cosmetic and functional customization was identified. However, the XAI method customization feature was used significantly more frequently than the others. Ultimately, the design principles were evaluated for their reusability with 61 experienced user interface and user experience designers. The design principles were rated positively, and the test subjects showed a high tendency to use them for suitable projects or recommend them to colleagues.

After the individual research articles that are part of the cumulative dissertation have been summarized in this chapter, the topic of design theory is presented in the next chapter. Since an ISDT represents a form of design theory, how the concept of design theories in DSR can be understood should be clearly explained.

4.2 Design Theories in Design Science Research

Since design theory is relevant in this cumulative dissertation, it will be separately described in this subsection. The main contribution of this cumulative dissertation is represented by the consolidation of the insights gained and generated design knowledge within the individual research articles into a design theory. Design theories can serve as intellectual tools by which the ISR community can not only contribute to technological innovations but also engage in tackling highly relevant real-world problems (Beck et al., 2013; vom Brocke et al., 2020; Goes, 2014; Gregor & Hevner, 2013; Lukyanenko & Parsons, 2020). Nevertheless, it is worth noting that within the DSR community, different camps with a stronger focus on either the artifact or design theory exist (Baskerville et al., 2018; Hevner et al., 2004). Moreover, DSR is often criticized for the exaggerated usage of the word "theory," which is also in line with the claimed "theory fetish" in the discipline of ISR (Ivari, 2020). Concluding from this, I adapt the perspective of Gregor and Hevner (2013), who state to the before mentioned DSR camps that their aim "[…] is to harmonize what we see as complementary rather than opposing perspectives, a repositioning that can enhance the conduct and reach of rigorous and impactful DSR." (p. 338). Following Gregor (2006), design theories can be categorized as theories for design and action, which can be informed by all other classes of theory and are strongly interrelated with theories for explaining and predicting. They represent prescriptive scientific knowledge, a desirable objective for theorizing about the progress of a class of artifacts (Baskerville et al., 2018; Goes, 2014). Moreover, like design principles, design theories can also be seen as technological knowledge since they are prescriptive in nature and invoke functional explanations (Baskerville & Pries-Heje, 2019).

Different types of design theories exist, such as explanatory design theories with varying focuses (Baskerville & Pries-Heje, 2010; Niehaves & Ortbach, 2016), utility theory (Venable, 2006), or the ISDT (Gregor & Jones, 2007). Ivari (2020) provides an in-depth conceptualization of the different types of design theories. In this cumulative dissertation, an ISDT for human-centered XUIs based on Gregor and Jones (2007) is developed. The decision is justified because their conceptualization of a design theory fits well with the focus of the cumulative dissertation since it highlights the importance of the artifact and theory taken together (Baskerville et al., 2018). In addition, the blueprint for a design theory by Gregor and Jones (2007) was applied and published in top IS journals, which further legitimized the decision (e.g., Avdiji et al., 2020; Chanson et al., 2019; Coenen et al., 2018; Giessmann & Legner, 2016; Kane et al., 2021; Lycett & Radwan, 2017; Mandviwalla, 2015; Meth et al., 2015; Morana et al., 2019; Venkatesh et al., 2017). The roots of information systems design theories go back to the important research contributions of Walls et al. (1992; 2004). Such design theories can be developed in combination with an associated instantiated artifact (e.g., Morana et al., 2019), by theorizing about a relevant phenomenon (e.g., Kane et al., 2021), or by observing already existing artifacts (e.g., Avdiji et al., 2020). Ultimately, design theories represent a set of abstract statements, enabling researchers and practitioners to tackle real-world problems in contexts different from where the design theory was developed (Lukyanenko & Parsons, 2020).

This subsection served to elaborate on the concept of design theories in DSR and the ISR community and thus appropriately classify the relevance of the subsequently developed ISDT for human-centered XUIs. The following subsection presents the successive development of the individual components of the ISDT. The section ends with merging the individual components into the ISDT for human-centered XUIs to give an overall overview.

4.3 Towards an Information Systems Design Theory for Human-Centered XUIs

4.3.1 Defining the Purpose and Scope

To formalize the ISDT for human-centered XUIs, I rely on the six core components and the two additional components proposed by Gregor and Jones (2007). The resulting design theory represents a consolidation of all research projects as part of the cumulative dissertation. The following subsections define the individual components, which result from a reconceptualization of the individual research contributions supplemented by the status quo of relevant research streams.

The *first component* describes the purpose and scope (the causa finalis) that specify the artifact type to which the design theory is applicable. This is typically achieved by concisely defining the aim pursued with the design theory (e.g., Meth et al., 2015; Morana et al., 2019). The pursued objective of the design theory is to guide the design (Giessmann & Legner, 2016) of human-centered XUIs for practitioners and design researchers. Therefore, the first component of the ISDT is used to describe what type of artifact can be used and to emphasize the relevance of the established ISDT for human-centered XUIs (Gregor & Jones, 2007).

In the different research papers that are part of this cumulative dissertation, the design of XUIs and their perception by relevant users in different application domains were examined. The systematic literature searches as part of the individual research projects produced a coherent picture. On the one hand, there is only very isolated, limited, generalizable design knowledge for designing XAI systems, particularly the XUIs, which represent the interaction interface for end users with the XAI system. On the other hand, the systematic literature search has shown that there is still much research to be done to examine and understand the perception of XUIs, different design configurations, and the interaction experience of users. These research gaps exist since there are many different XAI methods, new ones are constantly being introduced, and users have very different previous experiences, expectations, requirements, and knowledge. This results in a large number of possible research projects. The research gaps were processed through the individual research projects, part of the cumulative dissertation. The ISDT for human-centered XUIs developed in the cumulative dissertation thus represents an essential contribution to these research streams and gaps. Because established prescriptive design knowledge can be adopted for XUIs in various application domains, the influences and effects of design configurations in XUIs can be understood using the propositions included in the ISDT. Therefore, prescriptive design knowledge can be relevant for many XAI-based decision support systems in many application domains. However, it is important to emphasize that despite the generalizability of this design theory, the adoption of design knowledge is ideally accompanied by a human-centered approach to determining what goals should be achieved with specific design configurations for users to have a positive interaction experience or to work with the XAI system, and thus XUI in everyday work is optimally supported.

After the first components of the ISDT for human-centered XUIs have been described in this subsection, the causa finalis, the second component, is presented in the following subsection.

4.3.2 Establishing the Constructs

In the *second component*, constructs (the causa materialis), the representations of the entities of interest that are of interest for the design theory are presented. Frequently, this component contains the constructs used to evaluate the instantiated prescriptive design knowledge (e.g., Diederich et al., 2020; Venkatesh et al., 2017).

Constructs can be physical phenomena or abstract theoretical terms and represent the entities of interest within design theory (Gregor & Jones, 2007). The constructs stem from different research streams and are highly relevant for evaluating XAI (e.g., Vilone & Longo, 2021). Some constructs were derived from the body of literature on XAI, like associated goals of XAI such as trustworthiness, informativeness, or interactivity (Nazar et al., 2021), which all can be influenced by the designed explanations (Bunde, 2023; Meske & Bunde, 2023). Other constructs were identified by integrating relevant constructs like perceived usefulness from the Technology Acceptance Model (Davis, 1989; Venkatesh et al., 2003), constructs like the mental model with a focus on processes from the Mental Model Theory (Vitharana et al., 2016), or perceived cognitive effort relevant for decision strategies (e.g., Wang & Benbasat, 2009). It is important to note that a few constructs can be assigned to both goals and notions of XAI, such as interactivity, or interestingness, informativeness, satisfaction, user engagement, and customization (e.g., Al-Natour et al., 2022; Li et al., 2021; Shin et al., 2022; Vilone & Longo, 2021; Wang & Sundar, 2018), which are also important constructs in the Theory of Interactive Media Effects (Sundar et al., 2015). Relevant constructs for the reusability evaluation of the design principles were adapted from Ivari et al. (2021) and included accessibility, importance, novelty and insightfulness, actability and guidance, and effectiveness.

The constructs of perceived ease of use, usefulness, and intention to use were used for the human-centered quantitative evaluation in the second design cycle in A4. In the third design cycle of A4, the constructs of perceived cognitive effort, perceived informativeness, mental model (process), and trustworthiness were used in a between-subject experiment design. In addition, the constructs of accessibility, importance, novelty and insightfulness, actability and guidance, and effectiveness were used to evaluate the reusability of the design principles in A4. The identical constructs for reusability evaluation of design principles were also used in the research project presented in A5. In article A6, the constructs of perceived interestingness, perceived informativeness, perceived interactivity, and satisfaction with explanation were measured in the first design cycle. In addition, the task performance was calculated as accuracy. In the second design cycle, more constructs were used to investigate the effects of customization features in XUIs during the interaction experience. The influence of perceived interactivity, perceived interestingness, and perceived customization was examined on user engagement. Moreover, how user engagement predicts satisfaction was investigated. Finally, in this project, the design principles were evaluated by practitioners concerning reusability.

After the constructs were examined in this subsection, the third component follows in the following subsection, which includes the principles of form and function.

4.3.3 Principles of Form and Function

The next and *third component*, the principle of form and function (the causa formalis), summarizes the abstract blueprint or architecture that describes the design solution. Since design principles have evolved into a well-established contribution in terms of prescriptive design knowledge (Gregor et al., 2020), most ISDTs use this component to present the developed design principles (e.g., Coenen et al., 2018; Diederich et al., 2020). This design theory focuses on the artifact, i.e., XUIs. Therefore, the necessary and suitable information for this component is the artifact's properties, functionalities, features, or characteristics when it is instantiated, which are represented through the design principles (Gregor & Jones, 2007).

The here-introduced design principles are formalized according to the scheme by Gregor et al. (2020). In addition, the design principles are an amalgamation of those that focus on XAI-related design elements of the XUIs in the various research projects included in the cumulative dissertation. Each reconceptualized design principle is explained from which research projects they originate to achieve high degree of transparency and comprehensibility.

The first design principle is about communicating and presenting the AI performance as part of the explanation. The AI performance can be, for example, the accuracy of the AI, which it achieves on average, or the probability for a specific output, such as a classification. The relevance of this design principle is supported by articles A3, A4, and A5. In A3, the AI performance for a specific classification was realized by a visualized rating scale based on the traffic lights' design. The human-centered qualitative evaluation showed that users can easily and intuitively interpret this representation. In A4, the AI performance was represented by the probability of the AI in percentages to communicate to users how specific the AI is in the present case of hate speech detection. The term probability was replaced by confidence to make the value easier to understand. The human-centered evaluation in the first design cycle showed that users appreciate this value to assess the AI performance better and thus identify cases in which the AI could be wrong. In the second design cycle, a human-centered quantitative evaluation was carried out, and subjects had the opportunity to submit qualitative feedback through text fields. Here, too, comments were found that positively described the information regarding the confidence of the AI. In A5, an existing XAI-based medical decision support system was examined for the human-centered classification of skin lesions. The stakeholders from the patient and physician perspectives each described the representation of the probability as important information in the XUI. These findings motivate the first design principle. Table 5 presents the first design principle.

Design Principle Title	**Principle of Performance Communication**
Aim, implementer, and user	To allow users (users) to correctly comprehend the recommendation of an XAI-based system (aim),
Context	in the context of decision-making supported by a human-centered XAI-based system,
Mechanism	the system should communicate the performance of the underlying AI in an easy-to-understand way in the XUI, which requires no prior knowledge or expertise in the context of AI,
Rationale	so that the performance communication in XUIs enables a diverse user base to understand and interpret the AI's performance while requiring an appropriate cognitive effort and supporting the development of correct mental representations.

Table 5. Design principle of performance communication.

Articles A2, A3, A4, and A5 emphasized the relevance of involving relevant stakeholders when selecting the XAI method and the associated presentation format of the explanation. For example, the stakeholders involved in A2 and A3 communicated during the human-centered, qualitative semi-structured interviews, that explanations and their presentation format are essential to them when they are confronted with or use AI-based decision support. In article A4, it was also communicated through the semi-structured interviews during the human-centered evaluation in the first design cycle by the stakeholders involved, that the explanation was essential to them if automated hate speech detection was used in the moderation of social media platforms. In particular, the presentation format in the instantiated XUI was described as very positive and easy-to-understand, with post-hoc explanations being visualized as a heat map on the texts. Colors presented the classes, the strength or weakness of the color, and the relevance of the word marked in color. Similar sentiments towards the explanation were also found in the qualitative data from the text fields in the human-centered quantitative evaluation in the second design cycle. In article A5, different XAI methods were used in the XAI-based medical decision support system. Various individuals from the patient stakeholder group and the physicians communicated the relevance of the explanations. They also gave feedback on optimizing the explanations to make them easier to understand and interpret. Thus, the involvement of the targeted stakeholder groups through a human-centered approach is crucial to identifying the relevant XAI methods and the presentation format of the explanations, which motivates the following design principle: Table 6 presents the second design principle.

Design Principle Title	**Principle of Human-Centered Selection of XAI Methods and Presentation Format**
Aim, implementer, and user	To allow designers, developers, practitioners, and researchers (enactor) to select either the correct XAI method or mix of XAI methods and presentation format (aim) for their targeted user basis (users),
Context	in the context of decision-making supported by a human-centered XAI-based system,
Mechanism	they have to involve the users during the process of requirements elicitation, design, and evaluation of XUIs within an iterative process,
Rationale	so that they can take advantage of human-centeredness and identify individual information needs, expectations, biases, and acceptance barriers to consider these aspects during the design process.

Table 6. Design principle of human-centered selection of XAI methods and presentation format.

To make informed decisions, additional information is required in addition to the XAI design elements. It can be a great help to compare the facts to be processed with, for example, historical facts of a similar case or of a different case to support the decision-making process. This includes the possibility of case- and output-related comparisons. The motivation for this design principle comes from articles A2, A3, A4, and A5. For example, the qualitative, semi-structured interviews in A2 showed that users would perceive information positively when filtering content on multimodal mobility platforms, which shows the influence of different filters on the result. This allows users to estimate better how the output changes due to the adjustment of filters compared to the current output. Furthermore, outputs outside the actual filtering are also interesting for users, and these should be marked accordingly to enable a comparison with the current output.

In A3, various online news credibility assessment design elements were instantiated. Elements were designed to enable users to understand and assess the existing classification, for example, by providing sources of disproof if fake news was detected. In this way, users can compare the available news content and the classification with other sources and judge. Various design elements were instantiated in A4. This included, for example, the simple navigation between different cases of detecting hateful content to compare and process them quickly and easily. The comparison between these cases also enables users to interpret explanations and classifications to assess the performance and reliability of the AI. Participants from the human-centered qualitative study of the XAI-based medical decision support system explicitly communicated that a feature for comparing different cases would be beneficial to be able to compare them and thus better assess the AI and its performance. These different insights motivate the following design principle: Table 7 presents the third design principle.

Design Principle Title	**Principle of Case- and Output-Related Comparison**
Aim, implementer, and user	To support users (users) in developing a holistic understanding of the XAI-based output in the scope of the task to be performed (aim),
Context	in the context of decision-making supported by a human-centered XAI-based system,
Mechanism	the system should provide features that enable case- and output-related comparisons within the XUI,
Rationale	so that the users can understand how the XAI-based system has performed on different cases, gain new insights, and learn more about the system's functionality.

Table 7. Design principle of case- and output-related comparison.

Contextual information helps users develop a holistic understanding of the output of the task to be solved in the decision-making scenario. This information is elementary for decision-making and should be specifically integrated into XUIs to support the decision-making process. This goal is the basis of the following design principle: Different articles that are part of the cumulative dissertation justify this approach. This includes A3, A4, and A5. There are many ways to integrate context-related information into an XUI. In the case of A3, for example, additional information about the news content was integrated. This was the source of the news content, which is not always easily identifiable on social media, which was supplemented by a visual rating scale for source credibility. The sources of disproof can also be counted among the context-related information because if news content has been classified as fake news, then the sources of disproof can provide context-related information so that users can independently gain an appropriate understanding of the subject. In A4, context-related information was also included in the XUI for explainable hate speech detection. This includes, for example, the amount of hateful and non-hateful content that a user has published on the moderated platform, a historical analysis showing the development of the publication of hateful and non-hateful content, and the history of the actions that have already been taken against the particular user. All of these elements support the moderator of the social media platform in evaluating the user's behavior beyond a specific case to initiate appropriate action against him, for example, if he spreads hateful content to the user. In A5, the patients, in particular, expressed great interest in receiving context-related information concerning data processing. When patients upload private data and images to an XAI-based medical decision support system, they want to know how and by whom this data is processed and where it is stored and processed. The system examined also provided many other context-related details, such as the part of the body from which the photo of a conspicuous skin area originated. These insights and knowledge motivate the following design principles: Table 8 presents the fourth design principle.

Design Principle Title	**Principle of Contextual Information**
Aim, implementer, and user	To enable users (users) to familiarize themselves with all necessary information relevant to the task at hand (aim),
Context	in the context of decision-making supported by a human-centered XAI-based system,
Mechanism	the system should provide contextual information that is relevant to the task to be performed in the XUI,

Rationale	so that users can make informed decisions and either accept or reject the XAI-based recommendation.

Table 8. Design principle of contextual information.

Since the users of XAI-based decision support systems can have a wide variety of experiences, needs, expectations, or even cognitive biases, assistance with the correct interpretation of the explanation can be helpful and thus support users in developing a correct mental model. These findings come in particular from Articles A5 and partially A4. Since users have different levels of expertise in AI and digitization, a human-centered approach is precious to determine whether the targeted users need additional help or information to interpret the explanations provided correctly. In A4, for example, a hint for interpreting the explanation for hate speech, the heat map, which was projected in color onto the text, was offered. This element was tested in Design Cycle 2. Via the text fields offered in the human-centered quantitative evaluation, the explanation was described as easy-to-understand and interpret, so this information was removed again in this project. On the other hand, in A5, each participant from the stakeholder group of patients and doctors communicated that they would like information and assistance in interpreting the explanations. This information is relevant so the explanation can be interpreted correctly, and users can develop a correct mental model. This motivates the following design principle: Table 9 presents the fifth design principle.

Design Principle Title	**Principle of Assistance with Interpretation**
Aim, implementer, and user	To support users (users) in correctly interpreting the included information in the provided explanation (aim),
Context	in the context of decision-making supported by a human-centered XAI-based system,
Mechanism	the system should provide appropriate assistance depending on the XAI method or mix of XAI methods being used in the XUI,
Rationale	so that users can assess the soundness, completeness, and faithfulness of explanations or diagnosing and correcting flawed XAI-based outputs.

Table 9. Design principle of assistance with the interpretation.

The following two design principles were developed, instantiated, and quantitatively evaluated in a human-centered manner in A6. The statistical analysis of the data collected from the evaluation showed that the customization features significantly impacted the perception of explanation quality in this experiment and significantly increased task performance. Furthermore, it should be noted that the statistical analysis showed that the customization feature for cosmetic customization was used significantly more often than the functional customization feature. Since both customization features were instantiated together and planned as interconnected design configurations, the following two design principles are mainly motivated by A6. Moreover, the human-centered qualitative interviews in A2 have uncovered that involved stakeholders desire individualized explanations, supporting the relevance of customization of explanations in different application domains. Consequently, the following two design principles are established, and Tables 10 and 11 present the sixth and seventh design principles. Both design principles were developed, instantiated, and evaluated in Article A6.

Design Principle Title	**Principle of Cosmetic Customization**
Aim, implementer, and user	To allow users (users) to interact with AI-generated explanations for adjusting their visual representation from a set of pre-defined alternatives (aim),
Context	in the context of decision-making supported by a human-centered XAI-based system,

Mechanism	the system should provide easy-to-use settings for users to customize the visual representation according to their individual needs as well as preferences in the XUI,
Rationale	so that the cosmetic customization in XUIs can provide a high degree of interactivity, interestingness, informativeness, satisfy users, and ultimately improve the perceived explanation quality.

Table 10. Design principle of cosmetic customization.

Design Principle Title	**Principle of Functional Customization**
Aim, implementer, and user	To allow users (users) to interact with AI-generated explanations for adjusting the number of relevant features displayed in the XUI from a set of pre-defined alternatives (aim),
Context	in the context of decision-making supported by a human-centered XAI-based system,
Mechanism	the system should provide easy-to-use settings for users to customize the scope of relevant features displayed in the XUI,
Rationale	so that the functional customization in XUIs can provide a high degree of interactivity, interestingness, informativeness, satisfy users, and ultimately improve the perceived explanation quality.

Table 11. Design principle of functional customization.

Explanations for the same type of data and AI model can generate very different explanations. Since users have different expectations, information needs, demands, and cognitive biases, an explanation that is very easy to understand for one user can be difficult for another user to understand. In A5, where the XAI-based medical decision system was investigated, users rated positively that the system offers different types of explanations. In the second design cycle of article A6, inspired by this knowledge, a design principle was formalized and instantiated, enabling XAI method customization. The analysis showed that users actively used this design element. The following design principle is motivated by the collected knowledge. Furthermore, the design principle was developed, instantiated, and evaluated during the second design cycle in A6. The statistical analysis of the data regarding cosmetic, functional, and XAI method customization in XUI uncovered that the XAI method customization feature was used significantly more often. Table 12 presents the eighth and last design principles.

Design Principle Title	**Principle of XAI Method Customization**
Aim, implementer, and user	To allow users (users) to interact with AI-generated explanations for adjusting the XAI method displayed in the XUI from a set of pre-defined alternatives (aim),
Context	in the context of decision-making supported by a human-centered XAI-based system,
Mechanism	the system should provide easy-to-use settings for users to customize the presented explanation displayed in the XUI,
Rationale	so that the XAI method customization in XUIs can provide an engaging and satisfying user experience.

Table 12. Design principle of XAI method customization.

Eight design principles from the different research projects in the cumulative dissertation were brought together. They are thus anchored in the knowledge base of the individual research projects but were also evaluated using human-centered qualitative or quantitative methods. It is important to note that if future research or practitioners wish to adopt or extend these design principles, users should be involved through human-centered methods to support the selection or extension of the existing design principles.

This subsection presented an essential part of the ISDT for human-centered XUIs. These design principles, representing the principles of form and function in the ISDT, are relevant for practitioners and researchers in designing XAI systems and XUIs since they can be adopted, adapted, or expanded. In the following subsection, the fourth component, artifact mutability, is presented.

4.3.4 Artifact Mutability

On an overarching level, the *fourth component*, artifact mutability, describes to what extent changes to the artifact the theory holds and applies. The information presented within this component summarizes the potential for changes in the artifact and potential fields of applications, together with the necessary adjustments to the artifact's functionalities (e.g., Chanson et al., 2019; Kane et al., 2021).

This component of the design theory arises due to the dynamic nature of IS artifacts, which are ascribed to being mutable by nature (Gregor & Jones, 2007). Artifacts are also described as evolving, primarily through their high degree of flexibility and adaptability that can be enabled by feedback loops, which can be used to optimize either the design knowledge or the instantiated artifact (Simon, 1996). The concept of mutability is also essential when talking about (X)AI since methods and techniques from this field constantly evolve (Berente et al., 2021). Therefore, a human-centered approach by constantly involving relevant stakeholders during the design and evaluation phases can lead to revised and optimized design solutions in the form of prescriptive design knowledge or instantiated XUIs. This line of argumentation further emphasizes the relevance of human-centered approaches in (X)AI, which is also acknowledged by recent research on human-centered (X)AI (Nazar et al., 2021; Schneiderman 2020a; Schneiderman 2020b). Therefore, a clear definition of the degree of mutability is important (Lycett & Radwan, 2017).

The presented design theory is mutable in that it must not be adopted to its full capacity. By taking a human-centered perspective and involving relevant stakeholders in the early phases of the development of XUIs, the essential requirements and needs can be discovered, and suitable design principles can be adopted as required. Therefore, relevant stakeholders can be involved in different phases, such as design or evaluation. This allows valuable insights to be gained, which can help identify the proper design knowledge. Thus, design knowledge that should be adopted or learned from the existing design knowledge contained in the ISDT for human-centered XUIs can be identified. Since prescriptive design knowledge does not refer to specific XAI methods that must be used, the constant developments in the XAI field are considered and allow freedom in selecting suitable XAI methods. Furthermore, no application type, such as a desktop or a web application, is prescribed because the established prescriptive design knowledge is mutable insofar as it can be used for various application types, platforms, and end devices. In addition, the design theory has many other types of mutability, such as the type and amount of data processed, which can vary greatly depending on the application and domain. In such cases, practitioners or researchers can adopt the appropriate elements and components of the ISDT for human-centered XUIs. Design features were also established when the prescriptive design knowledge was instantiated in the research articles in the cumulative dissertation. The special thing about the design features is that they can be adjusted according to the user's requirements or the requirements of the project at hand. For example, certain context information is part of the prescriptive design knowledge, as it is in A4, where the XUI for hate speech detection was designed for use by social media moderators. Context information could be integrated into the XUI in a completely different application domain. Thus, the ISDT for human-centered XUIs offers specific instructions for designing XUIs but leaves much flexibility to adapt the prescriptive design knowledge for the specific application. This is where the human-centered perspective gains relevance, which can help identify design theory elements that may be relevant to a project or how they need to be adjusted.

In this subsection, artifact mutability was explained, representing the fourth component of the ISDT for human-centered XUIs. The testable propositions, the fifth component of the ISDT, are presented in the following subsection.

4.3.5 Testable Propositions

The *fifth component* presents the truth statements about the design theory in this subsection. Here, most researchers either establish propositions without involving an evaluation of the artifact (e.g., Kane et al., 2021) or hypotheses that were established for the evaluation of the artifact (e.g., Diederich et al., 2020). The design theory includes testable propositions about the artifact that can be tested by instantiating the prescriptive design knowledge, for example, during the evaluation (Gregor & Jones, 2007; Hevner & March, 2003). It is a well-established approach to derive propositions to validate them and therefore demonstrate that the proposed design knowledge in an instantiated form achieves the intended goals or performs better than existing solutions (e.g., Diederich et al., 2020; Lycett & Radwan, 2017; Meth et al., 2015; Venkatesh et al., 2017). The established propositions here are grounded in the individual articles in the cumulative dissertation. In addition, the status of relevant research from disciplines such as HCI and ISR is included here. However, it should be noted that when evaluating XAI and XUIs in specific domains, a broad range of evaluation characteristics exist from which practitioners or researchers can choose suitable ones (e.g., Arrieta et al., 2020; Haque et al., 2023; Minh et al., 2022; Vilone & Longo, 2021).

The first proposition derives from Article A4. Here, the constructs of perceived ease of use, perceived usefulness, and the intention to use were measured by a human-centered quantitative evaluation in the second design cycle. The statistical evaluation showed that the users rated these constructs very positively, related to a positive perception of the XUI. This proposition is further supported by the qualitative text fields used in the same evaluation, in which the users communicated a positive perception. The instantiated XUIs were also described as useful and easy-to-use in other human-centered evaluations, and users expressed their intention to use such XUIs and the associated applications, such as in A3. Consequently, the following proposition is established:

> ***Proposition 1:*** *Human-centered designed XUIs that consider the needs of the targeted stakeholders will lead to high acceptance levels, increasing the perceived ease of use, perceived usefulness, and intention to use.*

The second proposition can be traced back to Article A4. In the third design cycle, a between-subject experiment was carried out with two groups. One group interacted with a UI without XAI features and another with an XUI. The statistical evaluation of the data collected from the human-centered quantitative evaluation showed that the XUI was rated with a significantly lower perceived cognitive effort and the constructs of perceived informativeness, mental model (process), and trustworthiness with a significantly higher value. Consequently, the following proposition is established.

> ***Proposition 2:*** *Human-centered designed XUIs that consider the needs of the targeted stakeholders will positively influence the perceived cognitive effort, perceived informativeness, mental model, and trustworthiness when compared to UIs without explanations.*

In article A6, customization features for XUIs were conceptualized, designed, instantiated, and quantitatively evaluated in a human-centered manner. A between-subject experiment was carried out in the first design cycle. Two groups each interacted with the same XUI. They had to perform a classification task during the experiment. One group interacted with an XUI without customization features, and the other with an XUI with customization features. The perceived quality of the explanation was measured using the constructs of perceived interestingness, perceived informativeness, perceived interactivity, and satisfaction with the explanation. All constructs were rated significantly better by the group interacting with the XUI with customization features. In addition, the statistical evaluation showed that the group with access to the customization features achieved significantly higher task performance in the classification task, measured as accuracy. In addition, a second design cycle investigated the effects of customization features in XUIs on perceived interactivity, perceived interestingness, and perceived customization. The constructs of perceived interactivity and perceived customization predicted user engagement, which predicted user satisfaction. Consequently, the following three propositions are established:

> ***Proposition 3:*** *Human-centered designed XUIs that provide customization features will positively influence the perceived explanation quality measured through perceived interestingness, perceived informativeness, perceived interactivity, and satisfaction with the explanation compared to XUIs without customization features.*

> ***Proposition 4:*** *Human-centered designed XUIs that provide customization features will provide an engaging interaction experience predicted by perceived interactivity, perceived customization, and high satisfaction predicted through user engagement.*

> ***Proposition 5:*** *Human-centered designed XUIs that enable users to interact with the explanation, such as customization features, will positively influence task performance.*

The reusability of design principles is an essential property, which means that their evaluation is becoming more and more relevant. The relevance of the developed design principles was evaluated in articles A4, A5, and A6. The human-centered qualitative and quantitative evaluations of reusability were consistently rated very highly. In A4, a human-centered quantitative evaluation of the design principles was performed. In A5, a human-centered qualitative evaluation was performed first, followed by a human-centered quantitative evaluation. Both evaluations showed that the design principles were perceived as being very reusable. The design principles from A6 were also quantitatively evaluated in a human-centered manner and rated as very positive. Experienced individuals, including software developers, user interface designers, and user experience designers, were involved in all evaluations. Consequently, the following proposition is established:

> ***Proposition 6:*** *Design principles for human-centered XUIs that are formalized according to the status quo in DSR will lead to high levels of perceived reusability by potential implementers, measured through accessibility, importance, novelty and insightfulness, actability and guidance, and effectiveness.*

A total of six propositions were established, which are part of the ISDT for human-centered XUIs. These propositions are motivated by the human-centered qualitative and quantitative evaluations of the individual research projects that are part of the cumulative dissertation.

In addition to the design principles, the propositions brought together and set out in this subsection are another elementary part of the ISDT for human-centered XUIs. They provide information about effects that could be proven in individual research work through human-centered quantitative evaluations in surveys and experiments. Thus, the propositions provide insight into how users perceive the instantiated prescriptive design knowledge. The sixth component, the justificatory knowledge of the ISDT, is described in the following subsection.

4.3.6 Justificatory Knowledge

The last of the core components is the *sixth component*, where justificatory knowledge is presented. It refers to the underlying knowledge or integrated theories that provide a basis and explanation for the proposed design. Some information systems design theories use theories from other fields like social sciences (e.g., Diederich et al., 2020), rather broad descriptions of the theory in terms of research streams (e.g., Morana et al., 2019), or underlying knowledge that can be categorized as technical science-oriented (e.g., Lycett & Radwan, 2017).

An important part of mature design knowledge is kernel theory, also nearly synonymous with justificatory knowledge. It explains why the design works (Gregor & Jones, 2007). Such knowledge can be drawn from the knowledge base and should relate to the research goals (Gregor & Hevner, 2013). The different research projects in the cumulative dissertation draw their justificatory knowledge from a versatile knowledge base. Therefore, justificatory knowledge combines kernel theories, which can be theories from different fields, including the natural or social sciences, and justificatory knowledge from interdisciplinary research disciplines. The kernel theories include the Theory of Interactive Media Effects (Sundar et al., 2015), the Source Credibility Theory (Lowry et al., 2014), the Mental Model Theory (Vitharana et al., 2016), notions from the Technology Acceptance Model (Davis, 1989; Venkatesh et al., 2003), and aspects from research on decision strategies (Wang & Benbasat, 2009).

In addition, distinct knowledge sources have been used as justificatory knowledge to reference the design and explain how and why it works, which is an acceptable practice (Giessmann & Legner, 2016; Gregor & Jones, 2007; Ivari, 2020). The justificatory knowledge of the proposed ISDT for human-centered XUIs is composed of knowledge from research fields such as computer science where most of the XAI methods are developed (e.g., Bau et al., 2017; Huysmans et al., 2011, Kaur et al., 2020; Ribeiro et al., 2018; Vilone & Longo, 2022), research in HCI (e.g., Bradley et al., 2022; Cai et al., 2019; Eshan et al., 2022; Leichtmann et al., 2023; Sundar et al., 2015; Jorritsma et al., 2015), with a focus on XAI as well as XUIs (e.g., Adadi & Berrada, 2018; Arrieta et al., 2020; Dikmen & Burns, 2020; Haque et al., 2023; Kim et al., 2014; Minh et al., 2022) and ISR (e.g., Benbasat et al., 1987; Benbya et al., 2021; Berente et al., 2021; Dhaliwal & Benbasat, 1996; Meske et al., 2022; Wang & Benbasat, 2009). Consequently, the proposed ISDT for human-centered XUIs is grounded in kernel theories and interdisciplinary justificatory knowledge, all related to the design of (X)UIs and ISR.

This subsection presented the justificatory knowledge of the ISDT for human-centered XUIs, which is represented by the relevant knowledge base that informed the individual research projects and was supplemented by relevant and selected kernel theories. Together, they represent the justificatory knowledge of the ISDT for human-centered XUIs. In the following subsection, the seventh component of the ISDT is described: the implementation principles.

4.3.7 Principles of Implementation

The *six components* presented in the previous subsections represent the core components of an ISDT. Since the necessary knowledge is required to present the additional two components generated through the cumulative dissertation, the ISDT comprises all eight components. However, the last two components are additional, and it is accepted to only use the core components (e.g., Chanson et al., 2019) or all eight components (e.g., Kane et al., 2021). The *seventh component* provides principles of implementation (the causa efficiens) and describes concrete steps and processes necessary to implement the design theory in specific contexts (e.g., Avdiji et al., 2020; Mandviwalla, 2015). Consequently, the seventh component provides information on instantiating the prescriptive design knowledge and how the design is brought into being (Gregor & Jones, 2007). The description of the additional components is usually relatively short and refers to the instantiations that are part of the

DSR project or recommendations made for the implementation of the design theory (e.g., Diederich et al., 2020; Kane et al., 2021; Meth et al., 2015; Venkatesh et al., 2017).

The implementation process is flexible and can be adapted for any development project in practice or design-oriented research. Two things should be considered. On the one hand, a human-centered approach during development and evaluation is very valuable and recommended. On the other hand, the process should be structured iteratively, which is made possible, for example, by cyclical frameworks from DSR or iterative process models from software development such as SCRUM. The iterative approach has various advantages. Using human-centered approaches and methods, the target group of users can be involved early on, and their expectations, needs, requirements, backgrounds, and feedback can be used to optimize the developed artifact. Furthermore, the human-centered approach and the integration of relevant users can be used to identify which parts of the prescriptive design knowledge are to be adopted and whether it needs to be adapted or supplemented for a particular application. Furthermore, it can be determined in this way whether the expectations of the users and the defined goals or requirements are satisfied, and thus the desired effects occur.

Specific advice on possible approaches for implementation can be found in articles A3, A4, and A6, which are briefly summarized here. In A3, prescriptive design knowledge with an unobtrusive design was developed, instantiated, and human-centered qualitatively evaluated. Such XUIs could be part of a browser plugin or a web-based application. The XUI was designed using Adobe XD, focusing on design features that provide contextual information like similar topics or sources of disproof in the context of fake news detection. Moreover, the presentation of information in easy-to-comprehend visualizations such as rating scales was realized. In the research project presented in A4, mature XUIs were instantiated, targeting social media moderators as users in the context of explainable hate speech detection. Transfer learning was implemented and fine-tuned on a hate speech dataset. A post-hoc XAI method to generate explanations was used. Early versions of the XUI were designed using Adobe XD. Later, more mature versions of the XUI were developed using web technologies, including HTML, CSS-Bootstrap, Python Django, and JavaScript, thus making it an XUI that could be easily integrated with any social media platform provided through a web-based application. Using the same web technologies as in A4, the XUI in A6 was designed and instantiated. The focus lay on customization features in XUIs, which were designed using data visualization techniques and interactive XUI features. The interactive XUI used the same technology stack as the other web-based XUIs.

A notable aspect regarding the implementation principles of the ISDT for human-centered XUIs should be emphasized here. Although the prescriptive design knowledge contained in the ISDT was developed as part of the cumulative dissertation, mainly using web technologies with associated programming languages and frameworks, it can be transferred to almost any other form of application. The prescriptive design knowledge can be adopted or adapted for all conceivable applications containing XAI-based components. These include, for example, desktop applications, mobile applications, hybrid or cross-platform applications, and different operating systems. This leads to high flexibility, abstraction, and generalizability of prescriptive design knowledge, which are also important design knowledge characteristics and goals.

The seventh component, the principles of implementation, belongs to the last two components of the ISDT for human-centered XUIs, which belong to the optional components. In this subsection, possible approaches to implementation were shown as they were used in the different research projects that are part of the cumulative dissertation. However, it is essential to emphasize ,that, as already described in component four, artifact mutability, there is a large degree of freedom concerning the implementation of the ISDT. In the following subsection, the last component of the ISDT is presented.

4.3.8 Expository Instantiations

The last and *eighth components* focus on the expository instantiation of the design theory, which is mainly represented through a description of the instantiated prescriptive design knowledge (e.g., Coenen et al., 2018; Meth et al., 2015). Following Hevner et al. (2004), the instantiation of the design is essential for identifying potential problems in the design theory and demonstrating that the design is valuable in real-world settings. Gregor and Jones (2007) argue that instantiation is a possible component of an ISDT since it represents the theory or exposition.

No instantiation was carried out as part of the cumulative dissertation itself because different XUI design configurations were instantiated based on the prescriptive design knowledge, which was developed in the DSR projects and evaluated qualitatively or quantitatively in a human-centered manner. In addition to A1, in which the initial systematic literature search was carried out, A2 and A5 represent exceptions. Although prescriptive design knowledge was developed in these research projects, it was not instantiated and evaluated. This allows the following XUIs with their design configurations to be assigned to this component of the ISDT.

In A3, a prototypical XUI for the credibility assessment of online news content was instantiated. Inspired by the identified knowledge base and the online context, the XAI design elements in the XUI were designed as lightweight design elements that could be used, for example, by a browser plugin in any web browser and thus on a wide variety of platforms. The human-centered qualitative evaluation showed that users would prefer such systems to those they would have to install and set up separately or more complex information systems.

In A4, an XUI for hate speech detection on social platforms was developed to support moderators of social media platforms in detecting hateful content. The XUI was designed in a design cycle in an interactive mock-up and evaluated qualitatively in a human-centered manner. Based on the knowledge gained from the evaluation, the design was optimized and developed as a web-based XUI for the second and third design cycles and evaluated qualitatively and quantitatively in a human-centered manner. This decision was also influenced by the knowledge base and the knowledge gained from the exchange with moderators on social platforms. Because web pages or web apps are usually provided on social media platforms, the moderators also use web-based applications to moderate the platforms. For the sake of completeness, it should be mentioned that social media platforms can also be used through hybrid or native mobile apps. In this research project, however, no moderator said they use such mobile apps for moderating platforms.

Another instantiation was done in A6, developing XUIs that included previously conceptualized and developed customization features. The XUIs for the first and second design cycles were each provided as a web-based application. In the first and second design cycles, different focuses were set on application domains and data types to examine the concept of customization features in XUIs in depth. The human-centered quantitative evaluations showed that the customization features in XUIs have the hypothesized effect and thus positively influenced the interaction experience for the users.

This subsection presents the last component of the ISDT for human-centered XUIs. It was discussed how the expository instantiations are represented by the different XUIs with individual design configurations in the individual research papers, which are part of the cumulative dissertation. The following subsection presents a summary of the ISDT for human-centered XUIs. The individual components are composed and summarized in tabular form, as introduced by Gregor and Jones (2007).

4.3.9 The Information Systems Design Theory for Human-Centered XUIs

In this last subsection, the previously individually developed components of an ISDT based on Gregor and Jones (2007) are transferred to the established tabular form and summarized. The tabular representation names the component type and briefly summarizes the component. Table 13 presents the proposed ISDT for human-centered XUIs.

<table>
<tr><th>Component Type</th><th>Component</th></tr>
<tr><td>1) Purpose and scope (the causa finalis)</td><td>The cumulative dissertation explores the design and perception of XUIs in various application domains. Through systematic literature searches, it becomes evident that comprehensive design knowledge for XAI systems, specifically XUIs, is limited and not easily generalizable. The research highlights the need for further investigation into how users perceive different XUI designs and configurations, considering their diverse experiences, expectations, and knowledge.

The research projects within the cumulative dissertation address these gaps by developing an ISDT for human-centered XUIs. This design theory provides essential insights into designing effective XUIs across diverse contexts, enabling an understanding of different design configurations of XUIs and their impacts on the user. The ISDT contributes significantly to the existing research landscape by offering adaptable design knowledge applicable to various XAI-based decision support systems in various fields.

However, it is important to stress that while design theory can be broadly applied, a human-centered approach remains crucial. Tailoring design configurations to align with specific goals, needs, expectations, and requirements of the targeted users is essential to ensuring positive interaction experiences and optimal support when working with XAI systems and their associated XUIs.</td></tr>
<tr><td>2) Constructs (the causa materialis)</td><td>A large selection of constructs and notions is documented in the body of literature on XAI, which can be influenced by the design of XUIs. Based on the status quo of research on XAI and theoretical groundings in theories such as the Theory of Interactive Media Effects, Source Credibility Theory, Technology Acceptance Model, and decision strategies, suitable constructs were identified and investigated during the human-centered evaluations in the different research projects that are part of the cumulative dissertation.

Most constructs with a focus on the perception of the instantiated prescriptive design knowledge were evaluated through human-centered quantitative research methods:

▪ Perceived ease of use, perceived usefulness, and intention to use (Technology Acceptance Model)
▪ Perceived cognitive effort (Decision Strategies)
▪ Mental Model – Process (Mental Model Theory)
▪ Perceived Informativeness, perceived interestingness, perceived interactivity, trustworthiness, and satisfaction with the explanation (goals and notions of XAI, shared with the Theory of Interactive Media Effects)
▪ Perceived customization, user engagement, and satisfaction (goals and notions of XAI, shared with the Theory of Interactive Media Effects)
▪ Task performance (Measured as Accuracy)

Constructs used to evaluate the reusability of design principles have been used in human-centered qualitative and quantitative evaluations in the research, which is part of the cumulative dissertation:</td></tr>
</table>

	▪ Accessibility, importance, novelty and insightfulness, actability and guidance, and effectiveness (reusability evaluation of design principles)
3) Principle of form and function (the causa formalis)	**Design principle of performance communication** To allow users (users) to correctly comprehend the recommendation of an XAI-based system (aim) in the context of decision-making supported by a human-centered XAI-based system, the system should communicate the performance of the underlying AI in an easy-to-understand way in the XUI, which requires no prior knowledge or expertise in the context of AI, so that the performance communication in XUIs enables a diverse user base to understand and interpret the AI's performance while requiring an appropriate cognitive effort and supports the development of correct mental representations. **Design principle of human-centered selection of XAI methods** To allow designers, developers, practitioners, and researchers (enactors) to select either the correct XAI method or mix of XAI methods (aim) for their targeted user base (users) in the context of decision-making supported by a human-centered XAI-based system, they have to involve the users during the process of requirements elicitation, design, and evaluation of XUIs within an iterative process so that they can take advantage of human-centeredness and identify individual information needs, expectations, bias, and acceptance barriers to consider these aspects during the design process. **Design principle of case- and output-related comparison** To support users (users) in developing a holistic understanding of the XAI-based output in the scope of the task to be performed (aim) in the context of decision-making supported by a human-centered XAI-based system, the system should provide features that enable case- and output-related comparisons within the XUI, so that the users can develop an understanding of how the XAI-based system has performed on different cases, gain new insights, and learn more about the functionality of the system. **Design principle of contextual information** To enable users (users) to familiarize themselves with all necessary information relevant to the task at hand (aim) in the context of decision-making supported by a human-centered XAI-based system, the system should provide contextual information that is relevant for the task to be performed in the XUI, so that users can make informed decisions and either accept or reject the XAI-based recommendation. **Design principle of assistance with interpretation** To support users (users) in correctly interpreting the included information in the provided explanation (aim), in the context of decision-making supported by a human-centered XAI-based system, the system should provide appropriate assistance depending on the XAI method or mix of XAI methods being used in the XUI, so that users can assess the soundness, completeness, and faithfulness of explanations or diagnose and correct flawed XAI-based outputs. **Design principle of cosmetic customization**

	To allow users (users) to interact with AI-generated explanations for adjusting their visual representation from a set of pre-defined alternatives (aim) in the context of decision-making supported by a human-centered XAI-based system, the system should provide easy-to-use settings for users to customize the visual representation according to their individual needs as well as preferences so that the cosmetic customization in XUIs can provide a high degree of interactivity, interestingness, informativeness, satisfy users, and ultimately improve the perceived explanation quality. **Design principle of functional customization** To allow users (users) to interact with AI-generated explanations for adjusting the number of relevant features displayed in the XUI from a set of pre-defined alternatives (aim) in the context of decision-making supported by a human-centered XAI-based system, the system should provide easy-to-use settings for users to customize the scope of relevant features displayed in the XUI so that the functional customization in XUIs can provide a high degree of interactivity, interestingness, informativeness, satisfy users, and ultimately improve the perceived explanation quality. **Design principle of XAI method customization** To allow users (users) to interact with AI-generated explanations for adjusting the XAI method displayed in the XUI from a set of pre-defined alternatives (aim) in the context of decision-making supported by a human-centered XAI-based system, the system should provide easy-to-use settings for users to customize the presented explanation displayed in the XUI so that the XAI method customization in XUIs can provide an engaging and satisfying user experience.
4) Artifact mutability	The presented design theory is adaptable, not requiring complete adoption. Taking a human-centered approach and involving stakeholders early in XUI development allows for essential requirements and needs to be derived, enabling selective adoption of suitable design principles. Stakeholders can contribute throughout various phases, offering valuable insights to identify appropriate design knowledge. This knowledge can then be incorporated into the existing ISDT for human-centered XUIs, which do not specify particular XAI methods, granting flexibility in method selection as the field evolves. The design theory's mutability extends to application types, platforms, and data processing, making it versatile for diverse contexts. Practitioners and researchers can adopt elements from the ISDT that suit specific projects. Exemplary design features were established during instantiation in the research projects part of the cumulative dissertation, allowing adjustments based on user or project requirements. The ISDT provides specific design principles for XUI design, offering room for adaptation and customization and emphasizing the importance of a human-centered perspective. This perspective aids in identifying pertinent design theory elements for projects and determining necessary adjustments. Consequently, different technologies and approaches can be used to adapt or expand prescriptive knowledge for specific projects and targeted stakeholder groups.

5) Testable propositions	**Proposition 1** Human-centered designed XUIs that consider the needs of the targeted stakeholders will lead to high acceptance levels, increasing the perceived ease of use, perceived usefulness, and intention to use. **Proposition 2** Human-centered designed XUIs that consider the needs of the targeted stakeholders will positively influence the perceived cognitive effort, perceived informativeness, mental model (process), and trustworthiness when compared to UIs without explanations. **Proposition 3** Human-centered designed XUIs that provide customization features will positively influence the perceived explanation quality measured through perceived interestingness, perceived informativeness, perceived interactivity, and satisfaction with the explanation compared to XUIs without customization features. **Proposition 4** Human-centered designed XUIs that provide customization features will provide an engaging interaction experience predicted by perceived interactivity, perceived customization, and high satisfaction predicted through user engagement. **Proposition 5** Human-centered designed XUIs that enable users to interact with the explanation, such as customization features, will positively influence task performance. **Proposition 6** Design principles for human-centered XUIs that are formalized according to the status quo in DSR will lead to high levels of perceived reusability by potential implementers, measured through accessibility, importance, novelty and insightfulness, actability and guidance, and effectiveness.
6) Justificatory knowledge	Interdisciplinary justificatory knowledge and kernel theories from a versatile knowledge base influenced the ISDT for human-centered XUIs. Kernel's theories include the Theory of Interactive Media Effects, Source Credibility Theory, Mental Model Theory, notions from the Technology Acceptance Model, and dimensions of research on decision strategies. Justificatory knowledge was drawn from the research disciplines of computer science, ISR, and research in HCI, with a focus on XAI and XUI.
7) Principles of implementation (the causa efficiens)	The flexible implementation process can suit various practical development or research projects. It emphasizes a human-centered approach and iterative structuring facilitated by frameworks like DSR or SCRUM. Iterative procedures enable early user involvement, optimizing artifacts through their feedback. This approach helps identify which parts of the design knowledge to adopt, adapt, or supplement for specific applications, ensuring user expectations, goals, and desired effects are met.

	Although the prescriptive design knowledge contained in the ISDT for human-centered XUIs was instantiated mainly using web technologies with associated programming languages and frameworks, it can be transferred to almost any other form of application. The prescriptive design knowledge can be adopted or adapted for all conceivable applications where XAI-based decision support systems can be implemented. These include, for example, desktop applications, mobile applications, hybrid or cross-platform applications, and different operating systems. This leads to high flexibility, abstraction, and generalizability of prescriptive design knowledge.
8) Expository instantiation	XUIs with different design configurations were instantiated in the individual research projects that are part of the cumulative dissertation. In A3 a, a lightweight XUI for online news credibility assessment was instantiated, which was developed as a design for browser plugins that can be used through the web and across different platforms. In A4, an XUI aiding social media moderators in detecting hate speech was developed. Initially, an interactive mock-up was developed. In the later design cycles of the DSR project, the XUI design was optimized and instantiated as a web-based application. In A6, an XUI with customization features was instantiated and provided as a web-based application.

Table 13. An ISDT for human-centered XUIs (based on Gregor & Jones, 2007, p. 322).

After the individual components of the ISDT for human-centered XUIs were built in the last subsection, they were summarized in this last subsection as the core contribution of the cumulative dissertation. In the following section, the results of the cumulative dissertation are discussed, the theoretical and practical implications are addressed, and the limitations and potential for future research are reflected.

5 Discussion

5.1 On the Human-Centered Design of XUIs

A major point of discussion within the DSR community lies in the effective presentation and formalization of design knowledge in a way that makes it accessible for future design-oriented research as well as practice (Gregor & Jones, 2007; Gregor & Hevner, 2013; Kuechler & Vaishnavi, 2012). The generated design knowledge was formalized as an ISDT, according to Gregor and Jones (2007). It is not only an effective way to communicate the developed design knowledge, but it is also well-established in top ISR journals such as the Journal of the Association for Information Systems, European Journal of Information Systems, Information Systems Journal, or the MIS Quarterly (e.g., Avdiji et al., 2020; Coenen et al., 2018; Giessmann & Legner, 2016; Venkatesh et al., 2017). This means that the developed ISDT for human-centered XUIs follows the status quo on DSR within ISR. Moreover, the overarching research contribution of the cumulative dissertation adheres to the status quo on conducting and presenting DSR by following well-established guidelines and frameworks (vom Brocke & Maedche, 2019; Gregor & Hevner, 2013; Gregor et al., 2020; Ivari et al., 2021; Kuechler & Vaishnavi, 2012; Peffers et al., 2007; Venable et al., 2016). In doing so, an ISDT focuses on the solution space, which was also established (Avdiji et al., 2020; Meth et al., 2015; Morana et al., 2019).

Researchers from different disciplines have provided valuable contributions with a focus on human-centered design of XUIs (e.g., Leichtmann et al., 2023; Nazar et al., 2021; Schoonderwoerd et al., 2021;

Wells & Bednarz, 2021) and human-centered evaluations (e.g., Anjomshoae et al., 2019; Dosilovic et al., 2018; Vilone & Longo, 2021). Nevertheless, the research stream of XAI remains characterized by a lack of generalizable design knowledge and user studies (Nazar et al., 2021; van der Waa et al., 2021; Wells & Bednarz, 2021; Sperrle et al., 2021). This was one of the main motivations that ran through all the articles in the cumulative dissertation. Therefore, the aim of the proposed ISDT for human-centered XUIs is to provide abstract, coherent prescriptive design knowledge that summarizes the relevant components to develop human-centered XUIs with positive effects regarding the perception of users and their interaction experience (Gregor, 2006; Gregor & Hevner, 2013; Gregor & Jones, 2007).

Consolidating the developed knowledge in the individual research projects as part of the cumulative dissertation answers the overarching RQ1. Consequently, an ISDT for human-centered XUIs that contains all the components presented by Gregor and Jones (2007) is proposed. The subordinate RQs are addressed to a varying degree through the individual research articles included in the cumulative dissertation. Some research focused exclusively on developing prescriptive design knowledge for XUIs, whereas others aimed to develop and evaluate the instantiated XUIs. In any case, a human-centered approach was followed by involving relevant stakeholders during the elicitation and derivation of requirements, qualitative evaluations of prescriptive design knowledge and instantiated artifacts, or human-centered quantitative evaluations of prescriptive design knowledge and instantiated artifacts. The human-centered qualitative approaches mainly served to understand the needs and expectations of relevant stakeholders within the application domain. In addition, it was a valuable approach to obtain feedback for further optimization of the prescriptive design knowledge (A5) or to optimize the instantiated artifact (A4). Similarly, the human-centered quantitative research approaches evaluated the reusability of the proposed prescriptive design knowledge (A4, A5) or the instantiated artifacts by integrating theoretical concepts represented through the justificatory knowledge and kernel theories and the goals of XAI (A4, A6). Consequently, the generated contributions, supplemented by the status quo of research on XUIs, flowed into the proposed ISDT for human-centered XUIs.

The consolidated and presented design principles of the ISDT for human-centered XUIs address RQ1.1. The developed design principles are essential to the ISDT and aim to capture and communicate prescriptive design knowledge, their overarching objective (Chandra Kruse et al., 2015; Gregor et al., 2020). Moreover, they intend to provide actionable knowledge for designers who aim to build new versions of related systems (Chandra Kruse et al., 2022). The design principles are abstract prescriptive statements for design solutions and are part of the solution space. In contrast, the purpose and scope of the ISDT can be understood as a means to represent the problem space (Avdiji et al., 2020). Through the evaluation of the prescriptive design knowledge or instantiated artifacts, a connection between both spaces is established (vom Brocke & Maedche, 2019). This connection and the consideration of both the problem and solution spaces are not established in the literature on human-centered XUIs. Although there is limited prescriptive design knowledge, the proposed design principles have novel characteristics. Researchers from disciplines like HCI or computer science also develop XUIs but do not formalize the prescriptive design knowledge in such an accessible form as can be achieved through DSR (e.g., Alufaisan et al., 2021; Dikmen & Burns, 2020; Habers et al., 2010; Kim et al., 2014; Lim et al., 2009). Moreover, the resulting design knowledge from such disciplines is rarely evaluated by involving potential implementers.

Existing research that proposes prescriptive design knowledge also has limitations, which the here-presented ISDT for human-centered XUIs tries to overcome. For example, Schoonderwoerd et al. (2021) proposed a set of design patterns for UIs of clinical decision support systems in medical diagnosis. They strongly focus on clinical decision support systems for child health, and whether the proposed design knowledge is generalizable and applicable for other types of clinical decision support systems or even other application domains remains unanswered. In addition, the design knowledge was not evaluated with potential implementers, which leaves open the interrogation of how useable the introduced design knowledge can be in practice. Herm et al. (2022) developed a nascent design theory for explainable intelligent systems, one of the few research contributions closer to the ISDT for human-centered XUIs introduced in this cumulative dissertation. However, due to the structure of the DSR project by Herm et al. (2022), they only investigated one application domain and relied solely on qualitative methods for their evaluation. In addition, their broader scope on the design of explainable intelligent systems is not comparable to the in-depth focus on the design of human-centered XUIs, as achieved by consolidating the articles included in the cumulative dissertation. Another example stems from Barda et al. (2020), who introduced a qualitative research framework for designing user-centered displays of explanations in healthcare. Their research emphasized the relevance of including relevant stakeholders in the design of such displays. However, they did not introduce prescriptive design knowledge per se that could guide potential implementers, as design knowledge formalized according to the status quo of DSR. This distinguishes the ISDT for human-centered XUIs presented here strongly from existing research since a broad range of application domains, versatile stakeholder groups, and perspectives were considered. Novel design characteristics and configurations were introduced, including transfer learning for hate speech detection (A4), a multi-stakeholder perspective for human-centered XUIs of medical decision support systems (A5), or the concept of customization features in human-centered XUIs (A6). All these design configurations generated novel and valuable knowledge for designing human-centered XUIs and their effects on relevant stakeholders.

DSR projects can generate design knowledge in different forms, and design principles are am established way to formalize such knowledge (Gregor & Hevner, 2013; vom Brocke & Maedche, 2019; vom Brocke et al., 2020). The design knowledge should be relevant to a scientific audience and practitioners (Hevner et al., 2004). Consequently, design principles help researchers formalize and communicate research outcomes, build a cumulative body of knowledge, and ideally support potential implementers in designing similar artifacts (Chandra Kruse et al., 2022; Gregor et al., 2020). This makes the evaluation of instantiated artifacts but also (nascent) design theories, such as design principles, a key activity in DSR since it is an opportunity to receive feedback for optimization and assures rigor (Gregor & Hevner, 2013; Venable et al., 2016). Design principles can be evaluated with the targeted stakeholders—the potential implementer, for example—by measuring how actionable they are (Chandra Kruse et al., 2015). Ivari et al. (2021) introduced a framework for minimum reusability evaluation, which can be used to evaluate and ensure the practical relevance of design principles. Such an evaluation of the reusability and potential is not always conducted, even in DSR projects published in top IS journals (e.g., Germonprez et al., 2016; Giessmann & Legner, 2016; Seidel et al., 2015). In suitable DSR projects, the minimum reusability evaluation was applied, which consists of the following dimensions (Ivari et al., 2021): accessibility, importance, novelty and insightfulness, actability and guidance, and effectiveness.

This approach to evaluating design principles was used in A4, A5, and A6. In article A4, the design principles were quantitatively evaluated with 80 experienced software developers who rated all the reusability dimensions positively. In addition, 66 of 80 participants said they would adapt the design principles for a suitable software development project. A slightly different approach was followed for the reusability evaluation of the design principles in A5. It started with a human-centered qualitative reusability evaluation by conducting semi-structured interviews using the template of Ivari et al. (2021) as a basis for the interview guide with four experienced software developers from versatile domains, including web development, mobile app development, and machine learning engineers. The design principles were perceived positively, and only a little potential for optimization was uncovered. After minimum revisions of the design principles, the design principles were human-centered quantitatively evaluated with 66 participants. The positive perception from the qualitative evaluation was confirmed. A total of 51 participants stated that they would adapt the design principles for a suitable software development project, and 54 participants would recommend them to a colleague for a suitable project. An evaluation of the reusability of design principles was carried out in A6 with the involvement of experienced user interface and user experience designers. A total of 65 subjects took part, and the answers from 61 subjects were included in the statistical analysis. This practitioner positively rated the design principles. In this survey, 52 people said they would use the design principles for a suitable project, and 53 said they would recommend them to colleagues.

RQ1.2 is addressed by the propositions of the ISDT for human-centered XUIs. They represent a consolidation of the insights generated through the human-centered evaluations in the individual research projects that are part of the cumulative dissertation. Since explanations are only effective when they help the targeted stakeholders build appropriate trust, correct mental representations, and support understanding as well as comprehension, human-centered evaluations are highly relevant (Adadi & Berrada, 2018; Gunning & Aha, 2019; Vilone & Longo, 2021). Despite active research from different disciplines with a focus on evaluating and conducting quantitative experiments in the context of XAI (e.g., Bau et al., 2017; Huysmans et al., 2011; Kaur et al., 2020; Ribeiro et al., 2018; Vilone & Longo, 2022; Weitz et al., 2021), this research stream is still characterized by a lack of such human-centered evaluations (Anjomshoae et al., 2019; Dosilovic et al., 2018; Gilpin et al., 2018; Haque et al., 2023; van der Waa et al., 2021; Wang et al., 2019; Wells & Bednarz, 2021). The need for human-centered explanations is motivated and justified by versatile reasons, including the presentation formats of explanations that can have varying effects on individual stakeholder groups (Langer et al., 2021; Meske et al., 2022), the different information needs, expertise, and bias employed when working with explanations (Arrieta et al., 2020; Miller, 2019), or the sheer fact that there is no one-size-fits-all approach to designing explanations (Sokol & Flach, 2020). Moreover, various notions of explainability, XAI goals, and assessment methods exist (Ali et al., 2023; Vilone & Longo, 2021; Haque et al., 2023; Minh et al., 2022). This allows for almost infinite experiment designs with multiple foci and the integration of different theories as well as research disciplines (Barda et al., 2020; Fiok et al., 2022; Fügener et al., 2022; Leichtmann et al., 2023; Schoonderwoerd et al., 2021; Senoner et al., 2022). During the evaluations of the individual DSR projects, the opportunity to investigate such relevant notions of XAI and the perception of the instantiated prescriptive design knowledge by relevant stakeholders was utilized.

In A3, for example, an early version of an XUI prototype to support users in assessing the credibility of news articles in an online context was investigated. Three overarching goals were followed: supporting users in the source credibility assessment process, providing an unobtrusive design for the credibility assessment of the news content, and enabling the discovery of similar content outside the filter bubble in combination with sources of disproof in the case of detected fake news articles. The prototype was designed with Adobe XD software to achieve a realistic design and provide basic interactive functionalities. The qualitative evaluation showed that the 13 participants liked the proposed solution for assessing the credibility of news articles in an online context and described it as useful. The prescriptive design knowledge and the interview guide for the evaluation were grounded in the Source Credibility Theory. Participants said they would prefer a lightweight UI provided through a browser plugin over an additional information system. Moreover, since the credibility assessment was based on AI, they demand explanations to comprehend the reasons for the AI recommendation. Additionally, the participants liked rating the associated information sources and rather simple visualizations in rating scales since they are easy to understand.

A4 was a larger DSR project where an XUI was developed for social media moderators to support them in detecting hateful content. Three consecutive design cycles were run through, where transfer learning for hate speech detection and using local explanations were implemented. Again, Adobe XD was used to design the XUI for the first and second design cycles. In the first design cycle, the XUI was human-centered and qualitatively evaluated with 11 experienced moderators of social platforms. They provided valuable feedback for further optimization and communicated a positive sentiment toward the XUI. Moreover, the participants stated that such an XUI and the associated information systems would benefit their work as social media moderators. After working through the second design cycle and optimizing the XUI based on the feedback received from the evaluation of the first design cycle, the second evaluation was of a human-centered quantitative nature.

The second evaluation aimed to validate the utility provided by the XUI. Therefore, relevant dimensions from the Technology Acceptance Model were measured, including the perceived ease of use, usefulness, and intention to use. In addition, qualitative feedback was collected through text fields in the survey and analyzed using thematic analysis (Braun & Clarke, 2006). On the one hand, the analysis resulted in high measurements for the utilized constructs from the Technology Acceptance Model. On the other hand, very helpful and constructive feedback for further optimization of the XUI was collected. Overall, 190 participants were recruited for the second evaluation. After optimizing the XUI and implementing it as a web-based artifact using HTML, CSS-Bootstrap, Python Django, and JavaScript, a between-subject experiment was conducted as a form of a human-centered quantitative evaluation. With 360 participants divided into two groups, an AI version of the UI without XAI features was compared to the XUI version. It was hypothesized that the XAI component will lead to lower perceived cognitive effort, higher perceived informativeness, mental model (process), and trustworthiness. The statistical analysis supported all the hypotheses, showing the significance of explanations for using social media moderators in detecting hateful content. Moreover, this also plays into the mutability of this specific artifact; the social moderators mentioned several times that it would be beneficial if hate speech and other undesirable forms of content, such as sexism or racism, were detected.

In A6, the concept of customization was conceptualized for the design of XUIs. Different customizations were identified and operationalized: cosmetic, functional, and XAI method customization. Therefore, design requirements and design principles with associated design features were derived. The overarching goal was to improve the perceived explanation quality, operationalized through the constructs of perceived interestingness, informativeness, interactivity, and satisfaction with the explanations. The prescriptive design knowledge was instantiated as a web-based XUI using HTML, CSS-Bootstrap, Python Django, and JavaScript. To evaluate the effectiveness of the proposed design knowledge for customization features in XUIs, a between-subject experiment with a total of 180 participants was conducted. The participants were equally divided into two groups, where one group was provided with a baseline XUI with no customization features and another with customization features. The experimental task was to predict the attrition risk of employees, and all recruited participants had experience in human resources and hiring employees.

Not only did the group that was provided with the XUI with customization features have significantly higher measurements for all constructs, but they also achieved significantly higher task performance during the experimental classification task. Interestingly, the cosmetic customization feature was more often used than the functional one. Consequently, the proposed and instantiated prescriptive design knowledge for customization features in XUIs led to a significantly higher perceived explanation quality. In the second design cycle, as part of A6, user engagement and satisfaction were examined in more detail and explored in the context of XUIs with customization features. During the experiment, participants had to classify a salary for an unknown applicant in a binary classification task supported by an XUI with customization features. It was measured and showed how perceived interactivity and perceived customization could predict user engagement, which could not be proven for perceived interestingness. User engagement, for its part, positively predicted satisfaction. In the second design cycle, the XAI method customization feature was significantly more frequently used than the cosmetic and functional customization features.

In summary, the cumulative dissertation provides valuable research contributions and insights into the design of human-centered XUIs and how relevant stakeholders perceive them. The next subsection presents summarizes the theoretical contributions and implications.

5.2 Theoretical Contributions and Implications

The cumulative dissertation has several critical theoretical contributions and implications for research. Based on the DSR knowledge contribution framework introduced by Gregor and Hevner (2013), the ISDT for human-centered XUIs can be classified as an improvement. Such an improvement is achieved by developing new solutions for known problems, encompassing research opportunities and knowledge contributions. The design knowledge underlying the ISDT comprises design knowledge and configurations for XUIs in different application contexts where either no prescriptive design knowledge existed, or novel design components were introduced. As mentioned in the summary of the findings, research exists that guides the design of XAI systems and XUIs, such as the qualitative research framework for the design of user-centered displays in healthcare (Barda et al., 2020), design patterns for explanations of clinical decision support systems for child health (Schoonderwoerd et al., 2021), a nascent design theory for explainable intelligent systems (Herm et al., 2022), or design principles for interactive XUIs (Chromik & Butz, 2021). Although these contributions contain relevant design knowledge, the ISDT developed here addresses the limitations of the previously mentioned work. Some existing prescriptive design knowledge was either not evaluated with potential implementers, was not developed using a human-centered approach, was not instantiated in an evaluated artifact, or had a narrow focus on one particular type of information system. The proposed ISDT, on the contrary, is based on design knowledge that, in every case, was developed with a human-centered approach.

Relevant stakeholders were either involved during the requirement elicitation process, when evaluating the instantiated prescriptive design knowledge, or in evaluating the prescriptive design knowledge with potential implementers. Moreover, the design knowledge was developed in different application contexts and consolidated into the ISDT. In addition, the ISDT included actionable design principles that could be used for future research (Chandra Kruse et al., 2015; 2022; Gregor & Hevner, 2013; Ivari et al., 2021). In contrast to existing prescriptive design knowledge for XUIs, the design principles included in the ISDT do not only focus on the explanation itself but also on design features that can positively influence their perception by users, such as valuable contextual information, assistance for the interpretation of explanations, or customization features in XUIs.

Besides the DSR contributions in the form of prescriptive design knowledge that were ultimately consolidated into the ISDT for human-centered XUIs, the individual research articles have provided valuable insights regarding the perception of XUIs by relevant stakeholders. This was achieved through qualitative and quantitative research methods. On the one side, the qualitative methods were beneficial in understanding the problem space and environment of the DSR project (Hevner et al., 2004; Meth et al., 2015; Nazar et al., 2021). Therefore, they supported the individual DSR projects regarding identifying challenges for eliciting requirements and also provided valuable feedback for optimizing the individual XUIs (A2, A3, and A4). Moreover, through the human-centered qualitative research methods in the form of semi-structured interviews, the stakeholders, such as laymen, regular end users, experts, or professionals from different domains, emphasized their need for explanations when they use AI systems for decision support. Therefore, the excellent symbiosis of the human-centered approach with DSR made gaining such highly relevant insights indispensable. Additionally, human-centered quantitative methods such as surveys and between-subject experiments are well-established human-centered approaches (Vilone & Longo, 2021; Nazar et al., 2021). Interestingly, most experiments align with early research on explainability in the ISR discipline. For example, early related research also found that explainability can positively influence perceived usefulness, ease of use, satisfaction, or trust (Dhaliwal & Benbasat, 1996; Mao & Benbasat, 2000; Ye & Johnson, 1995). Similar findings were generated during research endeavors, such as A4 or A6. This is not self-evident since AI is a constantly evolving technology (Berente et al., 2021) and generates an entirely different level of attention than early knowledge-based or expert systems (Maslej et al., 2023). In addition, modern (X)AI penetrates far more domains with significantly more far-reaching consequences for industry, society, or individuals (Adadi & Berrada, 2018; Arrieta et al., 2020; Schneiderman, 2020a; 2020b). Despite interdisciplinary and active research focusing on XAI, there are still so many research opportunities and blind spots in this realm, including ensuring and evaluating appropriate levels of trust, also called calibrated trust (Naiseh et al., 2023). In addition, post-hoc explainability methods are also not free from critique, especially in high-stake scenarios, where the call for inherently explainable models gains relevance (Rudin, 2019).

After discussing theoretical contributions and implications, these points will be discussed from a practical perspective in the next subsection.

5.3 Practical Contributions and Design Implications

Similar to the theoretical contributions and research implications, the practical contributions and design implications are manifold as well. The cumulative dissertation contributes to the knowledge base for human-centered XUI design through the introduced design knowledge and insights generated by the evaluations of the instantiated XUIs and the prescriptive design knowledge. Therefore, relevant knowledge for the design of human-centered XUIs extends the current knowledge base (e.g., Barda et al., 2020; Chromik & Butz, 2021; Kulesza et al., 2015; Leichtmann et al., 2023; Nazar et al., 2021; Schoonderwoerd et al., 2021; Zacharias et al., 2022). In addition, through the human-centered evaluations, practitioners and researchers who want to adapt design knowledge from the proposed ISDT are provided with empirical evidence of the effects of individual design principles and design configurations. Therefore, the opportunity arises for specific design principles to be used and integrated into existing XAI systems or XUIs for varying application contexts.

While XAI is an interdisciplinary research subject, artifacts such as XAI methods, systems, and XUIs are primarily developed by computer science, HCI, ISR, and practitioners. All these parties can benefit from the ISDT, which highlights a human-centered perspective. This human-centered approach, in combination with the DSR paradigm, demonstrates how valuable and necessary it is to involve relevant stakeholders in the generation and testing cycle (Hevner et al., 2004). Many stakeholders with highly individual information needs, expectations, and demands can be potential users targeted by an XAI system (Arrieta et al., 2020; Langer et al., 2021). Therefore, design characteristics such as the established concept of personalized XAI (Meske et al., 2022) or the proposed design for customization features in XUIs (A6) can be of high value to reach users with different levels of expertise and desires, even when diverse stakeholders from the user base. In addition, integrating XAI can be essential to achieving high acceptance levels and positively influencing the perception of XUIs (A4). Consequently, the ISDT for human-centered XUIs emphasizes the added value of human-centeredness.

It is also important to emphasize that the proposed ISDT here will not be a final design solution for every XAI system. Since XAI methods, XAI systems, application contexts, and stakeholders can largely differ, it cannot be a one-size-fits-all solution, which is in line with prior research (Adadi & Berrada, 2018; Arrieta et al., 2020; Haque et al., 2023; Langer et al., 2021; Sokol & Flach, 2020). Therefore, the ISDT should be viewed as lessons learned, insights generated, and knowledge contributions produced during individual DSR projects (Gregor & Hevner, 2013; Gregor & Jones, 2007).

The next and final subsection of the discussion reflects the limitations of the cumulative dissertation and the research opportunities that arise from it in the future.

5.4 Limitations and Future Research

The individual research projects adhered to well-established DSR guidelines (Gregor et al., 2020; Gregor & Hevner, 2013; Hevner et al., 2004; Ivari et al., 2021; Kuechler & Vaishnavi, 2012; Peffers et al., 2007; Venable et al., 2016). In addition, well-established methods for the conduct of literature reviews (vom Brocke et al.; 2015; Schryen et al., 2017; Webster & Watson, 2002), qualitative methods such as semi-structured interviews and associated data analysis approaches (Bagayogo et al., 2014; Braun & Clarke, 2006; Patton, 2002), and quantitative research methods, including surveys and online experiments (Fink, 2022; Galliers & Land, 1987; Karahanna et al., 2018) were applied during the individual research projects. However, like other research endeavors, the individual research articles that are part of the cumulative dissertation and, hence, the cumulative dissertation itself are not free from limitations. Therefore, I want to emphasize four distinctive limitations through which many future research possibilities arise simultaneously.

The first limitation focuses on the maturity of the instantiated XUIs. The effort required and, thus, the maturity of the instantiated XUIs usually depended on whether the research project was planned as a conference paper or a journal article. For example, in A3, A5, or A6, one full design cycle was run through since they were conference articles. Therefore, prescriptive design knowledge was developed, usually instantiated, and evaluated, whereas A4 was a journal article. Here, the opportunity was taken to investigate the human-centered design of XUIs in the context of explainable hate speech detection more extensively. For all instantiated XUIs, established tools like Adobe XD or web development technologies such as HTML, CSS-Bootstrap, Python Django, and JavaScript were used. In any case, it was aimed at realistic and interactive XUIs. However, due to the detached nature of the instantiated artifacts, the opportunity arises for the prescriptive design knowledge to be adopted for XUIs in real-world settings and information systems used in versatile areas. This ultimately represents the second limitation of this cumulative dissertation. This limitation is recognized in the XAI literature: the need for more field studies and experiments in real-world settings and, hence, more realistic evaluations (Ali et al., 2023; Leichtmann et al., 2023; Sokol & Flach, 2020). A limitation of all the instantiated XUIs is that they are still rather artificial and provided in a controlled environment. This provided us with the basis to design rigorous and controlled evaluations, but at the same time, it took away the possibility of evaluations in more realistic settings (Karahanna et al., 2018). This limitation could be used as an opportunity for future research. Here, researchers and practitioners are invited to adopt suitable design knowledge from the ISDT for human-centered XUIs, using and evaluating it in real-world settings.

Another angle to describe limitations is stakeholder involvement. No effort was spared to involve the relevant stakeholders in individual research projects. For example, in A2, multimodal mobility platforms for travel planning were at the center of attention, and relevant stakeholders, such as regular end users with versatile backgrounds and experts, were recruited. In A3, the design of XUIs for services that support the credibility assessment process of news articles in an online context was investigated. Only stakeholders consuming their news content via online and social media platforms were involved. In the research project focusing on the design of XUIs for explainable hate speech detection on social media platforms presented in A4, many social media moderators were involved in the human-centered evaluation of the XUIs. In addition, experienced software developers were recruited to evaluate the prescriptive design knowledge in a human-centered manner. Similarly, in A5, where the focus was on the design of XUIs for medical decision support systems, the relevant stakeholder perspectives of physicians and patients were considered. Like in A4, experienced UI designers and experienced designers were involved in the human-centered evaluation of the proposed prescriptive design knowledge. The use case and task for the experiment to evaluate the effectiveness of customization features in XUIs in A6 were situated in the area of human resources. Therefore, participants with experience in human resources and within-hiring experience were recruited in both online experiments to evaluate the XUI. For the reusability evaluation of the design principles, experienced user interface and user experience designers were recruited. Therefore, only stakeholders with experience in financial investment. Consequently, the advantage was taken by professional networks and online platforms, such as CloudResearch in combination with Amazon Mechanical Turk as well as Prolific, to not only acquire stakeholders with relevant knowledge but also to achieve the necessary sample sizes (Fink, 2022; Karahanna et al., 2018).

However, a limitation regarding the involved participants is that the perspective of stakeholders, such as managers or regulators of (X)AI systems, was neglected (Arrieta et al., 2020; Meske et al., 2022). Therefore, future research can broaden the focus of stakeholders and involve, for example, managers and regulators who also need XAI to ensure aspects like accountability, fairness, or responsibility for the deployment of XAI systems, especially in high-stake scenarios (Langer et al., 2021). In addition, XAI's role and the XUIs' design could be relevant for future research on XAI driven by legislative motivations (Schneider et al., 2022). Lastly, despite the investigation of different domains and application contexts, the results of the cumulative dissertation are in line with prior research that emphasizes the high relevance of human-centered research in the field of XAI since there is most likely no one-size-fits-all approach to the design of explanations and XUIs (Nazar et al., 2021; Sokol & Flach, 2020; Vilone & Longo, 2021; van der Waa et al., 2021).

After discussing the cumulative dissertation, the next and final section presents the conclusion.

6 Conclusion

Since the beginning of the first research projects, which are part of the cumulative dissertation, and hence this cumulative dissertation, the research interest in XAI from different disciplines, including ISR, has enormously grown. As shown in A1, research focusing on explanations and explainability features in decision-making has a long history in ISR. However, we witness a resurgence of this research interest and an evolution of research focusing on XAI with its potential consequences, benefits, rising challenges, dangers, and many more intriguing facets. A plethora of unprecedented application and research opportunities arise. Due to the constant new developments and breakthroughs in AI and XAI, there is no end in sight to these emerging possibilities. New information systems that integrate (X)AI are constantly introduced not only by research but more and more often from industries with performances and growth rates in terms of users never seen before (Maslej et al., 2023). Such systems enter organizations, influencing and changing our work life, environment, and private life, how we consume digital content, what we buy, and which routes we take for traveling or supporting us in the most diverse decision-making processes.

Therefore, I am convinced that future research on the design, perception, and implications of (X)AI will continuously flourish. This is not only supported by the growing and interdisciplinary body of research on this subject (e.g., Ali et al., 2023; Arrieta et al., 2020; Haque et al., 2023; Meske et al., 2022; Miller, 2019; Minh et al., 2022; Samtani et al., 2023; Wells & Bednarz, 2021). It is also supported by two relevant learnings that were omnipresent throughout all conducted research projects that are part of the cumulative dissertation: (i) the high relevance of the stakeholder involvement with their diverse experiences, needs, expectations, or biases, hence making a human-centered approach indispensable; and (ii) the high relevance of interdisciplinary research teams and projects since XAI is such a multifaceted subject that requires knowledge from a diverse set of research disciplines, including ISR with behavioral science as well as DSR, computer science, economics, law, philosophy, political science, or sociology, to name just a few relevant disciplines. Finally, this cumulative dissertation contributes to this exciting and multifaceted research landscape through the individual research projects that are part of the cumulative dissertation and the consolidated ISDT for human-centered XUIs as a summarization.

7 References

Adadi, A., & Berrada, M. (2018). Peeking inside the Black-Box: A Survey on Explainable Artificial Intelligence (XAI). *IEEE Access, 6*, 52138-52160. https://doi.org/10.1109/ACCESS.2018.2870052

Agerfalk, P. J., Conboy, K., Crowston, K., Lundström, J. E., Jarvenpaa, S. L., Mikalef, P., & Ram, S. (2022). Artificial Intelligence in Information Systems: State of the Art and Research Roadmap. *Communications of the Association for Information Systems, 50*, 556-574. https://doi.org/10.17705/1CAIS.05017

Al-Natour, S., Benbasat, I., & Cenfetelli, R. T. (2022). Designing Caring and Informative Decision Aids to Increase Trust and Enhance the Interaction Atmosphere. *AIS Transactions on Human-Computer Interaction, 14*(1), 1-29. https://doi.org/10.17705/1thci.00159

Ali, S., Abuhmed, T., El-Sappagh, S., Muhammad, K., Alonso-Moral, J. M., Confalonieri, R., Guidotti, R., Del Ser, J., Diaz-Rodrigeuz, N., & Herrera, F. (2023). Explainable Artificial Intelligence (XAI): What we know and what is left to attain Trustworthy Artificial Intelligence. *Information Fusion*, Article 101805. https://doi.org/10.1016/j.inffus.2023.101805

Alufaisan, Y., Marusich, L. R., Bakdash, J. Z., Zhou, Y., & Kantarcioglu, M. (2021). Does Explainable Artificial Intelligence Improve Human Decision-Making? *Proceedings of the AAAI Conference on Artificial Intelligence, 35*(8), 6618-6626. https://doi.org/10.1609/aaai.v35i8.16819

Angelov, P. P., Soares, E. E., Jiang, R., Arnold, N. I., & Atkinson, P. M. (2021). Explainable Artificial Intelligence: an analytical review. *WIREs Data Mining and Knowledge Discovery, 11*(5), Article e1424. https://doi.org/10.1002/widm.1424

Anjomshoae, S., Najjar, A., Calbaresi, D., & Främling, K. (2019). Explainable Agents and Robots: Results from a Systematic Literature Review. *AAMAS'19: Proceedings of the 18th International Conference on Autonomous Agents and MultiAgent Systems,* 1078-1088.

Arnott, D., & Graham, P. (2012). Design Science in Decision Support Systems Research: An Assessment using the Hevner, March, Park, and Ram Guidelines. *Journal of the Association for Information Systems, 13*(11), 923-949. https://doi.org/10.17705/1jais.00315

Arrieta, A. B., Diaz-Rodriguez, N., Del Ser, J., Bennetot, A., Tabik, S., Barbado, A., Garcia, S., Gil-Lopez, S., Molina, D., Benjamins, R., Chatila, R., & Herrera, F. (2020). Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI. *Information Fusion, 58*, 82-115. https://doi.og/10.1016/j.inffus.2019.12.012

Avdiji, H., Elikan, D., Missonier, S., & Pigneur, Y. (2020). A Design Theory for Visual Inquiry Tools. *Journal of the Association for Information Systems, 21*(3), 695-734. https://doi.org/10.17705/1jais.00617

Bagayogo, F. F., Lapointe, L., & Bassellier, G. (2014). Enhanced Use of IT: A New Perspective on Post-Adoption. *Journal of the Association for Information Systems, 15*(7), 361-387. https://doi.org/10.17705/1jais.00367

Barda, A. J., Horvat, C. M., & Hochheiser, H. (2020). A qualitative research framework for the design of user-centered displays of explanations for machine learning model predictions in healthcare. *BMC Medical Informatics and Decision making, 20*, Article 257. https://doi.org/10.1186/s12911-020-01276-x

Barrett, M., Davidson, E., Prabhu, J., & Vargo, S. L. (2015). Service innovation in the digital age: key contributions and future directions. *MIS Quarterly, 39*(1), 135-154. https://doi.org/10.25300/MISQ/2015/39:1.03

Baskerville, R., Baiyere, A., Gregor, S., Hevner, A., & Rossi, M. (2018). Design Science Research Contributions: Finding a Balance between Artifact and Theory. *Journal of the Association for Information Systems, 19*(5), 358-376. https://doi.org/10.17705/1jais.00495

Baskerville, R., & Pries-Heje, J. (2010). Explanatory Design Theory. *Business & Information Systems Engineering, 2*, 271-282. https://doi.org/10.1007/s12599-010-0118-4

Baskerville, R., & Pries-Heje, J. (2019). Projectability in Design Science Research. *Journal of Information Technology Theory and Application, 20*(1), 53-76. https://aisel.aisnet.org/jitta/ol20/iss1/3

Bau, D., Zhou, B., Khosla, A., Oliva, A., & Torralba, A. (2017). Network Dissection: Quantifying Interpretability of Deep Visual Representations. *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR),* 3319-3327. https://doi.org/10.1109/CVPR.2017.354

Bauer, K., von Zahn, M., & Hinz, O. (2023). Expl(AI)ned: The Impact of Explainable Artificial Intelligence on Users' Information Processing. *Information Systems Research*, 1-21. https://doi.org/10.1287/isre.2023.1199

Beck, R., Weber, S., & Gregory, R. W. (2013). Theory-generating design science research. *Information Systems Frontiers, 15*, 637-651. https://doi.org/10.1007/s10796-012-9342-4

Becker, J., vom Brocke, J., Heddier, M., & Seidel, S. (2015). In Search of Information Systems (Grand) Challenges. *Business & Information Systems Engineering, 57*, 377-390. https://doi.org/10.1007/s12599-015-0394-0

Beede, E., Baylor, E., Hersch, F., Iurchenko, A., Wilcox, L., Ruamviboonsuk, P., & Vardoulakis, L. M. (2020). A Human-Centered Evaluation of a Deep Learning System Deployed in Clinics for the Detection of Diabetic Retinopathy. *CHI'20: Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems,* 1-12. https://doi.org/10.1145/3313831.3376718

Benbasat, I., Goldstein, D. K., & Mead, M. (1987). The Case Research Strategy in Studies of Information Systems. *MIS Quarterly, 11*(3), 369-386. https://doi.org/10.2307/248684

Benbya, H., Pachidi, S., & Jarvenpaa, S. L. (2021). Special Issue Editorial: Artificial Intelligence in Organizations: Implications for Information Systems Research. *Journal of the Association for Information Systems, 22*(2), 281-303. https://doi.org/10.17705/1jais.00662

Berente, N., Gu, B., Recker, J., & Santhanam, R. (2021). Managing Artificial Intelligence. *MIS Quarterly, 45*(3), 1433-1450. https://doi.org/10.25300/MISQ/2021/16274

vom Brocke, J., & Maedche, A. (2019). The DSR grid: six core dimensions for effectively planning and communicating design science research projects. *Electronic Markets, 29*, 379-385. https://doi.org/10.1007/s12525-019-00358-7

Bradley, C., Wu, D., Tang, H., Singh, I., Wydant, K., Capps, B., Wong, K., Agostinelli, F., Irvin, M., & Srivastava, B. (2022). Explainable Artificial Intelligence (XAI) User Interface Design for Solving Rubik's Cube. In Stephanidis, C., Antona, M., Ntoa, S., Salvendy, G. (Eds.), *HCI International 2022 – Late Breaking Posters, HCII 2022*, *Communications in Computer and Information Science*, vol. 1655 (pp. 605-612). Springer, Cham. https://doi.org/10.1007/978-3-031-19682-9_76

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology, 3*(2), 77-101. https://doi.org/10.1191/1478088706qp063oa

vom Brocke, J., Simons, A., Riemer, K., Niehaves, B., & Plattfaut, R. (2015). Standing on the Shoulders of Giants: Challenges and Recommendations of Literature Search in Information Systems Research. *Communications of the Association for Information Systems, 37*, 205-224. https://doi.org/10.17705/1CAIS.03709

vom Brocke, J., Winter, R., Hevner, A. R., & Maedche, A. (2020). Special Issue Editorial – Accumulation and Evolution of Design Knowledge in Design Science Research: A Journey Through Time and Space. *Journal of the Association for Information Systems, 21*(3), 520-544. https://doi.org/10.17705/1jais.00611

Buiten, M. C. (2019). Towards Intelligent Regulation of Artificial Intelligence. *European Journal of Risk Regulation, 10*(1), 41-59. https://doi.org/10.1017/err.2019.8

Bunde, E. (2021). AI-Assisted and Explainable Hate Speech Detection for Social Media Moderators – A Design Science Approach. *Proceedings of the 54th Hawaii International Conference on System Sciences,* 1264-1273. http://hdl.handle.net/10125/70766

Bunde, E. (2023). The Design and Perception of Customization Features in Explanation User Interfaces. Submitted to: *Human-Computer Interaction*.

Bunde, E., Eisenhardt, D., Sonntag, D., Profitlich, H.-J., & Meske, C. (2023). Giving DIAnA More TIME – Guidance for the Design of XAI-Based Medical Decision Support Systems. In Gerber, A., Baskerville, R. (Eds.), *Design Science Research for a New Society: Society 5.0., DESRIST 2023, Lecture Notes in Computer Science*, vol. 13873 (pp. 107-122). Springer, Cham. https://doi.org/10.1007/978-3-031-32808-4_7

Bunde, E., Kühl, N., & Meske, C. (2022). Fake or Credible? Towards Designing Services to Support Users' Credibility Assessment of News Content. *Proceedings of the 55th Hawaii International Conference on System Sciences,* 1883-1892. https://hdl.handle.net/10125/79569

Cai, C. J., Reif, E., Hegde, N., Hipp, J., Kim, B., Smilkov, D., Wattenberg, M., Viegas, F., Corrado, G. S., Stumpte, M. C., & Terry, M. (2019). Human-Centered Tools for Coping with Imperfect Algorithms During Medical Decision-Making. *CHI'19: Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems,* 1-14. https://doi.org/10.1145/3290605.3300234

Calisto, F. M., Santiago, C., Nunes, N., & Nascimento, J. C. (2021). Introduction of human-centric AI assistant to aid radiologists for multimodal breast image classification. *International Journal of Human-Computer Studies, 150*, Article 102607. https://doi.org/10.1016/j.ijhcs.2021.102607

Cater-Steel, A., Toleman, M., & Rajaeian, M. M. (2019). Design Science Research in Doctoral Projects: An Analysis of Australian Theses. *Journal of the Association for Information Systems, 20*(12), 1844-1869. https://doi.org/10.17705/1jais.00587

Chandra Kruse, L., Sandeep, P., & Seidel, S. (2022). How Designers Use Design Principles: Design Behaviors and Application Modes. *Journal of the Association for Information Systems, 23*(5), 1235-1270. https://doi.org/10.17705/1jais.00759

Chandra Kruse, L., Seidel, S., & Gregor, S. (2015). Prescriptive Knowledge in IS Research: Conceptualizing Design Principles in Terms of Materiality, Action, and Boundary Conditions. *2015 48th Hawaii International Conference on System Sciences,* 4039-4048. https://doi.org/10.1109/HICSS.2015.485

Chanson, M., Bogner, A., Bilgeri, D., Fleisch, E., & Wortmann, F. (2019). Blockchain for the IoT: Privacy-Preserving Protection of Sensor Data. *Journal of the Association for Information Systems, 20*(9), 1274-1309. https://doi.org/10.17705/1jais.00567

Chen, N.-C., Drouhard, M., Kocielnik, R., Suh, J., & Aragon, C. R. (2018). Using Machine Learning to Support Qualitative Coding in Social Science: Shifting the Focus to Ambiguity. *ACM Transactions on Interactive Intelligent Systems, 8*(2), Article 9. https://doi.org/10.1145/3185515

Chen, H., Gomez, C., Huang, C.-M., & Unberath, M. (2022). Explainable medical imaging AI needs human-centered design: guidelines and evidence from a systematic review. *npj Digital Medicine, 5*, Article 156. https://doi.org/10.1038/s41746-022-00699-2

Chromik, M., & Butz, A. (2021). Human-XAI Interaction: A Review and Design Principles for Explanation User Interfaces. In Ardito, C., Lanzilotti, R., Malizia, A., Petrie, H., Piccinno, A., Desolda, G., Inkpen, K. (Eds.), *Human-Computer Interaction, INTERACT 2021, Lecture Notes in Computer Science*, vol. 12933 (619-640). Springer, Cham. https://doi.org/10.1007/978-3-030-85616-8_36

Coenen, T., Coertjens, L., Vlerick, P., Lesterhuis, M., Mortier, A. V., Donche, V., Ballon, P., & De Maeyer, S. (2018). An information system design theory for the comparative judgement of competences. *European Journal of Information Systems, 27*(2), 248-261. https://doi.org/10.1080/0960085X.2018.1445461

Davis, F. D. (1989). Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology. *MIS Quarterly, 13*(3), 319-340. https://doi.org/10.2307/249008

Dhaliwal, J. S., & Benbasat, I. (1996). The Use and Effects of Knowledge-Based System Explanations: Theoretical Foundations and a Framework for Empirical Evaluation. *Information Systems Research, 7*(3), 342-362. https://doi.org/10.1287/isre.7.3.342

Diederich, S., Brendel, A. B., & Kolbe, L. M. (2020). Designing Anthropomorphic Enterprise Conversational Agents. *Business Information Systems & Engineering, 62*(3), 193-209. https://doi.org/10.1007/s12599-020-00639-y

Dietvorst, B. J., Simmons, J. P., & Massey, C. (2015). Algorithm aversion: People erroneously avoid algorithms after seeing them err. *Journal of Experimental Psychology: General, 144*(1), 114-126. https://doi.org/10.1037/xge0000033

Dikmen, M., & Burns, C. (2020). Abstraction Hierarchy Based Explainable Artificial Intelligence. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 64*(1), 319-323. https://doi.org/10.1177/1071181320641073

Dosilovic, F. K., Brcic, M., & Hlupic, N. (2018). Explainable artificial intelligence: A survey. *2018 41st International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO),* 210-215. https://doi.org/10.23919/MIPRO.2018.8400040

Drechsler, A., & Hevner, A. R. (2018). Utilizing, Producing, and Contributing Design Knowledge in DSR Projects. In Chatterjee, S., Dutta, K., Sundarraj, R. (Eds.), *Designing for a Digital and Globalized World. DESRIST 2018. Lecture Notes in Computer Science,* vol. 10844 (pp. 82-97). Springer, Cham. https://doi.org/10.1007/978-3-319-91800-6_6

Dudley, J. J., & Kristensson, P. O. (2018). A Review of User Interface Design for Interactive Machine Learning. *ACM Transactions on Interactive Intelligent Systems, 8*(2), Article 8. https://doi.org/10.1145/3185517

Dwivedi, Y. K., Kshetri, N., Hughes, L., Slade, E. L., Jeyaraj, A., Kar, A. K., Baabdullah, A. M., Koohang, A:, Raghavan, V., Ahuja, M., Albanna, H., Albashrawi, M. A., Al-Busaidi, A. S., Balakrishnan, J., Barlette, Y., Basu, S., Bose, I., Brooks, L., Buhalis, D., Carter, L., Chowdhury, S., Crick, T., Cunningham, S. W., Davies, G. H., Davison, R. M., Dé, R., Dennehy, D., Duan, Y., Dubey, R., Dwivedi, R., Edwards, J. S., Flavian, C., Gauld, R., Grover, V., Hu, M.-C., Janssen, M., Jones, P., Junglas, I., Khorana, S., Kraus, S., Larsen, K. R., Latreille, P., Laumer, S., Malik, F. T., Mardani, A., Mariani, M., Mithas, S., Mogaji, E., Nord, J. H., O´Connor, S., Okumus, F., Pagani, M., Pandey, N., Papagiannidis, S., Pappas, I. O., Pathak, N., Pries-Heje, J., Raman, R., Rana, N. P., Rehm, S.V., Ribeiro-Navarrete, S., Richter, A., Rowe, F., Sarker, S., Stahl, B. C., Tiwari, M. K., van der Aalst, W., Venkatesh, V., Viglia, G., Wade, M., Walton, P., Wirtz, J., Wright, R. (2023). *International Journal of Information Management, 71*, Article 102642. https://doi.org/10.1016/j.ijinfomgt.2023.102642

Einola, K., & Khoreva, V. (2023). Best friend or broken tool? Exploring the co-existence of humans and artificial intelligence in the workplace ecosystem. *Human Resource Management, 62*(1), 117-335. https://doi.org/10.1002/hrm.22147

Eshan, U., Wintersberger, P., Liao, Q. V., Watkins, E. A., Manger, C., & Ill, H. D. (2022). Human-Centered Explainable AI (HCAI): Beyond Opening the Black-Box of AI. *CHI EA'22: Extended Abstracts of the 2022 CHI Conference on Human Factors in Computing Systems*, 1-7. https://doi.org/10.1145/3491101.3503727

Fink, L. (2022). Why and How Online Experiments Can Benefit Information Systems Research. *Journal of the Association for Information Systems, 23*(6), 1333-1346. https://doi.org/10.17705/1jais.00787

Fiok, K., Farahani, F. V., Karwowski, W. & Ahram, T. (2022). Explainable artificial intelligence for education and training. *Journal of Defense Modeling and Simulation: Applications, Methodology, Technology, 19*(2), 133-144. https://doi.org/10.1177/15485129211028651

Fügener, A., Grahl, J., Gupta, A., & Ketter, W. (2022). Cognitive Challenges in Human-Artificial Intelligence Collaboration: Investigating the Path Toward Productive Delegation. *Information Systems Research, 33*(2), 678-696. https://doi.org/10.1287/isre.2021.1079

Galliers, R. D., & Land, F. F. (1987). Viewpoint: choosing appropriate information systems research methodologies. *Communications of the ACM, 30*(11), 901-902. https://doi.org/10.1145/32206.315753

Garibay, O. O., Winslow, B., Andolina, S., Antona, M., Bodenschatz, A., Coursaris, C., Falco, G., Fiore, S. M., Garibay, I., Grieman, K., Havens, J. C., Jirotka, M., Kacorri, H., Karwowski, W., Kider, J., Konstan, J., Koon, S., Lopez-Gonzalez, M., Maifeld-Carucci, I., McGregor, S., Salvendy, G., Schneiderman, B., Stephanidis, C., Strobel, C., Holter, C. T., & Xu, W. (2023). Six Human-Centered Artificial Intelligence Grand Challenges. *International Journal of Human-Computer Interaction 39*(3), 391-437. https://doi.org/10.1080/10447318.2022.2153320

Gasson, S. (2003). Human-Centered Vs. User-Centered Approaches to Information System Design. *The Journal of Information Technology Theory and Application, 5*(2), 29-46. https://aisel.aisnet.org/jita/vol5/iss2/5

Giacomin, J. (2014). What Is Human Centred Design? *The Design Journal, 17*(4), 606-623. https://doi.org/10.2752/175630614X14056185480186

Germonprez, M., Kendall, J. E., Mathiassen, L., Young, B., & Warner, B. (2016). A Theory of Responsive Design: A Field Study of Corporate Engagement with Open Source Communities. *Information Systems Research, 28*(1), 64-83. https://doi.org/10.1287/isre.2016.0662

Giessmann, A., & Legner, C. (2016). Designing business models for cloud platforms. *Information Systems Journal, 26*, 551-579. https://doi.org/10.1111/isj.12107

Gilpin, L. H., Bau, D., Yuan, B. Z., Bajwa, A., Specter, M., & Kagal, L. (2018). *Explaining Explanations: An Overview of Interpretability of Machine Learning*. *2018 IEEE 5th International Conference on Data Science and Advanced Analytics (DSAA)*, 80-90. https://doi.org/10.1109/DSAA.2018.00018

Goddard, K., Roudsari, A., & Wyatt, J. C. (2012). Automation bias: a systematic review of frequency, effect mediators, and mitigators. *Journal of the American Medical Informatics Association, 19*(1), 121-127. https://doi.org/10.1136/amiajnl-2011-000089

Goes, P. B. (2014). Editor's Comments: Design Science Research in Top Information Systems Journals. *MIS Quarterly, 38*(1), iii-viii. https://www.jstor.org/stable/26554865

Goldkuhl, G. (2019). The Generation of Qualitative Data in Information Systems Research: The Diversity of Empirical Research Methods. *Communications of the Association for Information Systems, 44*, 572-599. https://doi.org/10.17705/1CAIS.04428

Gregor, S. (2006). The Nature of Theory in Information Systems. *MIS Quarterly, 30*(3), 611-642. https://doi.org/10.2307/25148742

Gregor, S., Chandra, Kruse, L., & Seidel, S. (2020). Research Perspectives: The Anatomy of a Design Principle. *Journal of the Association for Information Systems, 21*(6), 1622-1652. https://doi.org/10.17705/1jais.00649

Gregor, S., & Hevner, A. R. (2013). Positioning and presenting design science research for maximum impact. *MIS Quarterly, 37*(2), 337-356. https://doi.org/10.25300/MISQ/2013/37.2.01

Gregor, J., & Jones, D. (2007). The Anatomy of a Design Theory. *Journal of the Association for Information Systems, 8*(5), 312-335. https://doi.org/10.17705/1jais.00129

Gu, H., Liang, Y., Xu, Y., Williams, C. K., Magaki, S., Khanlou, N., Ni, S., Yang, C., Yan, W., Zhang, X. R., Li, Y., Haeri, M., & Chen, X. A. (2023). Improving Workflow Integration with xPath: Design and Evaluation of Human-AI Diagnosis System in Pathology. *ACM Transactions on Computer-Human Interaction, 30*(2), Article 28. https://doi.org/10.1145/3577011

Guidotti, R. (2021). Evaluating local explanation methods on ground truth. *Artificial Intelligence, 291*, Article 103428. https://doi.org/10.1016/j.artint.2020.103428

Gunning, D., & Aha, D. (2019). DARPA's Explainable Artificial Intelligence (XAI) Program. *AI Magazine, 40*(2), 44-58. https://doi.org/10.1609/aimag.v40i2.2850

Gunning, D., Stefeik, M., Choi, J., Miller, T., Stumpf, S., & Yang, G.-Z. (2019). XAI – Explainable artificial intelligence. *Science Robotics, 4*, Article eaay7120. https://doi.org/10.1126/scirobotics.aay7120

Ha, T., Sah, Y. J., Park, Y., & Lee, S. (2022). Examining the effects of power status of an explainable artificial intelligence system on users' perceptions. *Behavior & Information Technology, 41*(5), 946-958. https://doi.org/10.1080/0144929X.2020.1846789

Habers, M., van den Bosch, K., & Meyer, J.-J. (2010). Design and Evaluation of Explainable BDI Agents. *2010 IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology,* 125-132. https://doi.org/10.1109/WI-IAT.2010.115

Hevner, A. R. (2007). A Three Cycle View of Design Science Research. *Scandinavian Journal of Information Systems, 19*(2), 87-92. https://aisel.aisnet.org/sjis/vol19/iss2/4

Hevner, A. R. (2021). The duality of science: Knowledge in information systems research. *Journal of Information Technology*, 36(1), 72-76. https://doi.org/10.1177/0268396220945714

Hevner, A. R., & March, S. T. (2003). The information systems research cycle. *IEEE Computer, 36*(11), 111-113. https://doi.org/10.1109/MC.2003.1244541

Hevner, A. R., March, S. T., Park, J., & Ram, S. (2004). Design Science in Information Systems Research. *MIS Quarterly, 28*(1), 77-105. https://doi.org/10.2307/25148625

Haque, A. K. M. B., Islam, A. K. M. N., & Mikalef, P. (2023). Explainable Artificial Intelligence (XAI) from a user perspective: A synthesis of prior literature and problematizing avenues for future research. *Technological Forecasting & Social Change, 186*, Article 122120. https://doi.org/10.1016/j.techfore.2022.122120

Herm, L.-V., Steinbach, T., Wanner, J., & Janiesch, C. (2023). A nascent design theory for explainable intelligent systems. *Electronic Markets, 32*, 2185-2205. https://doi.org/10.1007/s12525-022-00606-3

Herrmann, T., & Pfeiffer, S. (2022). Keeping the organization in the loop: a socio-technical extension of human-centered artificial intelligence. *AI & Society*, 1-20. https://doi.org/10.1007/s00146-022-01391-5

Huysmans, J., Dejaeger, K., Mues, C., Vanthienen, & Baesens, B. (2011). An empirical evaluation of the comprehensibility of decision table, tree and rule based predictive models. *Decision Support Systems, 51*, 141-154. https://doi.org/10.1016/j.dss.2010.12.003

Ivari, J. (2007). A Paradigmatic Analysis of Information Systems As a Design Science. *Scandinavian Journal of Information Systems, 19*(2), 39-64. http://aisel.aisnet.org/sjis/vol19/iss2/5

Ivari, J. (2020). Editorial: A Critical Look at Theories in Design Science Research. *Journal of the Association for Information Systems, 21*(3), 502-519. https://doi.org/10.17705/1jais.00610

Ivari, J., Hansen, M. R. P., & Haj-Bolouri, A. (2021). A proposal for minimum reusability evaluation of design principles. *European Journal of Information Systems, 30*(3), 286-303. https://doi.org/10.1080/0960085X.2020.1793697

Ivari, I., & Ivari, N. (2011). Varieties of user-centeredness: an analysis of four systems development methods. *Information Systems Journal, 21*(2), 125-153. https://doi.org/10.1111/j.1365-2575.2010.00351.x

Jarrahi, M. H. (2018). Artificial intelligence and the future of work: Human-AI symbiosis in organizational decision making. *Business Horizons, 61*(4), 577-586. https://doi.org/10.1016/j.bushor.2018.03.007

Jeyaraj, A., & Zadeh, A. H. (2020). Evolution of information systems research: Insights from topic modeling. *Information & Management, 57*(4), Article 103207. https://doi.org/10.1016/j.im.2019.103207

Jorritsma, W., Cnossen, F., & van Ooijen, P. M. A. (2015). Adaptive support for user interface customization: a study in radiology. *International Journal of Human-Computer Studies, 77*, 1-9. https://doi.org/10.1016/j.ijhcs.2014.12.008

Jussupow, E., Spohrer, K., Heinzl, A., & Gawlitza, J. (2021). Augmenting Medical Diagnosis Decisions? An Investigation into Physicians' Decision-Making Process with Artificial Intelligence. *Information Systems Research, 32*(2), 713-735. https://doi.org/10.1287/isre.2020.0980

Kane, G. C., Young, A. G., Majchrzak, A., & Ransbotham, S. (2021). Avoiding an Oppressive Future of Machine Learning: A Design Theory for Emancipatory Assistants. *MIS Quarterly, 45*(1), 371-396. https://doi.org/10.25300/MISQ/2021/1578

Karahanna, E., Benbasat, I., Bapna, R., & Rai, A. (2018). Opportunities and Challenges for Different Types of Online Experiments. *MIS Quarterly, 42*(4), iii-x.

Kaur, H., Nori, H., Jenkins, S., Caruana, R., Wallach, H., & Vaighan, J. W. (2020). Interpreting Interpretability: Understanding Data Scientists' Use of Interpretability Tools for Machine Learning. *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems,* 1-14. https://doi.org/10.1145/3313831.3376219

Kim, B., Rudin, C., & Shah, J. (2014). The Bayesian case model: a generative approach for case-based reasoning and prototype classification. *Proceedings of the 27th International Conference on Neural Information Processing Systems, 2*, 1952-1960. https://doi.org/10.48550/arXiv.1503.01161

Knapic, S., Malhi, A., Saluja, R., & Främling, K. (2021). Explainable Artificial Intelligence for Human Decision Support System in the Medical Domain. *Machine Learning & Knowledge Extraction, 3,* 740-770. https://doi.org/10.3390/make303037

Kocielnik, R., Amershi, S., & Bennett, P. N. (2019). Will You Accept an Imperfect AI? Exploring Designs for Adjusting End user Expectations of AI Systems. *CHI'19: Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, 1-14. https://doi.org/10.1145/3290605.3300641

Kuechler, W., & Vaishnavi, V. (2012). A Framework for Theory Development in Design Science Research: Multiple Perspectives. *Journal of the Association for Information Systems, 13*(6), 395-423. https://doi.org/10.17705/1jais.00300

Kulesza, T., Burnett, M., Wong, W.-K., & Stumpf, S. (2015). Principles of Explanatory Debugging to Personalize Interactive Machine Learning. *IUI'15: Proceedings of the 20th International Conference on Intelligent User Interfaces,* 126-137. https://doi.org/10.1145/2678025.2701399

Langer, M., Oster, D., Speith, T., Hermanns, H., Kästner, L., Schmidt, E., Sesing, A., & Baum, K. (2021). What do we want from Explainable Artificial Intelligence (XAI)? – A stakeholder perspective on XAI and a conceptual model guiding interdisciplinary XAI research. *Artificial Intelligence, 296*, Article 103473. https://doi.org/10.1016/j.artint.2021.103473

Leichtmann, B., Humer, C., Hinterreiter, A., Streit, M., & Mara, M. (2023). Effects of Explainable Artificial Intelligence on trust and human behavior in a high-risk decision task. *Computers in Human Behavior, 139*, Article 107539. https://doi.org/10.1016/j.chb.2022.107539

Li, S., Gu, X. (2023). A Risk Framework for Human-centered Artificial Intelligence in Education: Based on Literature Review and Delphi-AHP Method. *Educational Technology & Society, 26*(1), 187-202. https://doi.org/10.30191/ETS.202301_26(1).0014

Li, W., Mao, Y., & Zhou, L. (2021). The Impact of Interactivity on User Satisfaction in Digital Social Reading: Social Presence as a Mediator. *International Journal of Human-Computer Interaction, 37*(17), 1636-1647. https://doi.org/10.1080/10447318.2021.1898850

Lim, B. Y., Dey, A. K., & Avrahami, D. (2009). Why and why not explanations improve the intelligibility of context-aware intelligent systems. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems,* 2119-2128. https://doi.org/10.1145/1518701.1519023

Lins, S., Schneider, S., Szefer, J., Ibraheem, S., & Sunyaev, A. (2019). Designing Monitoring Systems for Continuous Certification of Cloud Services: Deriving Meta-requirements and Design Guidelines. *Communications of the Association for Information Systems, 44*, 460-510. https://doi.org/10.17705/1CAIS.04425

Litman, L., Robinson, J., & Abberbock, T. (2017). TurkPrime.com: A versatile crowdsourcing data acquisition platform for the behavioral sciences. *Behavior Research Methods, 49*, 433-442. https://doi.org/10.3758/s13428-016-0727-z

Lowry, P. B., Wilson, D. W., & Haig, W. L. (2014). A Picture is Worth a Thousand Words: Source Credibility Theory Applied to Logo and Website Design for Heightened Credibility and Consumer Trust. *International Journal of Human-Computer Interaction, 30*(1), 63-93. https://doi.org/10.1080/10447318.2013.839899

Lukyanenko, R., & Parsons, J. (2020). Research Perspectives: Design Theory Indeterminacy: What Is it, How Can it Be Reduced, and Why Did the Polar Bear Drown? *Journal of the Association for Information Systems, 21*(5), 1-59. https://doi.org/10.17705/1jais.00639

Lycett, M., & Radwan, O. (2017). Developing a Quality of Experience (QoE) model for Web Applications. *Information Systems Journal, 29*, 175-199. https://doi.org/10.1111/isj.12192

Maguire, M. (2001). Methods to support human-centered design. *International Journal of Human-Computer Studies, 55*(4), 587-634. https://doi.org/10.1006/ijhc.2001.0503

Mandviwalla, M. (2015). Generating and Justifying Design Theory. *Journal of the Association for Information Systems, 16*(5), 314-344. https://doi.org/10.17705/1jais.00397

Mao, J.-Y., & Benbasat, I. (2000). The Use of Explanations in Knowledge-Based Systems: Cognitive Perspectives and a Process-Tracing Analysis. *Journal of Management Information Systems, 17*(2), 153-179. https://doi.org/10.1080/07421222.2000.11045646

March, S. T., & Storey, V. C. (2008). Design Science in the Information Systems Discipline: An Introduction to the Special Issue on Design Science Research. *MIS Quarterly, 32*(4), 725-730. https://doi.org/10.2307/2518869

Maslej, N., Fattorini, L., Brynjolfsson, E., Etchemendy, J., Ligett, K., Lyons, T., Manyika, J., Ngo, H., Niebles, J. C., Parli, V., Shoham, Y., Wald, R., Clark, J., & Perrault, R. (2023). *The AI Index 2023 Annual Report*. AI Index Steering Committee, Institute for Human-Centered AI, Stanford University, Stanford, CA, April 2023.

Meske, C., & Bunde, E. (2023). Design Principles for User Interfaces in AI-Based Decision Support Systems: The Case of Explainable Hate Speech Detection. *Information Systems Frontiers, 25*, 743-77. https://doi.org/10.1007/s10796-021-10234-5

Meske, C., Bunde, E., & Ehmke, J. F. (2020). Improving Customers' Decision Making on Blackboxed Multimodal Mobility Platforms – A Design Science Approach. *ICIS 2020 Proceedings,* Article 12. https://aisel.aisnet.org/icis2020/digital_commerce/digital_commerce/12

Meske, C., Bunde, E., Schneider, J., & Gersch, M. (2022). Explainable Artificial Intelligence: Objectives, Stakeholders, and Future Research. *Information Systems Management, 39*(1), 53-63. https://10.1080/10580530.2020.1849465

Meth, H., Mueller, B., & Maedche, A. (2015). Designing a Requirement Mining System. *Journal of the Association for Information Systems, 16*(9), 799-837. https://doi.org/10.100705/1jais.00408

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

Minh, D., Wang, H. X., Li, Y. F., & Nguyen, T. N. (2022). Explainable artificial intelligence: a comprehensive review. *Artificial Intelligence Review, 55*, 3503-3568. https://doi.org/10.1007/s10462-021-10088-y

Mirchi, N., Bissonnette, V., Yilmaz, R., Ledwos, N., Winkler-Schwartz, A., Del Maestro, R. F. (2020). The Virtual Operative Assistant: An explainable artificial intelligence tool for simulation-based training in surgery and medicine. *PLoS ONE, 15*(2), Article e0229596. https://doi.org/10.1371/journal.pone.0229596

Morana, S., Kroenung, J., Maedche, A., & Schacht, S. (2019). Designing Process Guidance Systems. *Journal of the Association for Information Systems, 20*(5), 499-535. https://doi.org/10.17705/1jais.00542

Naisehe, M., Al-Thani, D., Jiang, N., Ali, R. (2023). How the different explanation classes impact trust calibration: The case of clinical decision support systems. *International Journal of Human-Computer Studies, 169*, Article 102941. https://doi.org/10.1016/j.ijhcs.2022.102941

Nazar, M., Alam, M. M., Yafi, E., & Su'ud, M. M. (2021). A Systematic Review of Human-Computer Interaction and Explainable Artificial Intelligence in Healthcare With Artificial Intelligence Techniques. *IEEE Access, 9*, 153316-153348. https://doi.org/10.1109/ACCES.2021.3127881

Niehaves, B., & Ortbach, K. (2016). The inner and the outer model in explanatory design theory: the case of designing electronic feedback systems. *European Journal of Information Systems, 25*(4), 303-316. https://doi.org/10.1057/ejis.2016.3

Norman, D. A. (2005). Human-centered design considered harmful. *Interactions, 12*(4), 14-19. https://doi.org/10.1145/1070960.1070976

Okoli, C. (2015). A Guide to Conducting a Standalone Systematic Literature Review. *Communications of the Association for Information Systems, 37*, 879-910. https://doi.org/10.17705/1CAIS.03743

Österle, H., Becker, J., Frank, U., Hess, T., Karagiannis, D., Krcmar, H., Loos, P., Mertens, P., Oberweis, A., & Sinz, E. J. (2011). Memorandum on design-oriented information systems research. *European Journal of Information Systems, 20*(1), 7-10. https://doi.org/10.1057/ejis.2010.55

Padmanabhan, B., Fang, X., Sahoo, N., & Burton-Jones, A. (2022). Machine Learning in Information Systems Research. *MIS Quarterly, 46*(1), iii-xix.

Patel, B. N., Rosenberg, L., Willcox, G., Baltaxe, D., Lyons, M., Irvin, J., Rajpurkar, P., Amrhein, T., Gupta, R., Halabi, S., Langlotz, Lo, E., Mammarappallil, A.J. Mariano, Riley, G., Seekings, J., Shen, L., Zucker, E., & Lungren, M. P. (2019). Human-machine partnership with artificial intelligence for chest radiograph diagnosis. *npj Digital Medicine, 2*, Article 111. https://doi.org/10.1038/s41746-019-0189-7

Patton, M. Q. (2002). *Qualitative research & evaluation methods*. Thousand Oaks, CA: Sage Publications.

Peer, E., Rothschild, D., Gordon, A., Evernden, Z., & Damer, E. (2022). Data quality of platforms and panels for online behavioral research. *Behavior Research Methods, 54*, 1643-1662. https://doi.org/10.3758/s13428-021-01694-3

Peffers, K., Tuunanen, T., & Niehaves, B. (2018). Design science research genres: introduction to the special issue on exemplars and criteria for applicable design science research. *European Journal of Information Systems, 27*(2), 129-139. https://doi.org/10.1080/0960085X.2018.1458066

Peffers, K., Tuunanen, T., Rothenberger, M. A., & Chatterjee, S. (2007). A Design Science Research Methodology for Information Systems Research. *Journal of Management Information Systems, 24*(3), 45-77. https://doi.org/10.2753/MIS0742-1222240302

Rahwan, I., Cebrian, M., Obradovich, N., Bongard, J., Bonnefon, J.-F., Breazeal, C., Crandall, J. W., Christakis, N. A., Couzin, I. D., Jackson, M. O., Jennings, N. R., Kamar, E., Kloumann, I. M., Larochelle, H., Lazer, D., McElreath, R., Mislove, A., Parkes, D. C., Pentland, A. S., Roberts, M. E., Shariff, A., Tenenbau, J. B., & Wellman, M. (2019). Machine behaviour. *Nature, 568*, 477-486. https://doi.org/10.1038/s41586-019-1139-y

Rai, A. (2017). Diversity of Design Science Research. *MIS Quarterly, 41*(1), iii-xviii.

Rai, A., Constantinides, P., & Sarker, S. (2019). Editor's Comments: Next-Generation Digital Platforms: Toward Human-AI Hybrids. *MIS Quarterly, 43*(1), iii-ix.

Ribeiro, M. T., Singh, S., & Guestrin, C. (2016). "Why Should I Trust You?": Explaining the Predictions of Any Classifier. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining,* 1135-1144. https://doi.org/10.1145/2939672.2939778

Ribeiro, M. T., Singh, S., & Guestrin, C. (2018). Anchors: High-Precision Model-Agnostic Explanations. *Proceedings of the AAAI Conference on Artificial Intelligence, 32*(1), 1527-1535. https://doi.org/10.1609/aaai.v32i1.11491

Riedl, M. O. (2019). Human-centered artificial intelligence and machine learning. *Human Behavior and Emerging Technologies, 1*(1), 33-36. https://doi.org/10.1002/hbe2.117

Rosemann, M., & Vessey, I. (2008). Toward Improving the Relevance of Information Systems Research to Practice: The Role of Applicability Checks. *MIS Quarterly, 32*(1), 1-22. https://www.jstor.org/stable/25148826

Rudin, C. (2019). Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. *Nature Machine Intelligence, 1*, 206-215. https://doi.org/10.1038/s42256-019-0048-x

Samtani, S., Zhu, H., Padmanabhan, B., Chai, Y., Chen, H., & Nunamaker Jr., J. F. (2023). Deep Learning for Information Systems. *Journal of Management Information Systems, 40*(1), 271-301. https://doi.org/10.1080/07421222.2023.2172772

Sarker, S., Xiao, X., & Beaulieu, T. (2013). Guest Editorial: Qualitative Studies in Information Systems: A Critical Review and Some Guiding Principles. *MIS Quarterly, 37*(4), iii-xviii. https://www.jstor.org/stable/43825778

Sarro, F., Moussa, R., Petrozziello, A., & Harman, M. (2020). Learning From Mistakes: Machine Learning Enhanced Human Expert Effort Estimates. *IEEE Transactions on Software Engineering, 48*(6), 1868-1882. https://doi.org/10.1109/TSE.2020.3040793

Schneider, J., Abraham, R., Meske, C., & vom Brocke, J. (2022). Artificial Intelligence Governance For Businesses. *Information Systems Management*, 1-21. https://doi.org/10.1080/10580530.2022.2085825

Schneiderman, B. (2020a). Human-Centered Artificial Intelligence: Three Fresh Ideas. *AIS Transactions on Human-Computer Interaction, 12*(3), 109-124. https://doi.org/1.17705/1thci.00131

Schneiderman, B. (2020b). Human-Centered Artificial Intelligence: Reliable, Safe & Trustworthy. International *Journal of Human-Computer Interaction, 36*(6), 495-504. https://doi.org/10.1080/10447318.2020.1741118

Schneiderman, B. (2020c). Design Lessons From AI's Two Grand Goals: Human Emulation and useful Applications. *IEEE Transactions on Technology and Society, 1*(2), 73-82. https://doi.org/10.1109/TTS.2020.2992669

Schoenherr, J. R., Abbas, R., Michael, K., Rivas, P., & Anderson, T. D. (2023). Designing AI Using a Human-Centered Approach: Explainability and Accuracy Toward Trustworthiness. *IEEE Transactions on Technology and Society, 4*(1), 9-23. https://doi.org/10.1109/TTS.2023.3257627

Schoonderwoerd, T. A. J., Jorritsma, W., Neerincx, M. A., & van den Bosch, K. (2021). Human-centered XAI: Developing design patterns for explanations of clinical decision support systems. *International Journal of Human-Computer Studies, 154*, Article 102684. https://doi.org/10.1016/j.ijhcs.2021.102684

Schryen, G., Benlian, A., Rowe, F., Gregor, S., & Larsen, K. (2017). Literature Reviews in IS Research: What Can Be Learnt from the Past and Other Fields? *Communications of the Association for Information Systems, 41*, 759-774. https://doi.org/10.17705/1CAIS.04130

Seidel, S., Chandra Kruse, L., Szekely, N., Gau, M., & Stieger, M. (2015). Design principles for sensemaking support systems in environmental sustainability transformations. *European Journal of Information Systems, 27*(2), 221-247. https://doi.org/10.1057/s41303-017-0039-0

Selbst, A. D., Boyd, D., Friedler, S. A., Venkatasubramanian, S., & Vertesi, J. (2019). Fairness and Abstraction in Sociotechnical Systems. *Proceedings of the Conference on Fairness, Accountability, and Transparency,* 59-68. https://doi.org/10.1145/3287560.3287598

Senoner, J., Netland, T., & Feuerriegel, S. (2022). Using Explainable Artificial Intelligence to Improve Process Quality: Evidence from Semiconductor Manufacturing. *Management Science, 68*(8), 5704-5723. https://doi.org/10.1287/mnsc.2021.4190

Shad, R., Cunningham, J. P., Ashley, E. A., Langlotz, C. P., & Hiesinger, W. (2021). Designing clinically translatable artificial intelligence systems for high-dimensional medical imaging. *Nature Machine Intelligence, 3*, 929-935. https://doi.org/10.38/s42256-021-00399-8

Shen, J., Iandoli, L., & Aguirre-Urreta, M. (2022). Human-Centered Design for Individual and Social Well-being: Editorial Preface. *Transactions on Human-Computer Interaction, 14*(4), 446-460. https://doi.org/10.17705/1thci.00175

Shin, H., Bunosso, I., & Levine, L. R. (2022). The influence of chatbot humour on consumer evaluations of services. *International Journal of Consumer Studies, 47*(2), 545-562. https://doi.org/10.1111/ijcs.12849

Simon, H. A. *The Sciences of the Artificial* (3rd ed.), MIT Press, Cambridge, MA, 1996.

Sokol, K., & Flach, P. (2020). One Explanation Does Not Fit All – The Promise of Interactive Explanations of Machine Learning Transparency. *KI – Künstliche Intelligenz, 34*, 235-250. https://doi.org/10.1007/s13218-020-00637-y

Song, X., Yu, A. S. L., Kellum, J. A., Waitman, L. R., Matheny, M. E., Simpsons, S. Q., Hu, Y., & Liu, M. (2020). Cross-site transportability of an explainable artificial intelligence model for acute kidney injury prediction. *Nature Communications, 11*, Article 5668. https://doi.org/10.1038/s41467-020-19551-w

Sovrano, F., & Vitali, F. (2022). Explanatory artificial intelligence (YAI): human-centered explanations of explainable AI and complex data. *Data Mining and Knowledge Discovery*, 1-28. https://doi.org/10.1007/s10618-022-00872-x

Sperrle, F., El-Assady, M., Guo, G., Borgo, R., Chau, D., H., Endert, A., & Keim, D. (2021). A Survey of Human-Centered Evaluations in Human-Centered Machine Learning. *Computer Graphics Forum, 40*(3), 543-568. https://doi.org/10.1111/cgf.14329

Spinner, T., Schlegel, U., Schäfer, H., & El-Assady, M. (2020). explAIner: A Visual Analytics Framework for Interactive and Explainable Machine Learning. *IEEE Transactions on Visualization and Computer Graphics, 26*(1), 1064-1074. https://doi.org/10.1109/TVCG.2019.2934629

Stark, L. (2018). Algorithmic psychometrics and the scalable subject. *Social Studies of Science, 48*(2), 204-231. https://doi.org/10.1177/0306312718772094

Strich, F., Mayer, A.-S., & Fiedler, M. (2021). What Do I Do in a World of Artificial Intelligence? Investigating the Impact of Substitutive Decision-Making AI Systems on Employees' Professional Role Identify. *Journal of the Association for Information Systems, 22*(2), 304-324. https://doi.org/10.17705/1jais.00663

Sujan, M., Furniss, D., Grundy, K., Grundy, H., Nelson, D., Elliott, M., White, S., Habli, I., & Reynolds, N. (2019). Human factors challenges for the safe use of artificial intelligence in patient are. *BMJ Health Care Informatics, 26*, Article e100081. https://doi.org/10.1136/bmjhci-2019-100081

Sundar, S. S. (2020). Rise of Machine Agency: A Framework for Studying the Psychology of Human-AI Interaction (HAII). *Journal of Computer-Mediated Communication, 25*(1), 74-88. https://doi.org/10.1093/jcmc/zmz026

Sundar, S. S., Jia, H., Waddell, T. F., & Huang, Y. (2015). Toward a theory of interactive media effects (TIME): Four models for explaining how interface features affect user psychology. In S. S. Sundar (Ed.), *The handbook of the psychology of communication technology* (pp. 47-86). Wiley Blackwell. https://doi.org/10.1002/9781118426456.ch3

Taylor, J. E. T., & Taylor, G. W. (2021). Artificial cognition: How experimental psychology can help generate explainable artificial intelligence. *Psychonomic Bulletin and Review, 28*, 454-475. https://doi.org/10.3758/s13423-020-0182-5

Te'eni, D., Seidel, S., & vom Brocke, J. (2017). Stimulating dialog between information systems research and practice. *European Journal of Information Systems, 26*, 541-545. https://doi.org/10.1057/s41303-017-0067-9

Toreini, P., Langner, M., Maedche, A., Morana, S., & Vogel, T. (2022). Designing Attentive Information Dashboards. *Journal of the Association for Information Systems, 23*(2), 521-552. https://doi.org/10.17705/1jais.00732

Tschandl, P., Codella, N., Akay, B. N., Argenziano, G., Braun, R. P., Cabo, H., Gutman, D., Halpern, A., Helba, B., Hofmann-Wellenhof, R., Lallas, A., Lapins, J., Longo, C., Malvehy, J., Marchetti, M. A., Marghoob, A., Menzies, S., Oakley, A., Paoli, J., Puig, S., Rinner, C., Rosendahl, C., Scope, A., Sinz, C., Soyer, H. P., Thomas, L., Zalaudek, I., & Kittler, H. (2019). Comparison of the accuracy of human readers versus machine-learning algorithms for pigmented skin lesion classification: an open, web-based, international, diagnostic study. *Lancet Oncology, 20*(7), 938-947. https://doi.org/10.1016/S1470-2045(19)30333-X

Tuunanen, T., Vartiainen, T., Kainulainen, S., & Ebrahim, M. (2023). Development of an Agile Requirements Risk Prioritization Method: A Design Science Research Study. *Communications of the Association for Information Systems, 52*, 123-149. https://aisel.aisnet.org/cais/vol52/iss1/24

Venable, J. R. (2006). The role of theory and theorizing in design science research. In *Proceedings of the 1st Conference on Design Science Research in Information Systems and Technology*.

Venable, J. R., Pries-Heje, J., & Baskerville, R. (2016). FEDS: a Framework for Evaluation in Design Science Research. *European Journal of Information Systems, 25*, 77-89. https://doi.org/10.1057/ejis.2014.36

Venkatesh, V., Aloysius, J. A., Hoehle, H., & Burton, S. (2017). Design and Evaluation of Auto-Id Enabled Shopping Assistance Artifacts in Customers' Mobile Phones. *MIS Quarterly, 41*(1), 83-114. https://www.jstor.org/stable/10.2307/26629638

Venkatesh, V., Morris, M. G., Davi, G. B., & Davis, F. D. (2003). User Acceptance of Information Technology: Toward a Unified View. *MIS Quarterly, 27*(3), 425-478. https://doi.org/10.2307/300336540

Vilone, G., & Longo, L. (2021). Notions of explainability and evaluation approaches for explainable artificial intelligence. *Information Fusion, 76*, 89-106. https://doi.org/10.1016/j.intfus.2021.05.009

Vilone, G., & Longo, L. (2022). A Novel Human-Centred Evaluation Approach and and Argument-Based Method for Explainable Artificial Intelligence. In Maglogiannis, I., Iliadis, L., Macintyre, J., Cortez, P. (Eds.), *Artificial Intelligence Applications and Innovations*, *AIAI 2022, IFIP Advances in Information and Communication Technology,* vol. 646 (pp. 447-460). Springer, Cham. https://doi.org/10.1007/978-3-031-08333-4_36

Vitharana, P., Zahedi, M. F., & Jain, H. K. (2016). Enhancing Analysts' Mental Models or Improving Requirements Elicitation: A Two-stage Theoretical Framework and Empirical Results. *Journal of the Association for Information Systems, 17*(12), 804-840. https://doi.org/10.17705/1jais.00444

van der Waa, J., Nieuwburg, E., Cremers, A., & Neerincx, M. (2021). Evaluating XAI: A comparison of rule-based and example-based explanations. *Artificial Intelligence, 291*, Article 103404. https://doi.org/10.1016/j.artint.2020.103404

Walls, J. G., Widmeyer, G. R., El Sawy, O. A. (1992). Building and Information System Design Theory for Vigilant EIS. *Information Systems Research, 3*(1), 36-59. https://doi.org/10.1287/isre.3.1.36

Walls, J., Widmeyer, G. R., & El Sawy, O. A. (2004). Assessing Information System Design Theory in Perspective: How Useful Was our 1992 Initial Rendition? *Journal of Information Technology Theory and Application, 6*(2), 43-58. https://aisel.aisnet.org/jitta/vol6/iss2/6

Wang, W., & Benbasat, I. (2009). Interactive Decision Aids for Consumer Decision making in E-Commerce: The Influence of Perceived Strategy Restrictiveness. *MIS Quarterly, 33*(2), 293-320. https://doi.org/10.2307/20650293

Wang, R., & Sundar, S. S. (2018). How does Parallax Scrolling Influence User Experience? A Test of TIME (Theory of Interactive Media Effects). *International Journal of Human-Computer Interaction, 34*(6), 533-543. https://doi.org/10.1080/10447318.2017.1373457

Wang, D., Yang, Q., Abdul, A., & Lim, B. Y. (2019). Designing Theory-Driven User-Centric Explainable AI. *CHI'19: Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems,* 1-15. https://doi.org/10.1145/3290605.3300831

Waters, D. (2011). *Quantitative Methods for Business*. Pearson Education Limited.

Weber, P., Carl, K. V., & Hinz, O. (2023). Applications of Explainable Artificial Intelligence in Finance – a systematic review of Finance, Information Systems, and Computer Science Literature. *Management Review Quarterly*, 1-41. https://doi.org/10.1007/s11301-023-00320-0

Webster, J., & Watson, R. T. (2002). Analyzing the Past to Prepare for the Future: Writing a Literature Review. *MIS Quarterly, 26*(2), xiii-xxiii.

Weitz, K., Schiller, D., Schlagwoski, R., Huber, T., & André, E. (2021). "Let me explain!": exploring the potential of virtual agents in explainable AI interaction design. *Journal of Multimodal User Interfaces, 15*, 87-98. https://doi.org/10.1007/s12193-020-00332-0

Wells, L., & Bednarz, T. (2021). Explainable AI and Reinforcement Learning – A Systematic Review of Current Approaches and Trends. *Frontiers in Artificial Intelligence, 4*, Article 550030. https://doi.org/10.3389/frai.2021.550030

Winter, R. (2008). Design science research in Europe. *European Journal of Information Systems, 17*, 470-475. https://doi.org/10.1057/ejis.2008.44

Wolfswinkel, J. F., Furtmueller, E., & Wilderom, C. P. M. (2013). Using grounded theory as a method for rigorously reviewing literature. *European Journal of Information Systems, 22*, 44-55. https://doi.org/10.1057/ejis.2011.51

Xu, W. (2019). Toward human-centered AI: a perspective from human-computer interaction. *Interactions, 26*(4), 42-46. https://doi.org/10.1145/3328485

Ye, L. R., & Johnson, P. E. (1995). The impact of explanation facilities on user acceptance of expert systems advice. *MIS Quarterly, 19*(2), 157-172. https://doi.org/10.2307/249686

Yigitcanlar, T., Degirmenci, K., & Inkinen, T. (2022). Drivers behind the public perception of artificial intelligence: insights from major Australian cities. *AI & Society*, 1-21. https://doi.org/10.1007/s00146-022-01566-0

Yuan, C. W. T., Bi, N., Lin, Y.-F., & Tseng, Y.-H. (2023). Contextualizing User Perceptions about Biases for Human-Centered Explainable Artificial Intelligence. *CHI'23: Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems,* 1-15. https://doi.org/10.1145/3544548.3580945

Zacharias, J., von Zahn, M., Chen, J., & Hinz, O. (2022). Designing a feature selection method based on explainable artificial intelligence. *Electronic Markets*, 1-26. https://doi.org/10.1007/s12525-022-00608-1

Zhang, P., Carey, J., Te'eni, D., & Tremaine, M. (2005). Integrating Human-Computer Interaction Development into the Systems Development Life Cycle: A Methodology. *Communications of the Association for Information Systems, 15*, 512-543. https://doi.org/10.17705/1CAIS.01529

Zimmermann, R., Mora, D., Cirqueira, D., Helfert, M., Bezbradica, M., Werth, D., Weitzl, W. J., Riedl, R., & Auinger, A. (2023). Enhancing brick-and-mortar store shopping experience with an augmented reality shopping assistant application using personalized recommendations and explainable artificial intelligence. *Journal of Research in Interactive Marketing, 17*(2), 273-298. https://doi.org/10.1108/JRIM-09-2021-0237

8 Anhang

8.1 Erklärung §10 (3) Promotionsordnung

Ich versichere an Eides statt, dass ich die eingereichte Dissertation „Human-Centered Design and Evaluation of Explanation User Interfaces – A Design Science Research Perspective" selbstständig verfasst habe. Anderer als der von mir angegebenen Quellen und Hilfsmittel habe ich mich nicht bedient. Alle wörtlich oder sinngemäß den Schriften anderer Autor:innen entnommen Stellen habe ich kenntlich gemacht. Diese Dissertation wurde anderweitig noch nicht als Prüfungsarbeit vorgelegt.

26. April 2024, Enrico Bunde

Datum, Name

8.2 Articles Published in the Cumulative Dissertation

Article 1: Explainable Artificial Intelligence: Objectives, Stakeholders, and Future Research Opportunities

Article Information

Title	Explainable Artificial Intelligence: Objectives, Stakeholders, and Future Research Opportunities
Year	2022
Outlet	Information Systems Management (VHB: C), 39:1, 53-63
Type	Journal Article
Status	Published
Full text available at	https://doi.org/10.1080/10580530.2020.1849465
License information	This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives License (http://creativecommons.org/licenses/by-ncnd/4.0/).

Co-Authorship

Author	**Location**	**Address**	**E-Mail**	**Contribution (CRediT)**
Christian Meske	Ruhr-Universität Bochum	Universitätsstr. 150, 44801 Bochum, Germany	christian.meske@rub.de	Conceptualization (lead), Data Curation (equal), Funding Acquisition (lead), Investigation (equal), Methodology (lead), Project Administration (lead), Supervision (lead), Writing – Original Draft Preparation (equal), Writing – Review & Editing (lead)
Enrico Bunde	Ruhr-Universität Bochum	Universitätsstr. 150, 44801 Bochum, Germany	enrico.bunde@rub.de	Conceptualization (support), Data Curation (lead), Funding Acquisition (-), Investigation (lead), Methodology (support), Project Administration (support), Supervision (support), Writing – Original Draft Preparation (equal), Writing – Review & Editing (support)
Martin Gersch	Freie Universität Berlin	Garystr. 21, 14195 Berlin	martin.gersch@fu-berlin.de	Conceptualization (support), Data Curation (support), Funding Acquisition (-), Investigation (equal), Methodology (support), Project Administration (support), Supervision (support), Writing – Original Draft Preparation (support), Writing – Review & Editing (support)

Johannes Schneider	Universität Liechtenstein	Fürst-Franz-Josef-Straße, 9490 Vaduz	johannes.schneider@uni.li	Conceptualization (support), Data Curation (support), Funding Acquisition (-), Investigation (equal), Methodology (support), Project Administration (support), Supervision (support), Writing – Original Draft Preparation (support), Writing – Review & Editing (support)

Alternative Usage

To the best of my knowledge, parts of this paper have not yet been used in other doctoral qualification procedures.

26. April 2024

Date

Article 2: Improving Customers' Decision Making on Blackboxed Multimodal Platforms – A Design Science Approach

Article Information

Title	Improving Customers' Decision Making on Blackboxed Multimodal Platforms – A Design Science Approach
Year	2020
Outlet	Proceedings of the International Conference on Information Systems (VHB: A) , pp. 1-9
Type	Conference Article, Short Paper
Status	Published
Full text available at	https://aisel.aisnet.org/icis2020/digital_commerce/digital_commerce/12

Co-Authorship

Author	Location	Address	E-Mail	Contribution (CRediT)
Christian Meske	Ruhr-Universität Bochum	Universitätsstr. 150, 44801 Bochum, Germany	christian.meske@rub.de	Conceptualization (equal), Data Curation (support), Funding Acquisition (-), Investigation (support), Methodology (support), Project Administration (lead), Supervision (lead), Writing – Original Draft Preparation (equal), Writing – Review & Editing (support)
Enrico Bunde	Ruhr-Universität Bochum	Universitätsstr. 150, 44801 Bochum, Germany	enrico.bunde@rub.de	Conceptualization (equal), Data Curation (lead), Funding Acquisition (-), Investigation (lead), Methodology (lead), Project Administration (support), Supervision (support), Writing – Original Draft Preparation (lead), Writing – Review & Editing (lead)
Jan Fabian Ehmke	Universität Wien	Universitätsring 1, 1010 Wien	jan.ehmke@univie.ac.at	Conceptualization (equal), Data Curation (support), Funding Acquisition (-), Investigation (support), Methodology (support), Project Administration (support), Supervision (support), Writing – Original Draft Preparation (equal), Writing – Review & Editing (support)

Alternative Usage

To the best of my knowledge, parts of this paper have not yet been used in other doctoral qualification procedures.

26. April 2024

Date

Article 3: Fake or Credible? Towards Designing Services to Support Users' Credibility Assessment of News Content

Article Information

Title	Fake or Credible? Towards Designing Services to Support Users' Credibility Assessment of News Content
Year	2022
Outlet	Proceedings of the Hawaii International Conference on Systems Sciences (VHB: C), pp. 1883-1892
Type	Conference Article
Status	Published
Full text available at	http://hdl.handle.net/10125/79569

Co-Authorship

Author	Location	Address	E-Mail	Contribution (CRediT)
Enrico Bunde	Ruhr-Universität Bochum	Universitätsstr. 150, 44801 Bochum, Germany	enrico.bunde@rub.de	Conceptualization (equal), Data Curation (lead), Funding Acquisition (-), Investigation (lead), Methodology (equal), Project Administration (equal), Supervision (support), Writing – Original Draft Preparation (lead), Writing – Review & Editing (lead)
Niklas Kühl	Universität Bayreuth	Universitätsstraße 30, 95447 Bayreuth	kuehl@uni-bayreuth.de	Conceptualization (equal), Data Curation (support), Funding Acquisition (-), Investigation (equal), Methodology (equal), Project Administration (equal), Supervision (support), Writing – Original Draft Preparation (support), Writing – Review & Editing (equal)
Christian Meske	Ruhr-Universität Bochum	Universitätsstr. 150, 44801 Bochum, Germany	christian.meske@rub.de	Conceptualization (equal), Data Curation (support), Funding Acquisition (-), Investigation (equal), Methodology (equal), Project Administration (equal), Supervision (lead), Writing – Original Draft Preparation (support), Writing – Review & Editing (equal)

Alternative Usage

To the best of my knowledge, parts of this paper have not yet been used in other doctoral qualification procedures.

26. April 2024

Date

Article 4: Design Principles for User Interfaces in AI-Based Decision Support Systems: The Case of Explainable Hate Speech Detection

Article Information

Title	Design Principles for User Interfaces in AI-Based Decision Support Systems: The Case of Explainable Hate Speech Detection
Year	2023
Outlet	Information Systems Frontiers (VHB: B), 25:743-773
Type	Journal Article
Status	Published
Full text available at	https://doi.org/10.1007/s10796-021-10234-5
License information	This is an Open Access article distributed under the terms of the Creative Commons Attribution 4.0 International License (https://creativecommons.org/licenses/by/4.0/).

Co-Authorship

Author	Location	Address	E-Mail	Contribution (CRediT)
Christian Meske	Ruhr-Universität Bochum	Universitätsstr. 150, 44801 Bochum, Germany	christian.meske@rub.de	Conceptualization (equal), Data Curation (support), Funding Acquisition (lead), Investigation (equal), Methodology (support), Project Administration (equal), Supervision (lead), Writing – Original Draft Preparation (equal), Writing – Review & Editing (equal)
Enrico Bunde	Ruhr-Universität Bochum	Universitätsstr. 150, 44801 Bochum, Germany	enrico.bunde@rub.de	Conceptualization (equal), Data Curation (lead), Funding Acquisition (support), Investigation (equal), Methodology (lead), Project Administration (equal), Supervision (support), Writing – Original Draft Preparation (equal), Writing – Review & Editing (equal)

Alternative Usage

To the best of my knowledge, parts of this paper have not yet been used in other doctoral qualification procedures.

26. April 2024

Date

Article 5: Giving DIAnA more TIME – Guidance for the Design of XAI-Based Medical Decision Support Systems

Article Information

Title	Giving DIAnA more TIME – Guidance for the Design of XAI-Based Medical Decision Support Systems
Year	2023
Outlet	International Conference on Design Science Research in Information Systems and Technology (DESRIST 2023). Lecture Notes of Computer Science (VHB: C), pp. 107-122
Type	Full Paper
Status	Published
Full text available at	https://doi.org/10.1007/978-3-031-32808-4_7

Co-Authorship

Author	Location	Address	E-Mail	Contribution (CRediT)
Enrico Bunde	Ruhr-Universität Bochum	Universitätsstr. 150, 44801 Bochum, Germany	enrico.bunde@rub.de	Conceptualization (lead), Data Curation (lead), Funding Acquisition (support), Investigation (equal), Methodology (lead), Project Administration (equal), Supervision (equal), Writing – Original Draft Preparation (equal), Writing – Review & Editing (equal)
Daniel Eisenhardt	Ruhr-Universität Bochum	Universitätsstr. 150, 44801 Bochum, Germany	daniel.eisenhardt@rub.de	Conceptualization (support), Data Curation (support), Funding Acquisition (support), Investigation (equal), Methodology (support), Project Administration (equal), Supervision (equal), Writing – Original Draft Preparation (equal), Writing – Review & Editing (equal)
Daniel Sonntag	Deutsches Forschungszentrum für Künstliche Intelligenz GmbH	Trippstadter Str. 122, 67663 Kauserslautern	sonntag@dfki.de	Conceptualization (support), Data Curation (support), Funding Acquisition (equal), Investigation (equal), Methodology (support), Project Administration (equal), Supervision (equal), Writing – Original Draft Preparation (equal), Writing – Review & Editing (equal)

Hans-Jürgen Profitlich	Deutsches Forschungszentrum für Künstliche Intelligenz GmbH	Trippstadter Str. 122, 67663 Kauserslautern	profitlich@dfki.de	Conceptualization (), Data Curation (support), Funding Acquisition (equal), Investigation (equal), Methodology (), Project Administration (), Supervision (), Writing – Original Draft Preparation (equal), Writing – Review & Editing (equal)
Christian Meske	Ruhr-Universität Bochum	Universitätsstr. 150, 44801 Bochum, Germany	christian.meske@rub.de	Conceptualization (support), Data Curation (support), Funding Acquisition (equal), Investigation (equal), Methodology (support), Project Administration (equal), Supervision (lead), Writing – Original Draft Preparation (equal), Writing – Review & Editing (equal)

Alternative Usage

To the best of my knowledge, parts of this paper have not yet been used in other doctoral qualification procedures.

26. April 2024
Date

Article 6: The Design and Perception of Customization Features in Explanation User Interfaces

Article Information

Title	The Design and Perception of Customization Features in Explanation User Interfaces
Year	2023
Outlet	Human-Computer Interaction (VHB: C)
Type	Journal Article
Status	Submitted

Co-Authorship

Author	Location	Address	E-Mail	Contribution (CRediT)
Enrico Bunde	Ruhr-Universität Bochum	Universitätsstr. 150, 44801 Bochum, Germany	enrico.bunde@rub.de	Conceptualization (lead), Data Curation (lead), Funding Acquisition (lead), Investigation (lead), Methodology (lead), Project Administration (lead), Supervision (lead), Writing – Original Draft Preparation (lead), Writing – Review & Editing (lead)

Alternative Usage

To the best of my knowledge, parts of this paper have not yet been used in other doctoral qualification procedures.

26. April 2024

Date

The Design and Perception of Customization Features in Explanation User Interfaces

Abstract: Explanation user interfaces (XUIs) represent an essential entry point into a meaningful human-computer interaction experience. Their design is a critical success factor for the AI system and can, for example, influence the acceptance of the system or its usability through different design configurations. Despite the great and growing interest in XAI, the research stream is characterized by a lack of design knowledge for the design of XUIs as well as human-centered user studies on the influence of explanations on people. This article presents a design science research (DSR) project that addresses these two research gaps. A focus was placed on customization, and design principles for customization features in XUIs were developed. The design principles instantiated in an interactive XUI were examined in two online experiments. In the first experiment, a between-subject experiment was carried out in the context of human resources with subjects who had experience in this context. The extent to which the developed and instantiated customization features influence the perceived explanation quality was examined. The perceived explanation quality was operationalized by the constructs of perceived interactivity, perceived interestingness, perceived informativeness, and the explanation satisfaction scale. The results showed a positive significant influence of the customization features on the perceived explanation quality. In the second online experiment, user engagement and satisfaction were examined in the context of XUI interaction. For this purpose, a structural model was developed and tested. The results showed that the constructs of perceived interactivity and customization positively predict user engagement, although such an effect was not identified for perceived interactivity. User engagement positively predicts satisfaction. In the final step, the design principles were evaluated with regard to their reusability, which was quantitatively implemented with experienced user interface or user experience designers. The evaluation revealed a positive perception of the design principles, i.e., high reusability and an acceptance and recommendation rate. In summary, this article presents diverse knowledge contributions for the design of XUIs with customization features and their human-centered evaluation, i.e., insights regarding human-XUI interaction.

Keywords: *Design Science Research, Human-Computer Interaction, Explanation User Interfaces, Customization Features, Explainable Artificial Intelligence*

1. Introduction

User interfaces (UIs) are the entry point for meaningful knowledge discovery when humans interact with artificial intelligence (AI) systems (Song et al., 2020). Their design is essential since it influences the adoption of AI systems (Ammar & Shaban-Nejad, 2020). This research project investigates a particular form of UIs, the explanation UIs (XUIs), because many state-of-the-art AI methods whose results are displayed in the UI are opaque and, therefore, often referred to as black box (Arrieta et al., 2020). This black box problem poses various challenges during the development, use, or consequences of AI and its spread, for example, regarding accountability, liability, transparency, or trustworthiness (Adadi & Berrada, 2018). As a solution, scholars from the research stream of explainable AI (XAI) introduce techniques to generate explanations either for the outputs (e.g., classifications) or for the inner learning procedures (e.g., learned weights in neural networks) of AI systems (Arrieta et al., 2020). For example, it is possible to visualize the AI model and explanations in UIs, which improves understandability, explainability, and interpretability (Alicioglu & Sun, 2021). With an adequate design, it is further possible to positively influence the perception of the UI regarding informativeness and trustworthiness (Meske & Bunde, 2023).

XUIs can be characterized by presenting essential information about the AI output, such as the label for a predicted class and the probability in percentage for the output (Bunde, 2021). In addition, one or more XAI methods can be used to integrate different explanations for the AI output, whereby, for example, local and global explanations can be combined (Adadi & Berrada, 2018). Since XUIs are an essential interface between humans on the one hand and the AI anchored in the system on the other, they can significantly impact how humans, including potential users or customers, perceive the system and ideally be supported in a task at hand (Gunning et al., 2019). Due to the high relevance of XUIs, two topics are gaining importance. The first is designing XUIs to exploit potential benefits, including increased task performance, usefulness, or comprehensibility (Gunning & Aha, 2019). Second, it is highly relevant to understand how the targeted stakeholders, such as system end users, perceive the XUI and what psychological effects arise during the interaction. Hence, the human-centered evaluation of XUIs becomes more relevant (Vilone & Longo, 2021). Human-centered evaluations can help measure how specific design features or explanation types influence dimensions such as trust or interactivity (Leichtmann et al., 2023; Vilone & Longo, 2021) or identify optimization potential for the design (Meske & Bunde, 2023). However, little knowledge of the guidance for designing XUIs exists, accompanied by a notable lack of user studies (van der Waa et al., 2020; Wells & Bednarz, 2021).

Especially the involvement of relevant stakeholders when designing and evaluating explanations or XUIs is vital since their characteristics, needs, and expectations should be considered (Gunning & Aha, 2019; Langer et al., 2021). A broad selection of dimensions, characteristics, and notions of explainability can influence the perception of explanations, including completeness, effectiveness, informativeness, or understandability (Arrieta et al., 2020; Vilone & Longo, 2021). All these aspects of explainability can influence the design of XUIs and users' perceptions. One design characteristic that has yet to receive much attention so far is customization. Customization is well-established in human-computer interaction and enables users to independently customize the information they receive (Sundar et al., 2015). Users may change the appearance, rearrange components, or manipulate data presented in UIs (Bolin et al., 2005). Customization can lead to versatile, positive effects, including improved enjoyment (Bailey et al., 2009), enhanced user loyalty, or a more effective interaction with the UI (Jorritsma et al., 2015; Teng, 2010). Personalization is a related concept to customization, which is already established in the research area of XAI (Meske et al., 2022; Sundar et al., 2015). The concepts of customization and personalization can be differentiated since customization is user-tailored, and personalization describes the content as system-tailored (Sundar et al., 2012). When transferring the concepts into the realm of

XAI, personalized explanations are tailored to a specific explainee (Meske et al., 2022), and customization provides the user with features to adjust elements independently (Sundar et al., 2015). This study aims to conceptualize customization for the design of XUIs by introducing design knowledge, including design principles, an instantiated artifact, and two quantitative human-centered evaluations that provide insights into the effects of customization in XUIs on users. The following three research questions guided the research endeavor:

> ***RQ1:*** *Which design principles can be established to guide the design of customization features in XUIs?*
>
> ***RQ2:*** *How do customization features in XUIs influence the perceived explanation quality?*
>
> ***RQ3:*** *How do XUIs with customization features influence user engagement and, ultimately, the satisfaction of users?*

A design science research (DSR) project was conducted to answer the research questions. The process of Kuechler and Vaishnavi (2012) was followed as an overarching framework. Through a literature review, the knowledge base for the project was developed (vom Brocke et al., 2015). Based on the insights gained, design requirements and associated evaluation metrics were derived, representing the applied goodness criteria for the solution acceptance (vom Brocke et al., 2020). The design requirements are addressed by design principles formalized according to Gregor et al. (2020). In addition to the design principles, exemplary design features were derived to instantiate them in a prototypical web-based XUI with customization features (Seidel et al., 2018). Motivated by the lack of human-centered user studies in the context of XAI and XUIs (Wells & Bednarz, 2021), two online experimental studies were conducted. Therefore, the Human Risk & Effectiveness evaluation strategy for DSR projects was followed since the major design risk was anticipated to be social or user-oriented (Venable et al., 2016). In the first evaluation, a between-subject experiment was conducted to compare the influence of customization features in XUIs on the perceived explanation quality. One group interacted with an XUI offering customization features and another interacted with an XUI without customization features. The second experiment explored how XUIs with customization features influence user engagement and, ultimately, users' satisfaction with the XUI.

The statistical analysis of the first experimental study showed that the customization features in XUIs significantly positively affected the perceived explanation quality compared to an XUI without customization features. The group that had access to the customization features also achieved a higher accuracy during the classification task in the experiment. The statistical analysis of the structural equation model (SEM) in the second experimental study uncovered how the designed XUI influenced user engagement through perceived interactivity and perceived customization and how user engagement influences users' satisfaction with the XUI. In the last step, the proposed design principles were quantitatively evaluated by experienced practitioners regarding their reusability and were rated positively. Overall, 52 participants stated that they would use the design principles for a suitable project, and 53 participants said that they would recommend the design principles for a suitable project to colleagues. Consequently, the DSR project has generated multifaceted knowledge contributions to the research on designing XUIs and their influence on users. Consequently, the contribution and output knowledge of the DSR project is versatile (vom Brocke & Maedche, 2019).

The remainder of the article is structured as follows. In Section 2, the related work and conceptual background are presented, customization is conceptualized for XUIs, and the overview of user-oriented studies and the influence and perception of XAI by end users is provided. The research approach and methods used are described in Section 3. The DSR process steps suggestion and development are summarized in Section 4, where design requirements and associated goodness criteria are established, as well as the derivation of the design principles and their instantiation. Section 5 describes the evaluation, which includes both experimental studies in the form of online experiments, the statistical analysis, and the results of the reusability evaluation of the proposed design knowledge. The findings, theoretical and practical implications, limitations, and future research are discussed in Section 6. The last Section 7 concludes the article.

2. Related Work and Conceptual Background

2.1 Explanation User Interfaces

The design of UIs is a critical aspect of the success of AI systems that can also influence the user experience and, ultimately, human decision-making (Dudley & Kristensson, 2018; Ferreira & Monteiro, 2020). Well-designed XUIs should support the user in understanding the explanation of the decision support for the task at hand (Keneni et al., 2019). XUIs commonly present information about the AI model itself and use one or more methods to generate explanations for providing users with meaningful information (Chromik & Butz, 2021). An appropriate design of XUIs significantly influences how users analyze, experience, understand, and interpret the outputs and let users verify them (Bradley et al., 2009; Füßl et al., 2023). An appropriate design can furthermore aid users in understanding black box AI systems and can affect different outcomes, including the ability of users to trust or debug an AI model (Adadi & Berrada, 2018; Haque et al., 2023). However, unique challenges arise when designing XUIs (Dudley & Kristensson, 2018). Each design decision can have profound implications and effects on the user experience and interaction with XUIs. For example, Zhang and Lim (2022) uncovered how relatable explanations, such as counterfactual samples with semantic cues, can improve decision quality without sacrificing time. Thus, many nuances must be considered when designing XUIs, which leads to another challenge. The challenge lies in the comprehensible presentation of explanations in different formats and styles, including explanations in text, by visual means, or by using examples to explain the rationale for the outputs of AI models (Arrieta et al., 2020).

Explanations can be generated through different XAI methods and either explain one specific AI output in the form of local explanations or the behavior of a whole AI model in the form of global explanations (Adadi & Berrada, 2018). Explanations that reveal AI models' behavior or underlying decision-making procedures can be generated through various visualization techniques, leading to more understandable, explainable, and interpretable explanations for end users (Alicioglu & Sun, 2022). For example, Alsallakh et al. (2014) developed an interactive exploration environment that presents information about classification results, class probabilities, and features, though no evaluation with users was conducted. In the healthcare domain and the study of adverse childhood experiences, Ammar and Shaban-Nejad (2020) showed how explainable features can enhance the ability of healthcare practitioners to comprehend and explain their decisions when using XAI systems. XUIs can also enable users to explore the provided information, correct the underlying AI model, or add background information (Gunning & Aha, 2019). Explanations in UIs have also been shown to lower the perceived cognitive effort and improve the perceived informativeness and trustworthiness compared to UIs without explanations (Meske & Bunde, 2023). Providing explanations can also lead to a rich information basis and support the decision-making process (Zimmermann et al., 2022). Many dimensions, such as the perception of XUIs by end users or the interaction experience, can be significantly influenced by their design, and there are many possible design options.

Therein lies another challenge regarding the design and evaluation of XUIs. Research focusing on the design and perception of XUIs is still described as scarce, and systematic knowledge is missing (Chromik & Butz, 2021; Füßl et al., 2023). In many research projects that deal with the design or evaluation of XUIs, no design knowledge, such as design principles, is developed. These include XUIs for contexts and domains, such as explainable learning systems, explainable question-answering systems, or explainable clinical decision support systems (Gunning & Aha, 2019; Panigutti et al., 2023). In these studies, however, no design knowledge was formalized and made available to a broader audience. There are also research projects that have formalized design knowledge in the form of design patterns (Schoonderwoerd et al., 2021) or developed frameworks (Barda et al., 2020). The design principles introduced here fit into this stream of research. In addition to the lack of design knowledge for XUIs, the research area of XAI is also characterized by a lack of user studies (van der Waa et al., 2020; Wells & Bednarz, 2021). In particular, human-centered evaluations are suitable for involving relevant stakeholders in the evaluation procedure (Vilone & Longo, 2022). The goals pursued by XAI and specific XUI design configurations can be varied. For example, one could pursue the goal that users develop increased trust in the decision support provided by AI (Meske & Bunde, 2023) or to improve task performance in a task (Leichtmann et al., 2023). In order to be able to define the relevant dimensions and goals, the involvement of relevant stakeholders with their individual expectations, experiences, assumptions, and cognitive biases is essential (Langer et al., 2021). Through the two experiential studies in the form of online experiments, the DSR project also provides relevant insights regarding the perception of XUIs with customization features by relevant stakeholders concerning the perceived explanation quality and the relationship between user engagement and satisfaction.

2.2 Customization in Human-Computer Interaction and Conceptualization for Explanation User Interfaces

Customization is a critical concept well-established in human-computer interaction and allows users to independently customize the information they receive (Sundar & Marathe, 2010). It is also described as a means to an end, which is not typically designed as an end in itself since it is assumed to be a secondary activity supporting users to tackle the task at hand (Marathe & Sundar, 2011). Moreover, it is considered a desirable attribute of media technologies, requiring users to exercise choice (Kang & Sundar, 2013) actively. Customization can be used to satisfy the needs of stakeholders with varying cognitive styles (Ku et al., 2016). The customization procedure is associated with users' possibilities to change the appearance, rearrange components, and insert or remove data from the UI (Bolin et al., 2005). For example, the customization options for plots in a UI could be the adjustment of colors, various values, or an option to download the results (Lopez-Giraldez & Townsend, 2011). Often, users can choose customization options in UIs through predefined options (Macias & Paterno, 2008). It is possible to differentiate between varying types of customizations and two of these types are adapted in this study (Sundar et al., 2015): (i) cosmetic customization for the adjustment of the appearance of the XUI and (ii) functional customization to modify utility aspects of the XUI.

The concept of personalization is already established in the research field of XAI. Therefore, this research aims to conceptualize customization as an additional design characteristic of XUIs. Consequently, it is essential to differentiate between these two concepts. Personalization is a system-initiated adaptation of content, whereas customization is a user-initiated adaptation (Sundar & Marathe, 2010). Consequently, individualization through personalization is a form of adaptivity to tailor the content, structure, or presentation automatically to an individual user. In contrast, customization provides a form of adaptability to let users modify versatile dimensions of the content themselves (Ku et al., 2016). It is also possible to design a hybrid approach of personalization and customization, where personalization is initiated by the system and approved by the user (Vitale et al., 2020). The following Figure 1 provides a schematic overview of personalization and customization in the context of XAI. Personalization is shown on the left. For example, data on the usage behavior of a user can be used to generate personalized explanations. If a specific user is offered various explanation types, and he always uses a specific type of explanation, such knowledge should flow into the design of a personalized explanation. On the other hand, customization is shown, whereby various customization features can be offered in an XUI. These can be predefined, for example, and can be called up by the user depending on their preferences. There is also a hybrid approach.

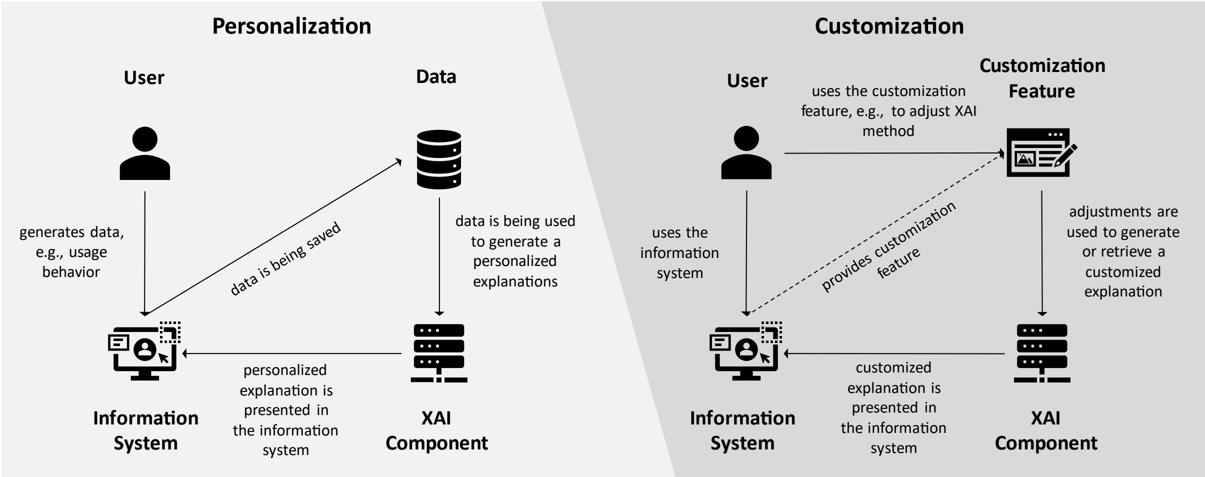


Figure 1. Schematic comparison of personalization and customization in the context of XAI.

Research investigating customization in the design of systems, services, or UIs has provided empirical insights regarding the positive effects for users. Studies have uncovered that customization creates strong emotional and cognitive appeal to users (Sundar et al., 2012). In addition, customization can reduce errors and increase user acceptance in human-computer interactions (Burkolter et al., 2014). The availability of customization features can influence subjective feelings and psychological indicators of emotion, ultimately influencing end users' enjoyment (Bailey et al., 2009). Customization can also affect a behavioral level, observed in users' browsing activities (Kalyanaraman & Sundar, 2006). By enabling users to customize services to themselves, an enhanced level of self-efficacy beliefs and perceived fit of the environment with their individual wants and needs can be achieved (Kang & Lee, 2015). Furthermore, customization can be useful for learning performance and perceptions (Ku et al., 2016). In a suitable context, customization can also enhance user loyalty, directly or indirectly improve immersion satisfaction, and enable more effective interaction with UIs (Jorritsma et al., 2015; Teng, 2010). Concerning task difficulty, Rivera (2005) showed that customization features in UIs can lead to a reduced perceived workload. In summary, customization can be a design characteristic that positively influences the perception of a system and supports users in their work with a system.

2.3 The Influence and Perception of Explainable Artificial Intelligence

The relevance and lack of user evaluations for XAI are emphasized throughout the XAI literature (e.g., van der Waa et al., 2021; Wells & Bednarz, 2021). The reasons for the high relevance are manifold. For example, versatile stakeholder groups are interested in AI-generated explanations and have varying information needs, prior knowledge, or backgrounds (Arrieta et al., 2021; Meske et al., 2022). In addition, the need for explanations may differ in various application scenarios depending on the criticality of the AI use, its possible consequences, and the used (X)AI methods (Adadi & Berrata, 2018; Vilone & Longo, 2021). To meet these demands, user- or human-centered approaches in designing and evaluating XAI systems are emerging (Barda et al., 2020; Schoonderwoerd et al., 2021). Table 1 presents an overview of the reviewed literature that evaluated different dimensions well-established in the literature of XAI. Since XAI is an interdisciplinary research subject, the overview aims to be representative and typify the large body of research on XAI (vom Brocke et al., 2015). The table is divided into three XAI categories: XAI methods, XAI systems, and XUIs. XAI methods include research projects in which the authors have introduced a new XAI method or further developed an existing one.

User studies are often carried out for a human-centered evaluation. XAI systems include research projects in which the authors have designed and human-centeredly evaluated an XAI system. The XUIs category includes works like this, which explicitly investigate UI design and user experience in an XAI context.

XAI methods used to generate explanations are typically introduced in the computer science discipline. User studies are commonly a part of the evaluation and aim to provide evidence for the added value of the proposed XAI method. For example, Ribeiro et al. (2018) introduced the XAI method anchors. In their user study, they found that anchors enable users to predict the behavior of an AI model with less effort and high precision. Kaur et al. (2020) took an interesting approach and evaluated XAI methods in a user study with data scientists. They investigated how data scientists use such interpretability tools to uncover issues when developing and evaluating AI models. Vilone and Longo (2022) compared their proposed argumentation framework for explainability with a decision tree as a baseline. They measured a broad set of notions of explainability, including actionability, causality, cognitive relief, and comprehensibility. The evaluation proved that the argumentation-based approach led to higher measurements for the notions of explainability.

Concerning XAI systems, Knapic et al. (2021) developed a decision support system with explanation features in the medical domain. Their user study found that users perceive different explanations for the same decision-making scenario differently, for example, regarding how understandable the explanations were to users. Explanations are also crucial in decision-making scenarios with a high risk, like deciding whether mushrooms are edible or poisonous. Leichtmann et al. (2023) developed a prototype app for this use case. They found that users provided with explanations outperformed those without access to explanations. Moreover, explanations have led to better-calibrated trust levels. With a focus on feature selection, Zacharias et al. (2022) developed a software artifact and evaluated it with users. They measured effectiveness, usefulness, understandability, emotional trust, and satisfaction during the evaluation, which were rated positively.

Dikmen and Burns (2020) developed a prototypical XUI focusing on abstraction hierarchy-based XAI. They measured versatile XAI characteristics, including confidence, understandability, satisfaction, behavioral intention, or perceived learning. On a higher level, they found that abstraction hierarchy-based explanations helped participants learn about the process at hand and improved the perceived quality of explanations. Van der Waa et al. (2021) investigated different explanations in the decision-making scenario of diabetes self-management. Here, they evaluated the effects of explanations on dimensions including system understanding, persuasive power, and task performance. While designing and evaluating an XUI in hate speech detection for social media moderators, Meske and Bunde (2023) found a significant influence of explanations on perceived cognitive effort, informativeness, mental model, and trustworthiness.

Consequently, a broad range of design characteristics and notions of explainability were investigated. However, since there are many characteristics and notions of XAI, this overview supports the view that more user studies are necessary to advance the field of XAI (Vilone & Longo, 2021; Wells & Bednarz, 2021). In addition, the concept of customization is not well-established in the discipline of XAI, which further emphasizes the identified research gap.

Evaluation Approach	Object of Evaluation	Evaluated Characteristics and Notions of Explainability	Reference
Human-Centered Evaluation of XAI	XAI Methods	Cognitive Effort/ Effort, Mental Model/ Mental Fit, Task Performance	Ribeiro et al. (2018)
		Cognitive Load, Confidence, Task Performance	Kaur et al. (2020)
		Confidence, Reaction Time, Task Performance	Huysmans et al. (2011)

		Interpretability	Bau et al. (2017)
		Actionability, Causality, Cognitive Relief, Comprehensibility, Informativeness, Interestingness, Persuasion, Simplicity/ Simplification, Understanding/ Understandability	Vilone and Longo (2022)
	XAI Systems	Quality, Task Performance	Senoner et al. (2022)
		Reaction Time, Satisfaction, Understanding/ Understandability	Knapic et al. (2021)
		Usability/ Usefulness	Spinner et al. (2020)
		Trust/ Trustworthiness, Understanding/ Understandability	Lim et al. (2009)
		Learning, Task Performance, Trust/ Trustworthiness	Leichtmann et al. (2023)
		Confidence, Ease of Use, Effectiveness, Satisfaction, Trust/ Trustworthiness, Understanding/ Understandability, Usability/ Usefulness	Zacharias et al. (2022)
		Controllability, Internality, Stability	Ha et al. (2022)
	XUIs	Confidence, Satisfaction, Understanding/ Understandability	Dikmen and Burns (2020)
		Information Amount, Learning, Persuasion, Task Performance, Understanding/ Understandability	van der Waa et al. (2021)
		Behavioral Intention, Cognitive Effort/ Effort, Ease of Use, Informativeness, Mental Model/ Mental Fit, Trust/ Trustworthiness, Usability/ Usefulness	Meske and Bunde (2023)
		Confidence, Reaction Time, Task Performance	Alufaisan et al. (2021)
		Understanding/ Understandability	Kim et al. (2014)
		Usability/ Usefulness	Habers et al. (2010)
		Behavioral Intention, Ease of Use, Trust/ Trustworthiness, Usability/ Usefulness	Bunde (2021)
		Customization, Informativeness, Interactivity, Interestingness, Satisfaction, Task Performance, User Engagement	This study

Table 1. Design-oriented studies in the context of XUIs and their evaluation focus.

3. The Setting of the Design Science Research Project

This DSR project follows the process described by Kuechler and Vaishnavi (2012). The overarching goal is to introduce customization features for the design of XUIs supplemented by associated design principles and empirical knowledge regarding their influence on users. Figure 2 illustrates the DSR process, including the general process steps and activities. The process of Kuechler and Vaishnavi (2012) was run through twice.

The first step in the initial design cycle was to build a firm foundation for developing design knowledge and, therefore, advancing knowledge (Webster & Watson, 2002). The first step, awareness of problem, started with a literature review. In the step of the suggestion, design requirements, goodness criteria, design principles, and design features were derived, which were used to specify the design knowledge in an accessible form, adapting the anatomical lens for design principles introduced by Gregor et al.

(2020). During the step of development, the design principles and design features were instantiated to prove that they could be implemented in a prototypical system and to demonstrate their suitability for its intended purpose (Baskerville et al., 2018). The evaluation represents the next step to rigorously investigate the achievement of the designed and instantiated artifact concerning the identified design requirements (Sonnenberg & vom Brocke, 2012). The goal of the first evaluation was to explore the influence of customization features in XUIs on the perceived explanation quality, which was operationalized through the perceived informativeness, perceived interestingness, perceived interactivity, and explanation satisfaction scale. Therefore, a between-subjects experiment design with two groups was organized. The gathered data was statistically analyzed using R.

The first step of the second design cycle focused on the literature on the Theory of Interactive Media Effects. Based on the findings, the design knowledge was expanded in the second step, and the XUI design was revised accordingly in the third step. A human-centered quantitative evaluation was also carried out in the form of an online experiment in the evaluation in the second design cycle. It examined to what extent the dimensions of perceived interactivity, perceived interestingness, and perceived customization influence user engagement, which in turn affects satisfaction. For this purpose, hypotheses and an SEM were developed and analyzed using R and R Jamovi. In addition, the design principles were quantitatively evaluated by software developers. The reusability (Ivari et al., 2021) was examined with accessibility, importance, novelty and insightfulness, actability and appropriate guidance, and effectiveness. Ultimately, the generated insights from the DSR project are made accessible and communicated through this research manuscript (Peffers et al., 2007).

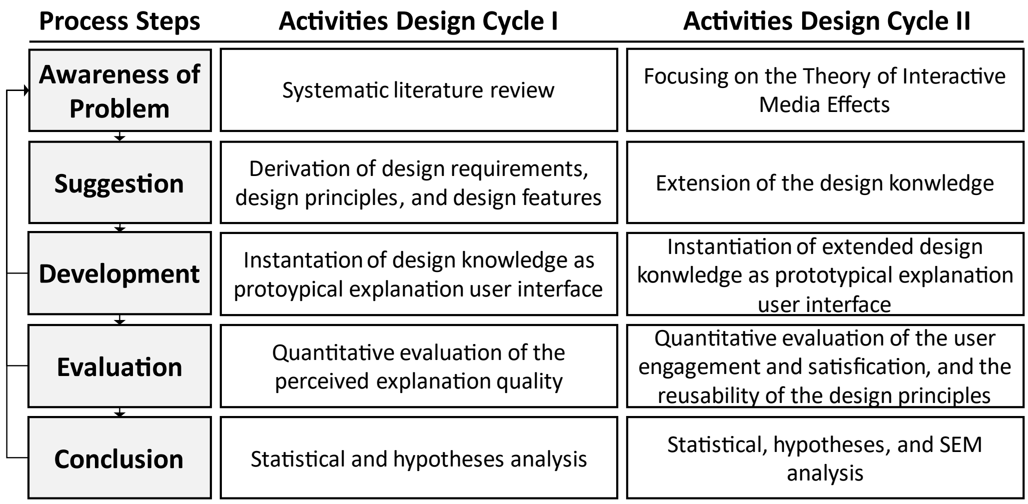


Figure 2. The adapted DSR process (based on Kuechler & Vaishnavi, 2012).

The evaluation in DSR aims to rigorously demonstrate the utility, quality, and efficacy of a designed artifact by using suitable, well-executed evaluation methods (Hevner et al., 2004). The overarching objective is to provide evidence for the usefulness of the designed artifact (Gregor & Hevner, 2013). The framework for evaluation in DSR (FEDS) was used for a rigorous evaluation of the proposed designed artifact and design principles. It provides evaluation strategies and a strategy choice process (Venable et al., 2016). The evaluation strategy of Human Risk & Effectiveness was chosen since the major design risk of the artifact was anticipated to be social or user-oriented (Venable et al., 2016, p. 82). The evaluations were organized as artificial, formative evaluations (Venable et al., 2016).

A web-based interactive XUI was developed for the first and second experimental studies. HTML, CSS, Bootstrap, Python, and JavaScript were used. A random forest was implemented as a machine learning (ML) model in the first and second experimental studies. Python and the sklearn (Pedregosa et al., 2011) library were used for this. In the first experimental study, a dataset was used in which the task was to classify the attrition risk for an anonymous person. It is a binary classification between low and high attrition risk (Kaggle, 2023a). In the second experimental study, a dataset was used to classify the salaries of anonymous individuals, which was also a binary classification task whereby a salary of less than or greater than $55,000 can be predicted (Kaggle, 2023b). Both datasets were identified on Kaggle. The random forest achieved a (weighted average) performance on the test data set for the classification of attrition risk: precision: 77%, recall: 83%, f1-score: 78%. On the second test data set for predicting salary, the following performance was achieved (weighted average): precision: 83%, recall 80%, f1 score: 82%.

The explanations were implemented with Python for both studies. Explanations for random examples from the respective test data set were used. In the first experimental study, the Python library was used for the XAI method local interpretable model agnostic explanations (LIME) (Ribeiro et al., 2016). For this purpose, the local explanations, i.e., the explanations for a concrete example, were incorporated into the XUI. In the second experimental study, a global explanation was added to the local explanation, i.e., an explanation that, for example, presents the relevant input features of the model. The global explanation was generated using the Python library sklearn with permutation feature importance and integrated into the XUI.

4. Suggestion and Development

4.1 Design Requirements for Customization Features of Explanation User Interfaces that Increase the Perceived Explanation Quality

Design requirements are essential for artifact-centric DSR projects, although different understandings and perspectives on this concept exist (Ivari, 2020). Within this DSR project, design requirements are used similarly to requirements in software engineering, which was done in prior DSR projects (Baskerville & Pries-Heje, 2010; Ivari, 2020). The design requirements are rooted in the investigated knowledge base in combination with theoretical concepts from the field of human-computer interaction (Meth et al., 2015). The overarching goal of the design requirements is to formulate requirements so that, if addressed correctly, they can positively influence individual nuances of the perceived quality of explanations. However, there are a variety of scales, constructs, and notions of explainability that can be used in evaluations. These include, for example, mental fit, efficiency, or soundness (Gunning et al., 2019; Vilone & Longo, 2021). For this study, the perceived interestingness, perceived informativeness, perceived interactivity, and the explanation satisfaction scale were operationalized to measure the perceived explanation quality. The perceived explanation quality was the focus of the evaluation in the first design cycle. The selected dimensions are established in research on XAI and human-computer interaction research. The design requirements that an XUI with customization features should meet are developed below.

Explanations are credited with increasing the perceived interestingness of AI systems (Minh et al., 2022). With this, visualizations are the most human-centered technique for designing visually exciting explanations (Adadi & Berrada, 2018). Visualizing AI outputs can lead to exciting ways of supporting the interpretation of opaque AI models (Arrieta et al., 2020). The interestingness of XUIs can be improved by providing customization features like “drill-down” or “zoom-in” functions, which support individual information needs and can minimize information overload (Barda et al., 2020). Prior studies have shown that customized content can lead to users being more involved since they perceive it as more interesting and essential (Kalyanaraman & Sundar, 2006). The level of interest and intrinsic interest can be even higher for users with a strong sense of self-efficacy when dealing with a specific activity (Bandura & Schunk, 1981; Kang & Lee, 2015). Moreover, users in prior experiments have described customization features as more interesting (Kleinsmith & Gillies, 2013). As a result, the following design requirement is established:

> **Design Requirement 1:** If an AI-based decision support system provides explanations, the system should provide the users with secondary activities that enable user-tailored customization of provided information in the XUI in order to enhance the user's perceived interestingness.

Explanations can significantly influence the perceived informativeness of decision aids (Al-Natour et al., 2022). Empirical evidence shows that UIs of decision aids that provide explanations achieve significantly higher perceived informativeness than UIs without explanations (Meske & Bunde, 2023). Moreover, by integrating visualizations of the output, such as the feature relevance in XUIs, it is possible to support the identification of informative and non-informative features (Alicioglu & Sun, 2022). In this context, it is vital to consider the individual information needs of the targeted stakeholders (Langer et al., 2021). By providing users with customization features, they can choose the appropriate amount of information they want or need (Barda et al., 2020). Therefore, customization features can avoid information overload, which can negatively affect the results of interactions with (X)AI systems, hindering the detection of faults (Poursabzi-Sangdeh et al., 2018). Moreover, customization can lead to an enhanced acceptance and a more critical examination of the provided information (Kang & Kim, 2020). As a result, the following design requirement is established:

> **Design Requirement 2:** If an AI-based decision support system provides explanations, the system should provide the users with customization features to tailor the provided information according to their individual needs in the XUI in order to improve the user's perceived informativeness.

The interactivity of explanations is a valuable aspect that can advance the research field of XAI (Adadi & Berrada, 2018). Moreover, XUIs are essential to enable and foster a practical interaction experience between users and XAI systems (Bradley et al., 2022). Through interactive visualizations, users can gain insights to identify classification errors, and by correcting them, the performance of the underlying AI can be improved (Alsallakh et al., 2014). Moreover, the interaction is significantly relevant for users to develop trust towards a decision aid (Al-Natour et al., 2022). Therefore, interactivity is an essential aspect of explanations that can be fostered through customization. Interaction experiences provided through customization features can lead to versatile responses from users (Bailey et al., 2009). Between the interaction of users with systems, customization can lead to reduced error rates or increased user acceptance (Burkolter et al., 2014). Moreover, customization features can afford more interactive exchanges between users and systems, resulting in a more positive stance towards customizable systems (Kalyanaram & Sundar, 2006). As a result, the following design requirement is established:

> **Design Requirement 3:** If an AI-based decision support system provides explanations, the system should provide the users with customization features as an engaging feature in the XUI in order to improve the user's perceived interactivity.

Satisfaction is an integral measurement for evaluating the quality of explanations (Arrieta et al., 2020). It can be used to investigate the explanation quality (Haque et al., 2023). In more detail, satisfaction can indicate whether users understood the system or how the system made a particular decision (Hoffman et al., 2018). The design of XUIs and decision aids, in a broader sense, is an integral part of improving the user's satisfaction (Al-Natour et al., 2022). When users interact with explanations that provide customization features like a "drill-down" function, they can individually satisfy their information needs, which is relevant for accepting XAI systems (Chromik & Butz, 2021). Providing users with customization features can improve performance and intrinsic motivation when performing tasks and increase overall satisfaction (Bailey et al., 2009). Studies have also highlighted that effectively designed customization features can lead to adequate user satisfaction (Jorritsma et al., 2015). As a result, the following design requirement is established:

> **Design Requirement 4:** If an AI-based decision support system provides explanations, the system should provide the users with customization features in a satisfactory manner to interact with the provided information in the XUI in order to increase the user's perceived satisfaction.

Table 2 provides an overview of the established design requirements, their overarching goals, and the associated evaluation goals. Ultimately, the evaluation goals represent the goodness criteria that serve as a way to assess the solution acceptance (vom Brocke et al., 2020), i.e., the influence of customization features in XUIs on the perceived explanation satisfaction. The constructs of perceived interestingness, perceived informativeness, perceived interactivity, and satisfaction are all relevant notions of explainability that can influence the perceived explanation quality (Minh et al., 2022; Vilone & Longo 2021). In summary, these are the overarching quality criteria that the XUI should meet, and they were used to evaluate the perceived explanation quality.

DR#	Overarching Goal of the Design Requirement	Evaluation Approach
DR1	Provide customization as an interesting feature for a secondary activity.	Measuring the perceived interestingness
DR2	Provide customization as possibility to adjust the presented information based on the user's needs.	Measuring the perceived informativeness
DR3	Provide customization as an engaging feature to provide a high degree of interactivity.	Measuring the perceived interactivity
DR4	Provide customization as lightweight and easy-to-use feature so that users can satisfy their needs.	Measuring the satisfaction

Table 2. Overview of the derived design requirements and the associated evaluation approaches.

4.2 Designing Customization Features of Explanation User Interfaces

The above-derived and described design requirements aim to support defining design principles that meet the associated evaluation goals. The status quo on formalizing design knowledge in the form of design principles was adhered to by formalizing them according to the scheme of Gregor et al. (2020). During the DSR project, two design principles for customization features were derived. The first design principle focuses on cosmetic customization, and the second on functional customization (Sundar et al., 2015). They enable users to tailor the explanations and presented content to and by themselves (Sundar, 2020). The design principles aim to guide how to operationalize customization features for the design explanations in XUIs.

The first design principle focuses on cosmetic customization. From a more general human-computer interaction perspective, cosmetic customization has been implemented by enabling users to change the appearance of UIs (Sundar et al., 2015). In the specific context of explanations integrated in XUIs, the concept of cosmetic customization is operationalized to adjust and change the presentation format of the explanation. Considering human factors and the user's information needs is highly relevant for designing XUIs and facilitating an appropriate human-XAI interaction (Langer et al., 2021; Schoonderwoerd et al., 2021). This is important since prior research has shown that well-designed explanations can influence human-computer interaction and ultimately increase their acceptance (Cramer et al., 2008). Visualizations are a well-established approach to present explanations, described as human-centered (Adadi & Berrada, 2018). Visual explanations, in combination with other techniques, are described as "[…] the most suitable way to introduce complex interactions within the variables involved in the model to users not acquainted to ML [machine learning] modeling." (Arrieta et al., 2020, p. 88). In addition, visualizations can improve explanations' understandability, explainability, and interpretability (Alicioglu & Sun, 2022). Consequently, it is possible that providing explanations can trigger different positive heuristics and thus lead to better user engagement (Sundar, 2020). Through this engagement, which is further facilitated through customization, users can manipulate visual objects in XUIs in various ways (Sundar et al., 2015). The adjustment of visual components is an established approach for customization and is described as one of the most common ones (Kleinsmith & Gillies, 2013; Teng, 2010). It is, therefore, well-suited to be transferred to the realm of XUIs. Therefore, it is anticipated that providing users with a set of visualization formats from which they can choose their individual needs and expectations that may differ (Barda et al., 2020; Langer et al., 2021) can be better satisfied in this way. Consequently, the following design principle is proposed:

> **Design Principle 1 – Cosmetic Customization:** To allow users to interact with AI-generated explanations for adjusting their visual representation from a set of pre-defined alternatives, in the context of decision-making supported by an XAI-based system, the system should provide easy to use settings for users to customize the visual representation according to their needs as well as preferences, so that the cosmetic customization in XUIs can provide a high degree of interactivity, interestingness, informativeness and satisfy users.

The second design principle focuses on functional customization, also established in human-computer interaction. It is described as modifying task-centered utility tools and has been implemented for users as features to enable interaction with the content, for example, by filtering or creating content (Sundar et al., 2015). It is anticipated that the customization of explanations could support them in identifying patterns used by the AI to produce specific outcomes (Ribeiro et al., 2016). When users can identify patterns, they are supported in identifying system failures and thus distinguish between correct but unexpected outputs from system malfunctions (Langer et al., 2021). Moreover, identifying such patterns could lead to the development of correct mental models of users, which can lead to outcomes including establishing appropriate trust, better comprehension, or better performance (Gunning & Aha, 2019). Individual users may have different backgrounds, prior experience with AI, and specific information needs that need to be considered when designing explanations (Meske et al., 2022; Sundar, 2020). Since not all these aspects are known during the development of XAI techniques or XAI systems, customization could be a valuable approach to consider human factors while giving the user a certain degree of freedom (Schoonderwoerd et al., 2021). Moreover, individual stakeholders may need more or less information presented in the explanation (Langer et al., 2021). By providing the customization of the amount of information provided in the explanation, the state of information overload could be avoided. Due to information overload, users may miss relevant information for the decision to make (Poursabzi-Sangdeh et al., 2021). Additionally, providing users with a customization feature can lead to an improved sense of control (Sundar et al., 2015). This perceived sense of control can influence the intrinsic motivation to explore the provided content and attitudes toward the system (Kang & Sundar, 2013). Users may discover novel knowledge by exploring the provided content (Vilone & Longo, 2021). Such an opportunity for discovery engages users' attention, and customization can be an approach to design this activity in an interesting way (Kang & Lee, 2015; Kleinsmith & Gillies, 2013; Vilone & Longo, 2021). In addition, prior research has shown that explanatory facilities can improve satisfaction with the decision process and transparency with the decision advice (Li & Gregor, 2011). This makes satisfaction vital when measuring the explanation quality and designing satisfying explanations (Haque et al., 2023). Consequently, the following design principle is proposed:

Design Principle 2 – Functional Customization: To allow users to interact with AI-generated explanations for adjusting the number of relevant features displayed in the XUI from a set of pre-defined alternatives, in the context of decision-making supported by an XAI-based system, the system should provide easy to use settings for users to customize the scope of relevant features displayed in the XUI, so that the functional customization in XUIs can provide a high degree of interactivity, interestingness, informativeness and satisfy users.

During the research, a third type of customization was conceptualized and instantiated as part of the second design cycle. This is the customization of the XAI method itself, which could be achieved, for example, by different XAI methods generating explanations and users being able to choose an adequate one, such as local and global explanations (Adadi & Berrada, 2018). Research has shown that different explanation types can lead to different effects in the same use case. For example, van der Waa et al. (2021) reported that rule-based explanations positively affected system understanding, whereas example-based explanations did not. In addition, both explanation types seem to persuade users to follow the AI recommendation even when incorrect. While designing an XUI in the healthcare context, Barda et al. (2020) concluded that varying stakeholders had different explanation goals and information needs, which were influenced by their knowledge. As previously described, users also have very different demands, information needs, and cognitive abilities to use the information presented in an XUI (Langer et al., 2021). Therefore, leaving the choice of the XAI method to the users can have positive consequences. Such interactions can also lead to users engaging with the explanation more intensively, thereby developing a correct mental model and, thus, in addition to a balanced level of trust, enabling users to recognize errors (Gunning & Aha, 2019; Gunning et al., 2019). By correcting errors, the user can also actively influence the quality of the underlying AI and help optimize it (Meske & Bunde, 2023). Furthermore, users can select explanations that provide the appropriate mental and cognitive fit (Vilone & Longo, 2021). Such a type of XAI method customization could continue to design XUIs to equally cover the needs of different stakeholders (Bunde et al., 2023).

Design Principle 3 – XAI Method Customization: To allow users to interact with AI-generated explanations for adjusting the XAI method displayed in the XUI from a set of predefined alternatives, in the context of decision-making supported by an XAI-based system, the system should provide easy to use settings for users to customize the presented explanation displayed in the XUI, so that the XAI method customization can provide an engaging and satisfying user experience.

After proposing design principles for customization features in XUIs, the following conceptualization step is the derivation of design features. The design features are specific capabilities of the artifact that satisfy the design principles and the instantiation of the prescriptive design knowledge (Meth et al., 2015; Seidel et al., 2018).

The first design feature focuses on presenting the explanations in a visual form, a human-centered approach to creating visually appealing explanations (Adadi & Berrada, 2018; Arrieta et al., 2020). A wide range of visualization techniques can be used to generate visual explanations that can support different objectives, including selecting features, analyzing the performance, or developing an understanding of the AI model (Alicioglu & Sun, 2022). In combination with the circumstance that humans have preferences and employ cognitive biases when evaluating and selecting explanations while having different information needs (Langer et al., 2021; Miller, 2019), it is anticipated that providing different visualizations as part of cosmetic customization can lead to an improved perceived explanation quality. This leads to the establishment of the first design feature:

Design Feature 1: Provide a pre-defined selection of visual explanations from which users can choose a visualization according to their preferences.

The motivations and reasons for interacting with explanations can vary significantly from use case to use case and among stakeholders (Adadi & Berrada, 2018; Langer et al., 2021). Therefore, individual users can have varying needs regarding the scope of information they desire or need, which is further influenced by the motivation to examine explanations (Arrieta et al., 2020). By allowing users to choose how much information they perceive, the size of the explanation can be adjusted based on the number of features incorporated in the explanation (Barda et al., 2020). Combined with visual explanations, this can aid users in identifying informative and non-informative data features (Alicioglu & Sun, 2022). Therefore, customization features to adjust the scope of information could avoid information overload (Poursabzi-Sangdeh et al., 2018). Consequently, it is anticipated that providing different information scopes through data feature sets as part of the functional customization can lead to an improved perceived explanation quality. This leads to the establishment of the second design feature:

Design Feature 2: Provide a pre-defined selection of data feature sets from which users can choose a scope of data features according to their need for information.

When users are provided with customization features, these features should be designed to be effortless, as an effortful customization procedure can have negative side effects like decision fatigue (Sundar et al., 2012). By providing accessible and intuitively usable customization features in UIs like "drill-downs", it is further possible to consider individual information needs, which in turn can positively influence the information processing of users (Barda et al., 2020). Moreover, through customization by a predetermined choice set, users can easily adapt the presented visualizations or the amount of information included in the explanation (Barda et al., 2020; Lopez-Giraldez & Townsend, 2011; Sundar et al., 2012). Therefore, it is anticipated that providing the cosmetic and functional customization feature easily and intuitively can lead to an improved perceived explanation quality. This leads to the establishment of the third design feature:

Design Feature 3: Provide the pre-defined selections of visual explanations and feature sets through an intuitively operable function in the XUI.

It is not to be assumed that users have previous knowledge of AI since users are very individual, depending on the use case (Langer et al., 2021). Therefore, one should not assume that users know different types of explanations or can distinguish or differentiate them from one another during the design. For this reason, the XAI methods should be represented in the XUI by names or labels that are as easy to interpret as possible. For example, existing terms such as local and global explanations can be used, whereby the users should be more familiar with these terms. Users can use local explanations to understand a specific output, such as classification, by presenting the most relevant features for the output. This can, for example, be the weighting of the words for a sentiment classification, which is comparable to XAI methods such as LIME (Ribeiro et al., 2016). On the other hand, users can use global explanations to understand the most relevant features for an entire AI model, which can be achieved, for example, with permutation feature importance. It is assumed that customization features in XUIs positively influence the interaction experience for users. This leads to the establishment of the fourth design feature:

Design Feature 4: Provide the pre-defined selections of explanation types through an intuitively operable function in the UI.

Figure 3 provides an overview of the established design requirements, proposed design principles, and derived design features.

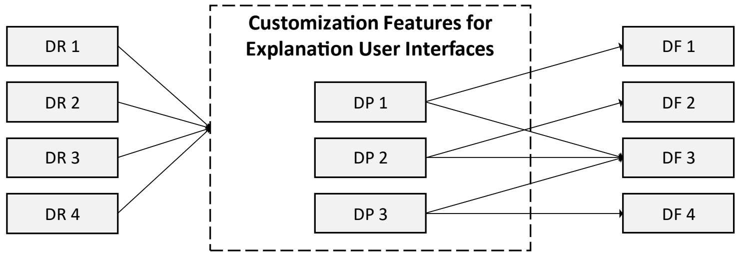


Figure 3. Overview of the relationship between design requirements, design principles, and design features.

4.3 Artifact Description: Customization Features in Explanation User Interfaces

During the DSR project, the initial goal was to develop an XUI with customization features that positively impact the perceived explanation quality. For this purpose, the design requirements were established, and the perceived explanation quality was operationalized through an adequate selection of dimensions. The initial design for an XUI with customization features can be seen in Figure 4. The XUIs were set up for the experiments and presented to the participants in such a way that they represent a new AI feature of human resource software. It is a web-based interactive XUI. The input data for the present case can be seen at the top left. All characteristics and their relevance for the classification are expanded in a tabular format by clicking the dark gray button. The top right is the recommendation or prediction of the AI. During the experiment, the subject had to decide whether to follow the recommendation based on the information presented or make a different decision. Below is the most exciting and essential part of this research project. This is the explanation, and there, the first and second design principles are instantiated using the associated design features. On the one hand, the chart type could be changed using cosmetic customization. On the other hand, functional customization made it possible to visualize different amounts of features in the chart.

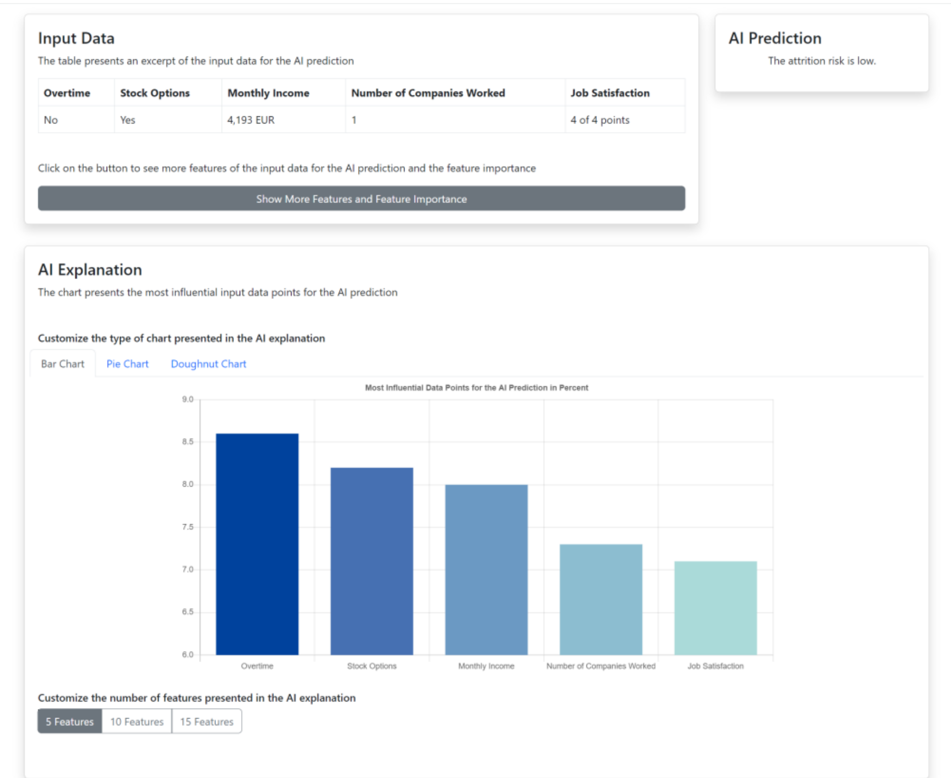


Figure 4. Exemplary XUI with customization features from design cycle one.

Figures 5 and 6 show two example XUIs from the second design cycle. It can be seen that the basic design remained untouched. The upper part of the design was retained and updated with the data relevant to the use case. In addition to the first and second design principles, the new third design principle has also been integrated into the lower part. Users have also been able to customize the chart type through cosmetic customization and change the number of features. Since the data set contained fewer features, this selection was adjusted here. The customization of the XAI method was new. Users could choose between local explanations (Figure 5) or global explanations (Figure 6). Information on how to interpret these statements and how they differ was provided to participants at the outset.

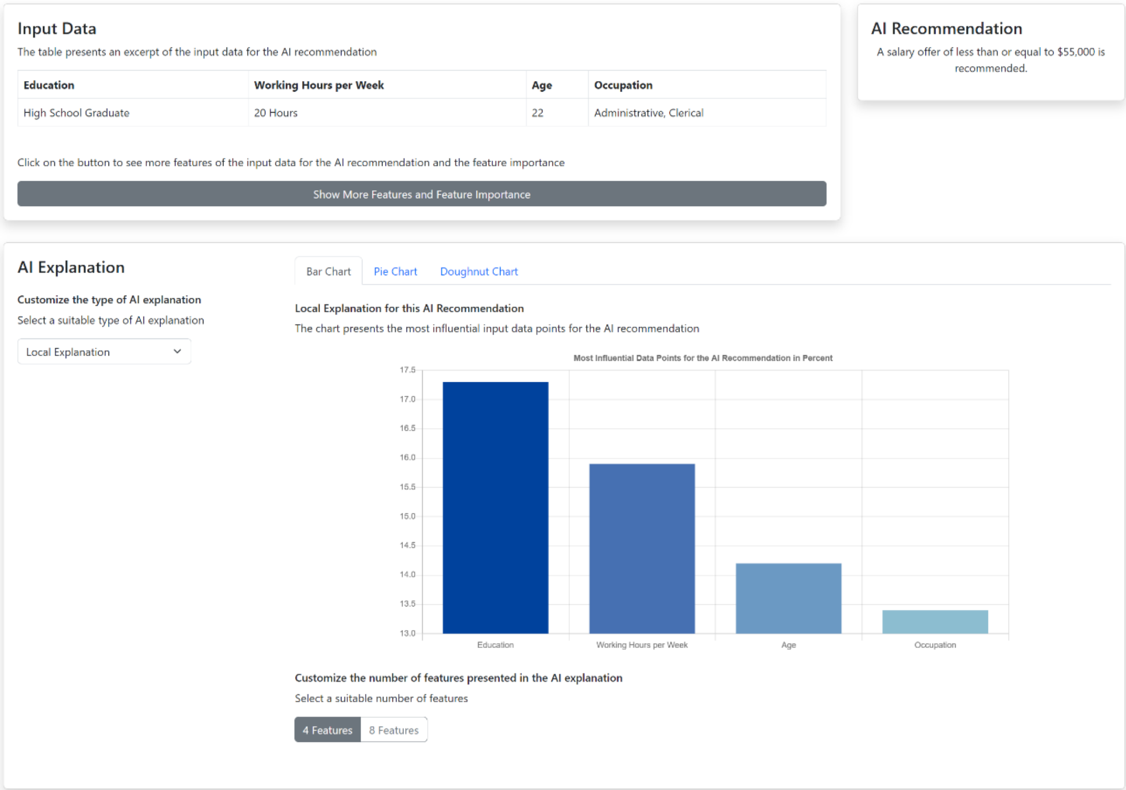


Figure 5. Exemplary XUI with customization features and a local explanation from design cycle two.

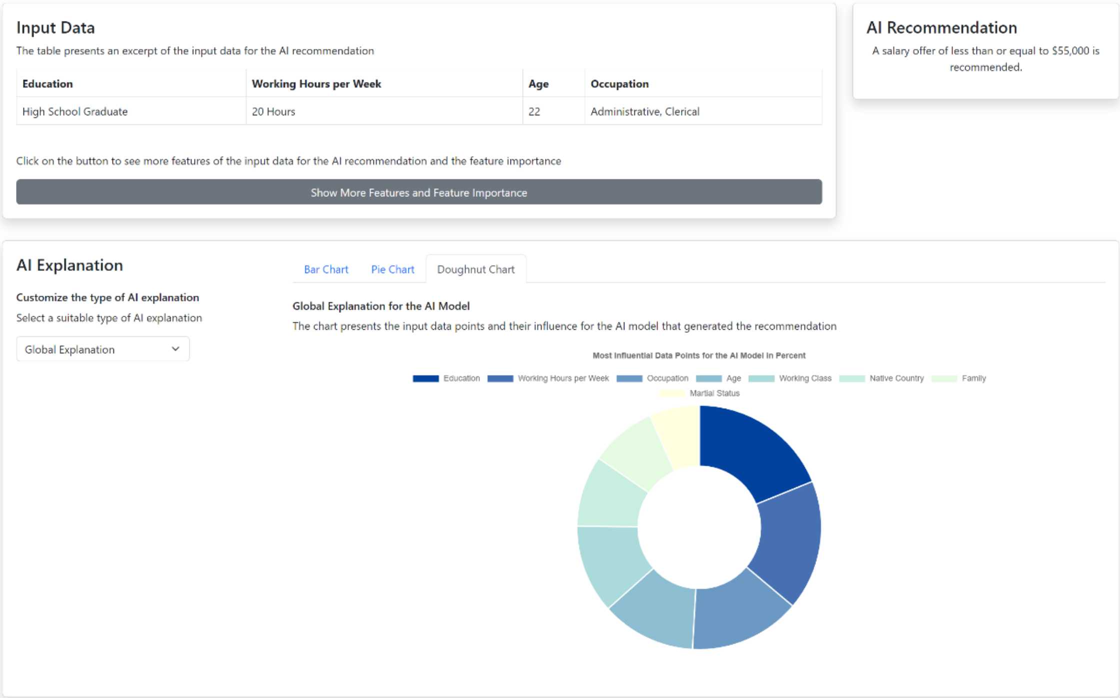


Figure 6. Exemplary XUI with customization features and a global explanation from design cycle two.

5. Evaluation

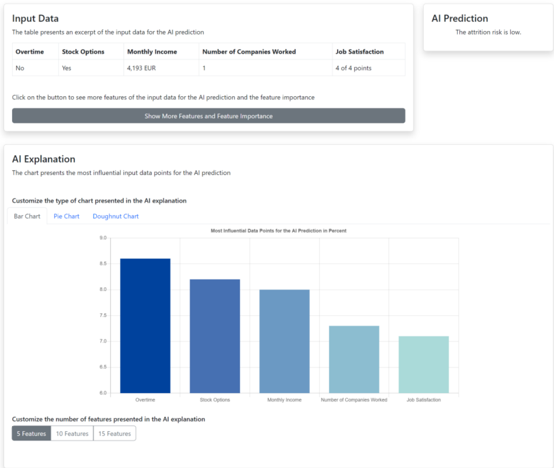
5.1 Participants and Procedures

The evaluations were planned and carried out using experimental studies in the form of online experiments. Online experiments are excellent for investigating human-computer interaction and human behavior during interaction with UIs (Fink, 2022). The participants for the online experiments were recruited on the Prolific platform. Subjects who can be recruited on this platform are demonstrably characterized by high attention and comprehension, which can positively affect the quality of the data collected (Peer et al., 2022). Motivated by the data sets, a human resources management use case was set up in both experimental studies. The key inclusion criterion for both experimental studies was that subjects should have experience in hiring processes, a criterion built into Prolific. In the first experimental study, a between-subject experiment design was chosen. It was, therefore, only possible for subjects to participate once in one of the two groups. Subjects who participated in the first experimental study were not admitted to the second experimental study. Other inclusion criteria applied to the reusability evaluation of the design principles. The criteria built into Prolific to know about user interface or user experience design. Appendix A presents the demographic data for the participants for the first experimental study, and appendix B presents the data for the second experimental study.

5.2 Experimental Study 1

5.2.1 Experiment Design, Flow and Task

In the binary classification task in the first online experiment, subjects had to classify the attrition risk as high or low. A between-subject experiment design was chosen to investigate the influence of the customization features in XUIs on the perceived quality of the explanation. A group interacted with XUIs, which had customization features. This version of the XUI can be seen in Figure 7a. A second group interacted with an identical XUI, which did not have any customization features. This XUI version can be seen in Figure 7b. Appendix C provides a selection of screenshots from the XUI from the first design cycle.



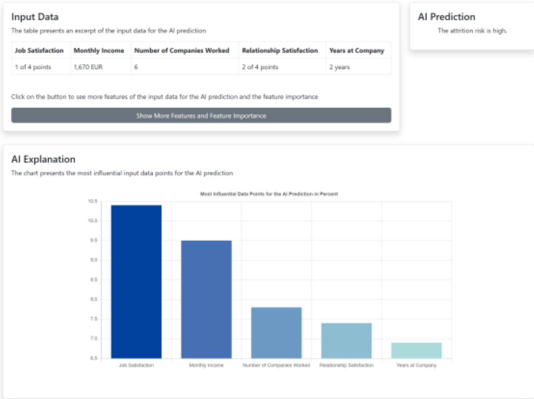


Figure 7a. XUI with customization features. **Figure 7b.** XUI without customization features.

In order to make the perceived quality of the explanation measurable, selected dimensions were identified, which can be understood as a link between the evaluation of XAI methods and customization features. These include perceived interestingness (Shin et al., 2022), perceived informativeness (Al-Natour et al., 2022), perceived interactivity (Sheng & Joginapelly, 2012), and explanation satisfaction (Hoffman et al., 2018). An overview of the constructs and items used is provided in Appendix D. All constructs were measured on a Likert scale from 1 to 7. In order to determine to what extent the designed customization features also influenced the perceived customization, the perceived customization (Theodosiou et al., 2019) was measured and compared between the groups as a manipulation check. In addition, the use of the customization features was measured within the group that had access to these features. The usage was measured independently as the number of mouse clicks on the cosmetic and functional customization features in order to be able to compare them with each other. Finally, the accuracy was measured during the binary classification task during the experiment. The subjects had to classify four examples, the order of which was randomized. In the examples, there was a correct classification for a low and high attrition risk and an incorrect classification for a low and high attrition risk. The experimental procedure began with subjects being recruited through Prolific. They were then directed to an external survey site. Here, they had to read through a short introduction to the study and a brief human resource management scenario description. In the next step, the subjects were provided with an explanation of the XUI and the different information, which was not yet part of the classification task. The classification task followed, then the measurement of the constructs, a demographic question, and a farewell.

5.2.2 Anticipated Influence of Customization Features in Explanation User Interfaces

Overall, the ex-post evaluation aimed to investigate the influence of the two design principles, focusing on customization on perceived interestingness, perceived informativeness, perceived interactivity, and satisfaction. In the following, hypotheses about the influence of customization features in XUIs will be presented and tested in an online experiment to evaluate the DSR process (Kuechler & Vaishnavi, 2012; Peffers et al., 2007).

When AI models become explainable, they are described to become more interesting for users (Minh et al., 2022). Interestingness is a notion of explainability related to the usefulness of the data and ultimately supports users in discovering interesting knowledge (Freitas, 2006). Moreover, it is described as the ability of XAI methods to enable users to discover novel knowledge and engage users' attention

(Arrieta et al., 2020). The design of explanations can also influence the perceived interestingness of the system itself (Cramer et al., 2008). Customization, in turn, gives users a sense of control over the UI, impacting user efficiency and ease of use (Hui & See, 2015). Prior research has further uncovered that when users work with XUIs, they are interested in working with typical customization features, including "drill-down" or "zoom-in" (Barda et al., 2020; Sundar et al., 2015). Therefore, it is anticipated that users perceive them as more interesting through well-designed customization features in XUIs.

> **H1:** The design feature for customization will positively influence the interestingness and of explanations in the XUI.

An overarching objective that XAI pursues is adequate informativeness of explanations (Arrieta et al., 2020). It can be described as the ability of XAI methods to provide users with useful information (Lipton, 2018). Therefore, explanations should support users in gathering information and gaining knowledge (Adadi & Berrada, 2018). Informativeness is a notion of explainability that can be measured to compare different XAI methods, while high values for informativeness are desirable (Vilone & Longo, 2022). By providing explanations in UIs, it is also possible to enhance the perceived informativeness (Meske & Bunde, 2023). Additionally, customization features in XUIs can support the satisfaction of information needs of individual stakeholder groups (Barda et al., 2020). This can be achieved since the customization features of XUIs determine the amount and accuracy of information presented to address the user's needs (Burkolter et al., 2014). By letting users customize the information to their needs, they can access the information that is more likely to be useful (Kang & Sundar, 2013). This results in the anticipation that an explanation with customization features will positively influence users' perceptions. Consequently, the following hypothesis is established:

> **H2:** The design feature for customization will positively influence the informativeness of explanations in the XUI.

The concept of interactivity is essential for designing XAI systems since it supports system understandability (Mohseni et al., 2021). By providing XUIs with interactive features, users are enabled to explore the explanations by themselves (Adadi & Berrada, 2018). Additionally, interactive features can support users in correcting errors, leading to improved performance (Alsallakh et al., 2014). Additionally, when users can integrate their knowledge during the interaction with AI models, it can further improve the knowledge discovery process pipeline (Holzinger, 2016). Customization is anticipated to influence interactivity since users have versatile information needs (Meske et al., 2022). The varying information needs can be satisfied by allowing users to independently manipulate the data presented in the XUI, which can be achieved through customization (Bolin et al., 2005). Prior research has shown that when users can customize and control the information they receive, they perceive it as more interactive (Kalyanaraman & Sundar, 2006). Additionally, customization has led to higher perceived interactivity in empirical studies (Sundar et al., 2012). Therefore, customization features in XUIs are anticipated to influence the perceived interactivity. Consequently, the following hypothesis is established:

> **H3:** The design feature for customization will positively influence the interactivity of explanations in the XUI.

Satisfaction is a well-established aspect with high importance when explanations in intelligent systems are to be evaluated (Mohseni et al., 2021). Explanation satisfaction can be described as the "[…] subjective rating of explanation completeness, usefulness, accuracy, and satisfaction." (Gunning & Aha, 2019; p. 54) of users. Studies have demonstrated that explanations can significantly influence the satisfaction of users with AI systems when compared to AI systems without explanation (Wells & Bednarz, 2021). Similarly, customization can have a direct and indirect effect on satisfaction. This effect was proven in quantitative studies (Teng, 2010) and qualitative studies (Fukazawa et al., 2009). Users who can control the design of UIs can lead to higher perceived efficiency, which in turn results in higher overall satisfaction (Hui & Lee, 2015). Prior research has also highlighted that customization can directly affect satisfaction (Chung & Shin, 2008). Therefore, customization features in XUIs are anticipated to influence explanation satisfaction. Consequently, the following hypothesis is established:

> **H4:** The design feature for customization will positively influence the satisfaction of users with the explanation in the XUI.

Figure 8 summarizes the research model for the first online experiment as part of the evaluation. The four constructs, perceived interestingness, perceived informativeness, perceived interactivity, and the explanation satisfaction scale, were operationalized to explore the influence of customization features in XUIs on the perceived explanation quality.

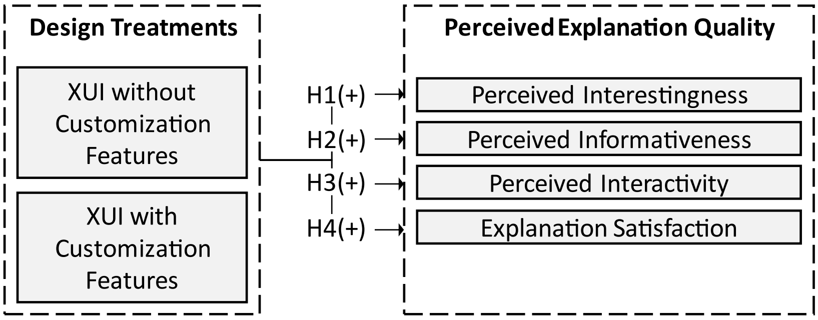


Figure 8. The proposed research model for the comparison of the perceived explanation quality.

5.2.3 Statistical Analysis

Overall, 180 participants were recruited for the experiment as part of the ex-post evaluation. The 180 participants were randomly assigned into one of the groups. Either they were assigned to the group that interacted with the XUI baseline design or to the group that interacted with the XUI customization design. Three participants in the baseline group failed the IMCs, and four participants in the customization group. Their answers were not included in the statistical analysis. Additional participants were recruited to achieve the target sample size of 90 participants per group. The participants who were subsequently recruited were not allowed to have taken part in the experiment before, and all passed the IMCs. Since the experimental task focuses on classifying the attrition risk, the inclusion criteria of having experience in hiring processes were established. The hiring experience distributed over the two experiment groups was analyzed using Welch's t-test. The hiring experience for the customization group ($M = 5.07$ years, $SD = 3.8$ years) and baseline group ($M = 5.31$ years, $SD = 4.16$ years) did not significantly differ showed the results, $t(180) = -0.41191$, $df = 176$, $p = 0.6595$, $d = 0.06$. This test was necessary since the experience level in hiring processes could influence the task performance. As a manipulation check, the perceived customization was measured. The customization group ($M = 5.1$, $SD = 1.2$, $\alpha = 0.84$) and the baseline group ($M = 4.4$, $SD = 1.2$, $\alpha = 0.82$) showed significant differences, which proves the effectiveness of the instantiated design features, $t(180) = 3.8048$, $df = 177.98$, $p < .001$, $d = 0.57$.

After the participants finished the experimental task, they were provided with a post-survey. This survey contained the notions of explainability that were operationalized. Table 3 summarizes the descriptive measurements for the constructs and provides the means, standard deviation, and Cronbach's α. The analysis resulted in good to excellent values for Cronbach's α. The measurements for the notions of explainability were rated higher in the group that interacted with XUI with customization features compared to the XUI baseline design.

Design Treatment	Construct	Mean	Standard Deviation	Cronbach's α
Baseline	Interestingness	4.6	± 1.3	0.88
Customization		5.3	± 1	0.79
Baseline	Informativeness	4.6	± 1.2	0.88
Customization		5.7	± 1	0.89
Baseline	Interactivity	4.6	± 1	0.86
Customization		5.7	± 0.9	0.89
Baseline	Satisfaction	4.5	± 1.2	0.92
Customization		5.2	± 1.1	0.93

Table 3. Descriptive analysis of the gathered data.

For a comparative evaluation of the measured notions of explainability, Welch's one-sided t-test was used. The following Table 4 provides an overview of relevant statistical values. The 90 participants using the customization design, compared to the 90 participants using the baseline design, demonstrated significantly higher perceived interestingness, $t(180) = 4.1672$, $p < .001$ with a medium effect size ($d = 0.62$); informativeness, $t(180) = 4.593$, $p < .001$ with a medium effect size ($d = 0.69$); interactivity, $t(180) = 7.2961$, $p < .001$ with a large effect size ($d = 1.08$); and satisfaction $t(180) = 3.9626$, $p < .001$ with a medium effect size ($d = 0.57$). Consequently, the XUI design with customization features was significantly better rated than the XUI baseline design.

Design Hypothesis	Construct	t	df	p Value	Cohens' d
Customization Design → Baseline Design	Interestingness	4.1672	167.27	< .001	0.62
	Informativeness	4.5953	166.95	< .001	0.69
	Interactivity	7.2961	176.39	< .001	1.08
	Satisfaction	3.9626	177.98	< .001	0.57

Table 4. Results of the Welch's one-sided t-test.

Through the online experiment as part of the ex-post evaluation, it was possible to generate evidence for the anticipated effects of customization features in XUIs. Table 5 provides an overview of the above-introduced hypotheses, the hypothesized effects, and their result. Based on the results of the statistical analysis, all hypotheses are supported. Consequently, customization features in XUIs can positively influence the perception of explanations and, ultimately, highly relevant notions of explainability.

Hypothesis	Hypothesized Effect	Support
H1	Customization will lead to a higher degree of perceived interestingness.	Yes
H2	Customization will lead to a higher degree of perceived informativeness.	Yes
H3	Customization will lead to a higher degree of perceived interactivity.	Yes
H4	Customization will lead to a higher degree of satisfaction.	Yes

Table 5. Summary of hypotheses and results.

Besides the above-introduced scales for the notions of explainability, three further dimensions have been measured in the online experiment. A statistical analysis of these measurements generated exciting insights into the interaction between humans with XUIs and the potential influence of customization features on task performance. The first dimension was the interaction behavior with the design feature to retrieve the fifteen most relevant features in a tabular overview for the presented case. This design feature was instantiated in both artifacts. There was no significant difference between the customization group ($M = 0.91$, $SD = 0.59$) and the baseline group ($M = 0.93$, $SD = 0.56$) regarding the usage of this specific design feature, $t(180) = 0$, $df = 177.57$, $p = .5$, $d = 0$. In addition, the interaction with the cosmetic and functional customization feature in the group that interacted with the XUI with customization features was analyzed. Here, for the customization group, the usage of the cosmetic feature ($M = 3.06$, $SD = 2.76$) and the usage of the functional feature ($M = 1.55$, $SD = 1.4$) was also examined. The analysis uncovered that the cosmetic feature was significantly more frequently used, t(180) = 5.9838, df = 175.84, $p < .001$, d = 0.89. Another fascinating aspect that was uncovered during the statistical analysis was the difference regarding the achieved accuracy. The customization group (M = 80%, SD = 17.26%) achieved a significantly higher accuracy in the classification task compared to the baseline group (M = 66%, SD = 19.99%) showed the analysis, t(180) = 5.0894, df = 174.28, $p < .001$, d = 0.76.

5.3 Experimental Study 2

5.3.1 Experiment Design, Flow and Task

In the binary classification task in the second online experiment, subjects had to decide whether they offered an anonymous applicant a salary equal to or lesser than \$55,000 or more than \$55,000. An SEM was developed and statistically analyzed in order to investigate the influence of the extended customization features in XUIs. The participants interacted with the identical XUI. Figure 9 presents an exemplary XUI with a local explanation and four features as a doughnut chart. Figure 10 shows an XUI with a global explanation and eight features as a pie chart. Appendix E presents screenshots from the XUI for an exemplary case from the online experiment.

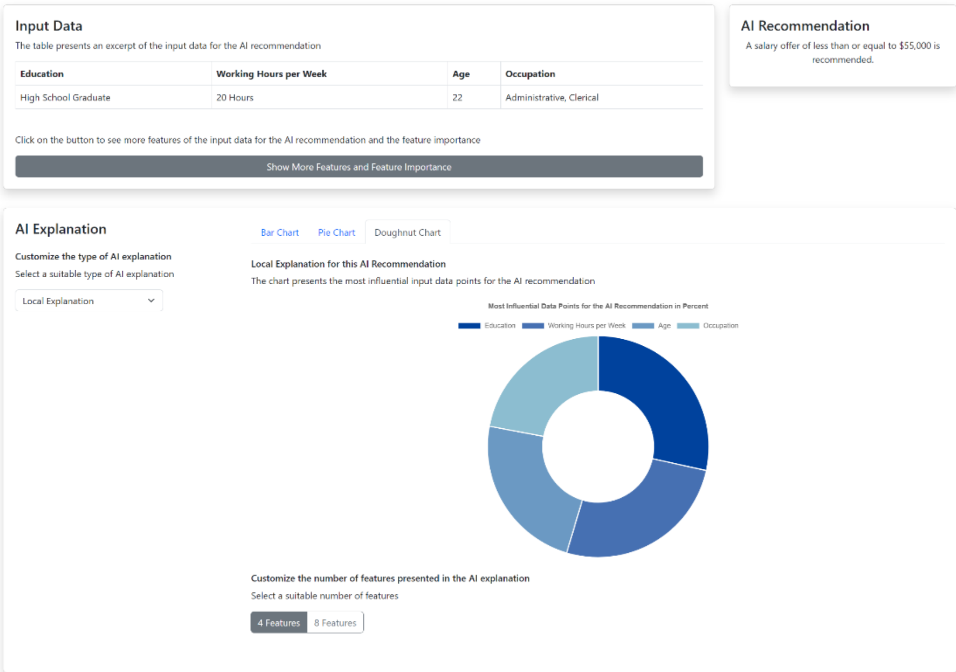


Figure 9. XUI with local explanation and four features as doughnut chart.

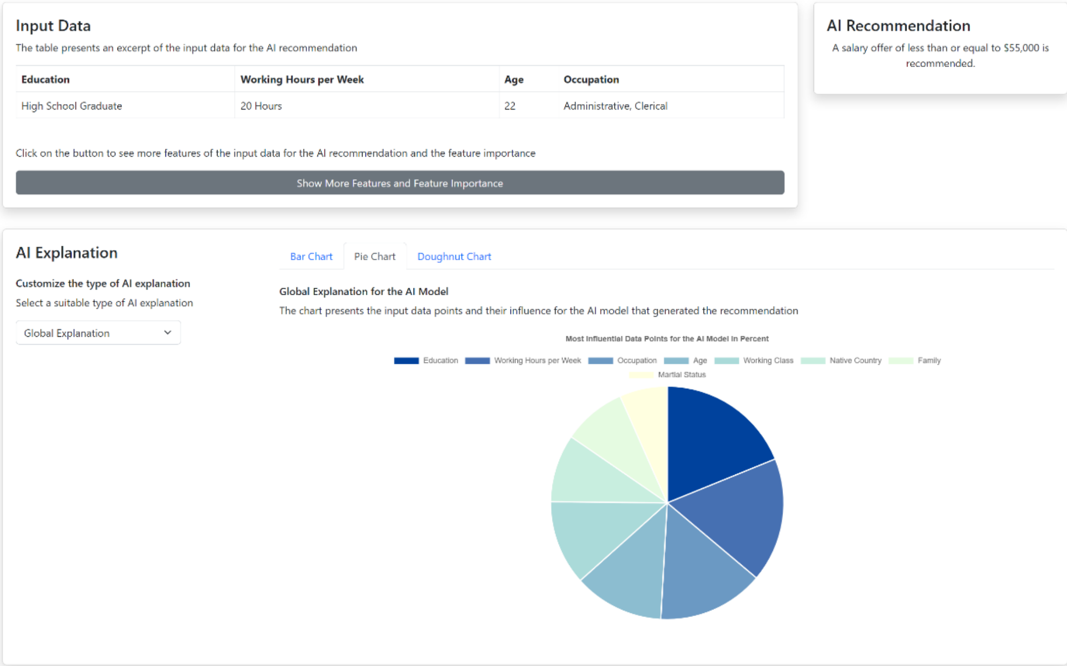


Figure 10. XUI with global explanation and eight features as pie chart.

The constructs measures include perceived interestingness (Shin et al., 2022), perceived customization (Theodosiou et al., 2019), perceived interactivity (Sheng & Joginapelly, 2012), user engagement (Wang & Sundar, 2017), and satisfaction (Li et al., 2021). An overview of the constructs and items used is provided in Appendix F. All constructs were measured on a Likert scale from 1 to 7. The usage of the three customization features was measured independently through the mouse clicks on the cosmetic, functional, or XAI method customization feature. The accuracy was also measured during the binary classification task part of the online experiment. The subjects had to classify six examples, the order of which was randomized. In the examples, there were two correct classifications for a salary equal to or less than $55,000 and more than $55,000. In addition, there was also one incorrect example for both cases. The experimental procedure was identical to the first experimental study.

5.3.2 Anticipated Influence of Customization Features in Explanation User Interfaces on User Engagement and Satisfaction

In the second design cycle, the focus was on examining two relevant dimensions for XAI. For this purpose, hypotheses were set up, and an SEM was developed to determine the influence of perceived interactivity, perceived interestingness, and perceived customization on user engagement and how strongly user engagement influences satisfaction. A moderation analysis was also carried out. The two dimensions of user engagement and satisfaction are essential for both XAI and UI research.

One critical aspect of well-designed explanations and UIs is user engagement, which plays a pivotal role in shaping the effectiveness and usability of XUIs to advance the XAI field (Adadi & Berrada, 2018). User engagement describes the extent to which users actively interact with a system and enjoy their interaction experience (O’Brien & Toms, 2008). The impact of user engagement can be versatile. For example, when users engage in the interaction experience with UIs, they explore the information more thoroughly (Sundar et al., 2015). In addition, interactivity allows users to understand the AI output, develops trust, and enables effective human-AI collaboration (Gunning & Aha, 2019). Research has already shown that explainable XUI design features with high user engagement foster trust, essential for their acceptance (Noori & Albahri, 2023). Therefore, perceived interactivity is an essential aspect that allows for meaningful interaction experiences and ideally leads to greater user engagement (Oh & Sundar, 2015; Sundar et al., 2015). Consequently, the following hypothesis is established:

H1: The perceived interactivity positively predicts the user engagement.

The perceived interestingness can be used to evaluate XAI (Vilone & Longo, 2021), and it is an essential factor when influencing user engagement or behavior (Arapakis et al., 2014). Interesting XUIs could, for example, stimulate the curiosity of users or lead to user retention (Constantin et al., 2019; Niu & Al-Doulat, 2021). When users are interested in the content and information they browse, they can be encouraged to dive deeper, leading to more focused attention (McCay-Peet et al., 2015). Consequently, they explore more XUI features when interested, which can influence the user's interaction behavior (Darejeh & Salim, 2016). This interaction experience can lead to users developing an understanding of the AI or improve engagement (Vilone & Longo, 2021; 2022; van der Waa et al., 2021). Since the perceived interestingness can be described as an essential driver for user engagement (Karnowski et al., 2017), it is anticipated that it positively predicts the user engagement of XUIs with customization features. Consequently, the following hypothesis is established:

H2: The perceived interestingness positively predicts the user engagement.

The perceived customization can encompass a sense of control for users over the interaction with UIs, leading to greater engagement (Oh & Sundar, 2015; Sundar et al., 2015). The customization features in the XUIs allow users to adjust the information in ways that make sense to them, which can support users in assessing the model's credibility and utility (Barda et al., 2020). Moreover, they could be one mechanism to design explanations that fit users' needs and expectations (Naiseh et al., 2023). Customization features for decision-making are also described to make it more efficient (Kang & Lou, 2022). The interactivity that emerges through the customization features influences user engagement and further aspects, such as intrinsic motivation or attitudes toward a UI (Sundar et al., 2012). Ultimately, customization can enhance the intrinsic motivation to engage with technology (Sundar et al., 2015). User engagement is also influenced by customization's behavioral effects, which can result from an altered user activity (Kalyanaraman & Sundar, 2006). Due to the influence of customization features on the interaction experience, activities, and, ultimately, user engagement (Sundar et al., 2022), perceived customization is anticipated to predict user engagement with XUIs with customization features positively. Consequently, the following hypothesis is established:

H3: The perceived customization positively predicts the user engagement.

Satisfaction is a crucial indicator of user interaction experience with a decision aid (Al-Natour et al., 2022). It is a well-established aspect of research on XAI and the evaluation of XAI (Alufaisan et al., 2021; Vilone & Longo, 2021). When explanation features are perceived as unsatisfactory, users will not increase their trust in the underlying AI (Panigutti et al., 2023). Therefore, satisfaction is an essential aspect of evaluating the quality of UIs. It also indicates user perception since satisfied users are more likely to have a positive view (Sundar et al., 2015). Prior research has shown that user engagement is a good predictor of user satisfaction (Masrek et al., 2018). Therefore, it is anticipated to predict satisfaction with XUIs with customization features positively. Consequently, the following hypothesis is established:

H4: The user engagement positively predicts the satisfaction.

Figure 11 summarizes the hypotheses in the research model for the second experimental study and illustrates the structural model.

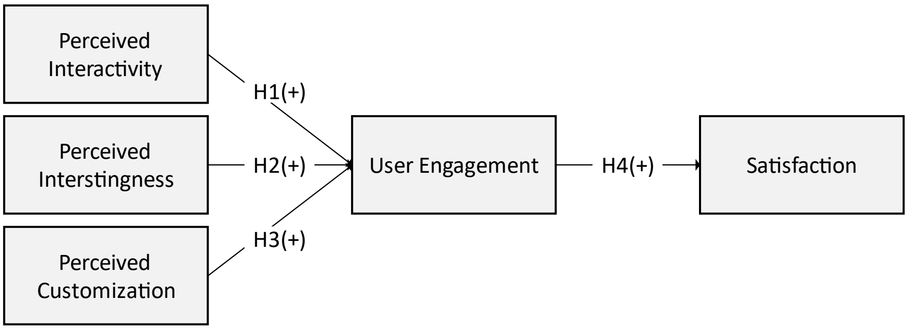


Figure 11. Research model for the second experimental study.

5.3.3 Statistical Analysis

Overall, 224 participants were recruited for the experiment as part of the second ex-post evaluation. The 224 participants were recruited on Prolific with the same inclusion criteria. Nine participants failed the IMCs, so their answers were excluded from the statistical analysis. Consequently, the answers of 215 participants built the foundation for the statistical analysis. Moreover, only participants were allowed to participate in the second study that had not participated in the first experimental study. The process of the online experiment was planned and executed identically to the experiment in the first design cycle. The only difference was that only one group was necessary in the second online experiment. To participate in the experiment, participants had to have experience in hiring processes (M = 2.1 years, SD = 1.7 years).

During the experiment, the participants in the online experiment had to classify six examples, with the answers being saved, from which the task performance, for example, in the form of accuracy, is calculated. Furthermore, while users interacted with the XUI, the use of the three customization features was tracked, and the mouse clicks on the cosmetic, functional, or XAI method customization feature were saved. After the classification task, the participants had to rate the different constructs in a post-survey on a Likert scale from 1 to 7. Table 6 shows the measured constructs, their means, and standard deviation. First of all, it can be seen that all constructs were rated very positively, and initially, only perceived interestingness (M = 4.93, SD = 1.44) differed from the ratings of the other measured constructs. Therefore, the statistical analyses and their results are described in the following steps, which explore the constructs and their influences. The participants achieved an accuracy during the classification task of (M = 62%, SD = 14.7%).

Construct	**Mean**	**SD**
Perceived Interestingness	4.93	± 1.44
Perceived Customization	5.53	± 1.34
Perceived Interactivity	5.32	± 1.43
User Engagement	5.45	± 1.29
Satisfaction	5.62	± 1.27

Table 6. Measurements for mean and standard deviation.

For the statistical analysis of the research model, SEM was adopted. Similar to research using the same approach, the two-step procedure introduced by Anderson and Gerbing (1988) was followed. The first step focused on investigating the measurement model, which is done to ensure the validity of the latent constructs. The structural relationship between the latent constructs was analyzed in the following step. Therefore, a confirmatory factor analysis (CFA) was conducted. Following the recommendations of Hu and Bentler (1999) and Schreiber (2008), the fit indices reached the threshold (CFI = 0.997, NFI = 0.998, RMSEA = 0.013, SRMR = 0.033). These results indicate that the quality of the measurement model was acceptable. Following the suggestions of Fornell and Larcker (1981), convergence and discriminant validity were evaluated before analyzing the relationship between the latent constructs in the structural model. Table 7 shows the analysis results and the standardized factor loadings of all items, which all exceeded a value of 0.7. The composite reliability (CR) for every construct is greater than 0.6, and the measured average variance extracted (AVE) is higher than 0.5. Consequently, the statistical analysis showed an adequate level of convergence validity. In addition, the skewness and kurtosis are presented. Most of the measurements are excellent, with values between -1 and +2. Other values are slightly over this range but still between -2 and +2. Therefore, the values are still acceptable (Hair et al., 2022).

Construct	Item	Standardized Factor Loading	Composite Reliability (CR)	Average Variance Extracted (AVE)	Skewness	Kurtosis
Perceived Interestingness	PINTE1	0.907	0.9190553	0.79	-0.4196989	-1.019439
	PINTE1	0.868				
	PINTE3	0.893				
Perceived Customizability	PCUS1	0.889	0.9159931	0.785	-1.149127	0.3303111
	PCUS2	0.877				
	PCUS3	0.891				
Perceived Interactivity	PINTA1	0.889	0.9521544	0.769	-0.9914019	-0.174894
	PINTA2	0.853				
	PINTA3	0.876				
	PINTA4	0.878				
	PINTA5	0.886				
	PINTA6	0.875				
User Engagement	UE1	0.874	0.970684	0.769	-1.216557	0.4512799
	UE2	0.872				
	UE3	0.875				
	UE4	0.858				
	UE5	0.881				
	UE6	0.891				
	UE7	0.973				
	UE8	0.872				
	UE9	0.883				
	UE10	0.885				
Satisfaction	SAT1	0.886	0.903502	0.759	-1.208995	0.5546395
	SAT2	0.841				
	SAT3	0.885				

Table 7. Results for the verification of the measurement model and convergence validity.

The discriminant validity was computed through the heterotrait-monotrait ratio of correlations with R and based on Henseler et al. (2015). The results are presented in Table 8. In the last step, the potential for outliers was investigated, and the results indicated that there were no outliers ($|z| < 3.0$) (Field, 200).

	Perceived Interestingness	Perceived Customization	Perceived Interactivity	User Engagement	Satisfaction
Perceived Interestingness	1.000				
Perceived Customization	0.118	1.000			

Perceived Interactivity	0.105	0.091	1.000		
User Engagement	0.024	0.060	0.026	1.000	
Satisfaction	0.086	0.090	0.034	0.037	1.000

Table 8. Heterotrait-monotrait ratio of correlations and discriminant validity.

In the next step, the structural model was analyzed. The model fit for the structural model was acceptable (CFI = 0.998, NFI = 0.998, RMSEA = 0.023, SRMR = 0.033). Table 9 and Figure 12 summarize the main results of the statistical analysis and hypotheses evaluation. When controlling the individual path coefficients, it is evident that the perceived interactivity significantly predicted user engagement ($\beta = 0.2847$, $p < .01$). Therefore, H1 is supported. Despite the positive measurement for perceived interestingness, there was no significant effect regarding the prediction of user engagement by perceived interestingness ($\beta = 0.0109$, $p = 825$), and H2 is rejected. Furthermore, the results show that H3 is supported since the perceived customization has significantly predicted user engagement ($\beta = 0.5005$, $p < .001$). Lastly, the results for the last hypothesis support H4 as user engagement significantly predicts satisfaction ($\beta = 0.9988$, $p < .001$).

Hypothesis	Paths	Standardized Regression Coefficient	Regression Coefficient	Standardized Error	Hypothesis Testing
H1	Perceived Interactivity → User Engagement	0.2847**	0.2778	0.0884	Supported
H2	Perceived Interestingness → User Engagement	0.0109	0.0108	0.0488	Rejected
H3	Perceived Customization → User Engagement	0.5005***	0.5027	0.0918	Supported
H4	User Engagement → Satisfaction	0.9988***	0.9823	0.0884	Supported

Table 9. Evaluation of the hypotheses (Note: ***p < .001, **p < .01).

Consequently, three of the four hypotheses are supported. In addition, the developed and tested model explains 53% of the variance in user engagement ($R^2 = 0.539$) and 99% of user satisfaction ($R^2 = 0.998$).

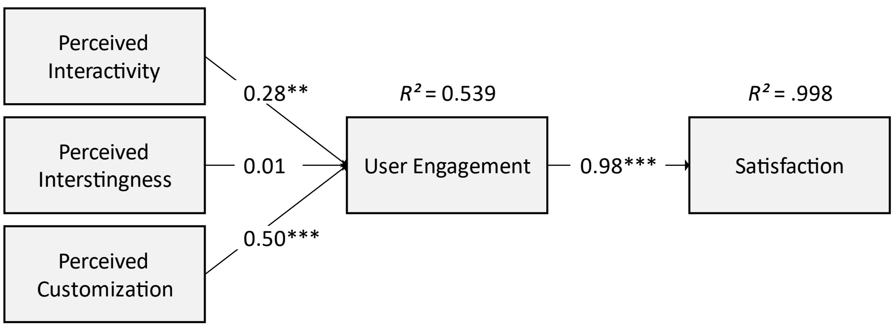


Figure 12. Results for the structural model (Note: ***p < .001, **p < .01).

R Jamovi was used for the mediation analysis, and 5000 bootstrap samples were executed, generating mediating effects and bias-corrected confidence intervals (95%). Table 10 provides an overview of the results of the mediation analysis. The results show a direct standardized effect (β = 0.5877, CI: [0.49777, 0.678]) of perceived interactivity on user engagement. In addition, the results show a significant indirect standardized effect for perceived interactivity on satisfaction mediated by user engagement (β = 0.0673, CI: [0.00598, 0.129]), for the perceived interestingness did not result a significant standardized direct effect on user engagement (β = 0.0479, CI: [-0.0667, 0.1624]) or indirect standardized effect on satisfaction, mediated by user engagement. (β = 0.0180, CI: [-0.0268, 0.0629]). Lastly, the perceived customization had a significant standardized effect on user engagement directly (β = 0.6278, CI: [0.53767, 0.718]) and also a significant standardized indirect effect on satisfaction (β = 0.0721, CI: [0.00703, 0.137]), moderated by user engagement.

Model Path	**Standardized Path Effect (β)**	**Standardized Error**	**95% CI**		**Effect Size of Mediation Effect**
			Lower Bounds	**Upper Bounds**	
Perceived Interactivity → User Engagement	0.5877***	0.0459	0.49777	0.678	
Perceived Interactivity → User Engagement → Satisfaction	0.0673*	0.0313	0.00598	0.129	0.114
Perceived Interestingness → User Engagement	0.0479	0.0584	-0.0667	0.1624	
Perceived Interestingness → User Engagement → Satisfaction	0.0180	0.0229	-0.0268	0.0629	
Perceived Customization → User Engagement	0.6278***	0.0460	0.53767	0.718	
Perceived Customization → User Engagement → Satisfaction	0.0721*	0.0332	0.00703	0.137	0.115

Table 10. Test of mediation of indirect, direct and bias-corrected 95% confidence interval (95%) (Note: ***p < .001, **p < .01, *p < .05).

As in the first online experiment, the use of the customization features in the XUI was also analyzed in the second experiment. The mouse clicks on the cosmetic, functional, and XAI method customization features were saved independently. The following values were obtained concerning the usage of the three customization features: cosmetic customization (M = 9.8, SD = 5.0), functional customization (M = 8.9, SD = 4.5), and XAI method customization (M = 11.1, SD = 5.4). The values for the three customization features were compared using an ANOVA. Table 11 presents the results of the ANOVA and shows that the XAI method customization features significantly more often (p < .05) than the functional customization feature and also significantly more often than the cosmetic customization feature (p < .001). Table 11 summarizes the results.

Customization Type	**Estimate Std.**	**Error**	***t* Value**	***p* Value**	**Effect Size (η^2)**
Cosmetic → Functional	-0.9349	0.4806	-1.945	.1270	0.18
XAI Method → Functional	1.2791	0.4806	2.661	< .05	
XAI Method → Cosmetic	2.2140	0.4806	4.606	< .001	

Table 11. Results of the ANOVA for the comparison of the customization feature usage.

5.4 The Reusability of the Proposed Design Principles

To ensure that the developed design principles are comprehensible and useful for the targeted stakeholders, they were evaluated regarding reusability (Ivari et al., 2018). For this purpose, the framework introduced by Ivari et al. (2021) was used, including accessibility, importance, novelty and insightfulness, actability and guidance, and effectiveness. These dimensions were rated using a Likert scale from 1 to 7. Since the design principles represent design knowledge for XUIs, i.e., a class of UIs, the inclusion criteria were set on Prolific that the participants had to have experience in user interface design or user experience design. A total of 65 participants were recruited. Of these 65 participants, the responses of 4 participants were ignored in the statistical analysis because they completed the evaluation in under 3 minutes, which represented an unrealistic survey completion time. Appendix G shows the demographics of the participants.

Consequently, 61 responses were included in the statistical analysis of the collected data. The recommendations of Ivari et al. (2018; 2021) were followed, and the participants first received a brief introduction to the topic of design principles and XUIs so that they could appropriately assess and evaluate the design principles and their role. The dimensions for the reusability evaluation were measured for each of the three design principles so that they could be evaluated independently. Appendix H provides an overview of the constructs and items for the evaluation. Table 12 provides an overview of the assessment of the reusability dimensions. The dimensions, the means, standard deviations, and Cronbach's alpha are shown. Overall, the design principles were rated positively, which is also reflected in the answers to two descriptive questions. Overall, 52 participants stated that they would use the evaluated design principles for a suitable project, and 53 participants stated that they would recommend the design principles for a suitable project to colleagues.

Construct	M	SD	Cronbach's Alpha
Design Principle 1 – Cosmetic Customization			
Accessibility	5.3	± 1.1	0.84
Importance	5.4	± 0.9	0.65
Novelty and Insightfulness	5.3	± 1.1	0.79
Actability and Appropriate Guidance	5.4	± 1.1	0.92
Effectiveness	5.4	± 1.1	0.91
Design Principle 2 – Functional Customization			
Accessibility	5.6	± 0.9	0.80
Importance	5.4	± 1.1	0.76
Novelty and Insightfulness	5.5	± 1.0	0.79
Actability and Appropriate Guidance	5.5	± 0.9	0.91
Effectiveness	5.5	± 1.0	0.64
Design Principle 3 – XAI Method Customization			
Accessibility	5.6	± 0.8	0.76
Importance	5.6	± 0.9	0.74
Novelty and Insightfulness	5.6	± 0.9	0.73
Actability and Appropriate Guidance	5.6	± 0.9	0.90
Effectiveness	5.6	± 0.9	0.88

Table 12. Reusability evaluation of the design principles (N = 61).

6. Discussion

6.1 The Design and Perception of Customization Features in Explanation User Interfaces

The DSR projects presented in this article generated and communicated a wide range of knowledge. From a DSR perspective, three design principles are introduced for customization features in XUIs. The design principles were developed based on the status quo of DSR research. For this purpose, design requirements were first derived, containing the design and XUI goals. The goals can also be described as goodness criteria, which aim to assess the solution acceptance (vom Brocke et al., 2020). The design principles have addressed the design requirements and were based on the scheme by Gregor et al. (2020) formalized. To instantiate the design principles in a technical artifact, i.e., an XUI, design features were derived that demonstrate how the design principles can be instantiated (Seidel et al., 2018). To ensure that the design principles are also useful and helpful for potential users (Chandra Kruse et al., 2022; Gregor et al., 2020), they were quantitatively evaluated by practitioners regarding reusability. The subjects had to have experience in user interface design or user experience design. The three design principles were measured independently using the dimensions of accessibility, importance, novelty and insightfulness, actability and appropriate guidance, and effectiveness. A Likert scale from 1 to 7 was used, and the design principles were rated very positively. Additionally, 52 of the 61 surveyed subjects indicated they would use the design principles for a suitable development project, and 53 indicated they would recommend them to colleagues. Furthermore, the contribution from the DSR perspective can still be characterized as a knowledge contribution in the form of exaptation since a known solution (i.e., customization) is transferred to the XAI research field (Gregor & Hevner, 2013). More specifically, the XUI with customization feature can be described as a DSR contribution of an artifact (Level 1) and the design principles as nascent design theory or knowledge as operational principles/ architecture (Level 2) based on Gregor and Hevner (2013).

The DSR projects presented in this article generated and communicated a wide range of knowledge. From a DSR perspective, three design principles are introduced for customization features in XUIs. The design principles were developed based on the status quo of DSR research. For this purpose, design requirements were first derived, containing the design and XUI goals. The goals can also be described as goodness criteria, which aim to assess the solution acceptance (vom Brocke et al., 2020). The design principles have addressed the design requirements and were based on the scheme by Gregor et al. (2020) formalized. To instantiate the design principles in a technical artifact, i.e., an XUI, design features were derived that demonstrate how the design principles can be instantiated (Seidel et al., 2018). To ensure that the design principles are also useful and helpful for potential users (Chandra Kruse et al., 2022; Gregor et al., 2020), practitioners quantitatively evaluated them regarding reusability. The subjects had to have experience in user interface design or user experience design. The three design principles were measured independently using the dimensions of accessibility, importance, novelty and insightfulness, actability and appropriate guidance, and effectiveness. A Likert scale from 1 to 7 was used, and the design principles were rated very positively. Additionally, 52 of the 61 surveyed subjects indicated they would use the design principles for a suitable development project, and 53 indicated they would recommend them to colleagues. Furthermore, the contribution from the DSR perspective can still be characterized as a knowledge contribution in the form of exaptation since a known solution (i.e., customization) is transferred to the XAI research field (Gregor & Hevner, 2013). More specifically, the XUI with customization feature can be described as a DSR contribution of an artifact (Level 1) and the design principles as nascent design theory or knowledge as operational principles/ architecture (Level 2) based on Gregor and Hevner (2013).

Concerning XAI, customization has not been rigorously conceptualized and narrowed down in previous research. However, customization has already been studied in many other areas, such as service environments (Kang & Lee, 2015), personal data curation (Vitale et al., 2020), online contexts (Ku et al., 2016; Teng et al., 2010), or customization of web portals (Kalyanaraman & Sundar, 2006; Sundar et al., 2012). In this research, customization is often associated with positive effects, for example, when the interaction experience for users is positively influenced. Concerning the perceived customization of the investigated XUIs with customization features, the statistical analysis of the first experiment shows a significant difference in perceived customization in the two groups of the between-subject experiment of the first design cycle. Furthermore, in the second experimental study, perceived customization significantly directly affected user engagement and also had a significantly positive impact on satisfaction, moderated by perceived customization. However, the other constructs that were part of the structural model helped to examine the perception of XUIs with customization features in terms of user engagement and satisfaction. Here, the results align with HCI research that found positive effects of customization (e.g., Burkolter et al., 2014; Kalyanaraman & Sundar, 2006; Kang & Lee, 2015; Sundar et al., 2015). Prior research has also shown that there can be cases where users use customization not effectively or do not use it at all (Jorritsma et al., 2015). Contrary to these findings, participants with access to customization features actively used them during the evaluation. More precisely, the customization features were actively used in both design cycles. In the first design cycle, cosmetic and functional customization was tracked by mouse clicks on the XUI elements. The results showed that users used the cosmetic customization feature significantly more often than the functional one. In the second design cycle, the three customization features were measured similarly. A comparison showed that the significant difference in the use of the cosmetic customization feature and the functional customization feature could not be reproduced. However, in the second experimental study, the XAI method customization feature was used significantly more frequently than both other customization features.

Consequently, this DSR project contributes design knowledge that could be reused in future research for artifacts of the same type (Chandra Kruse et al., 2022).

6.2 Implications for Research and Practice

The DSR project presented in this article encompasses a diverse set of knowledge about the design of XUIs with customization features and their impact on users. Design principles for these customization features were developed, instantiated, and evaluated with practitioners regarding reusability. Furthermore, a between-subject experiment was used to examine the influence of customization features in XUIs regarding perceived explanation satisfaction compared to XUIs without customization features. In a second experimental study, it was examined how user engagement is an essential part of the interaction experience with XUIs and how it influences satisfaction. Various implications for research and practice can be derived from these complex findings.

In more detail, the ex-post evaluation results of both online experiments align with HCI and ISR research that highlighted the positive effects of customization features. In different use cases, these positive effects were represented through a positive influence on aspects including the perceived enjoyment (Bailey et al., 2009), the perceived fit of the digital environment with the wants and needs of users (Kang & Lee, 2015), or the learning performance (Ku et al., 2016). The results of the first experiment showed how the customization features in XUIs positively influenced the perceived explanation quality, which was operationalized through the perceived interactivity, perceived interestingness, perceived informativeness, and the explanation satisfaction scale. Likewise, these dimensions are also essential notions of XAI and explainability (Arrieta et al., 2020; Haque et al., 2023; Minh et al., 2022). They are essential in evaluating human-centered XUIs and explanations (Gunning & Aha, 2019; Gunning et al., 2019). Therefore, the findings on the positive influence of the designed customization features in XUIs on the perception of the explanation are relevant for both research and practice. When future design-or behavioral-oriented studies or experiments related to XUIs are carried out, the DSR project presented here can provide a reference point. The second experiment examined the findings on the perception and influence of customization features in XUIs on users in more depth. It explored how perceived interactivity, perceived interestingness, and perceived customization positively influence and predict user engagement. In addition, it was examined to what extent user engagement positively influences or predicts satisfaction. The structural model developed for this purpose was evaluated and showed that all hypothesized effects could be confirmed, except for perceived interestingness. These findings can also have many implications for future XUI designs.

The DSR project has also expanded the empirical field around the Theory of Interactive Media Effects topic because this theory was essential to conceptualizing customization features for XUIs (Sundar et al., 2015; Sundar, 2020). The theory also influenced relevant decisions during the development of design knowledge and was used to anchor the design principles. Ultimately, the two experimental studies contribute to the growing research on the Theory of Interactive Media Effects. The focus was on the investigation and empirical evaluation of the interaction experience with the customization features in XUIs and their effect on the users. Because XUIs are context and application-independent, these findings can help inform future research projects.

Another relevant implication of the DSR project is to examine the design of explanations in a larger context. XAI methods are often carried out in simplified experiments (e.g., Ribeiro et al., 2016). However, perception through explanation can be influenced by many different factors. The experimental study from the first design cycle showed how the customization features led to significantly better scores for perceived interactivity, perceived interestingness, perceived informativeness, and the explanation satisfaction scale. Therefore, many factors must be considered when developing human-centered XAI systems or XUIs (Vilone & Longo, 2021). Different users have varying needs, backgrounds, and expectations when interacting with XUIs, which must be considered (Langer et al., 2021; Meske et al., 2022).

6.3 Limitations and Future Research Opportunities

Although this study adhered to established guidelines, including the recommendations for communicating DSR projects (vom Brocke & Maedche, 2019; Hevner et al., 2004), followed an established DSR process (Kuechler & Vaishnavi, 2012) and formalized the design principles according to the scheme of Gregor et al. (2020), the DSR project has limitations. For example, despite the XUI design being based on existing research with a focus on XUIs (e.g., Dikmen & Burns, 2020; Harbers et al., 2010; Kim et al., 2014; van der Waa et al., 2021), a somewhat simplified prototypical XUI was developed for the experiments. Future research can use the introduced prescriptive design knowledge to develop more mature XUIs for new use cases. Moreover, online experiments were chosen for the ex-post evaluation of the proposed design. Experimental research is well-established in the information systems discipline, and online experiments have evolved into an essential methodology for studying human behavior (e.g., human-computer interaction), which justifies the evaluation design (Fink, 2022; Karahanna et al., 2018). However, future research can adapt the design and investigate the influence of customization features in XUIs using methods like field studies, laboratory experiments, or qualitative methods like interviews.

In connection with evaluations and experiments in future research with customization features in XUIs, further notions of explainability can be investigated, including justifiability, user's mental model, actionability, or cognitive relief (Gunning & Aha, 2019; Minh et al., 2022; Vilone & Longo, 2021). Since user studies are still scarce in the research stream of XAI (van der Waa et al., 2021; Wells & Bednarz, 2021), they can lead to valuable contributions to research and practice. The use case for the ex-post evaluation with the classification of attrition risks was simplified. Nonetheless, it was a realistic use case since AI is already used and investigated in human resources, for example, to support hiring processes (e.g., Sipior et al., 2021). Since AI is already established and researched in different areas, there is the potential to use customization features for XUIs in these areas. For example, XAI has already arrived in areas like healthcare (Barda et al., 2020; Schoonderwoerd et al., 2021), manufacturing (Senoner et al., 2022), transport, or finance (Adadi & Berrada, 2018). Consequently, the concept of customization in XUIs may reveal hidden potential regarding the interaction of humans with XAI-based systems in the before-named and further areas.

Similar to previous research in HCI, customization features in XUIs have also been shown to influence the explanation's perception positively. However, it is also important to emphasize that customization as a concept also has its downsides, which need to be considered in future research on customization for XAI. For example, users only sometimes effectively use customization or do not customize (Jorritsma et al., 2015). Prior research has also shown that customization requires users' active exercise of choice. When users have to make constant personal choices, it can deplete the inner resources required for self-control (Kang & Sundar, 2013). In addition, the expertise of the targeted users can also influence their perception of customization features in UIs (Fukazawa et al., 2009). Moreover, it is crucial to consider the number of customization features as they can call for additional forms of effortful decision-making and can ultimately lead to decision fatigue (Sundar et al., 2012). Consequently, future research on customization features in XUIs needs to consider the potential adverse effects of customization features. Lastly, in this study, the concept of customization as a whole was the object of investigation. Therefore, future research in the context of customization for XAI can also take a more nuanced perspective and investigate the effects of cosmetic and functional customization independently. This could also lead to exciting insights since the analysis of the ex-post evaluation showed that the cosmetic customization feature was significantly more frequently used than the functional customization feature. In addition, hybrid customization could be an exciting approach to designing AI-generated explanations.

7. Conclusion

In this DSR study, customization features for XUIs were conceptualized, instantiated, and examined in two experimental studies using online experiments. The structured and iterative DSR process by Kuechler and Vaichnavi (2012) was used for this purpose. Three design principles for customization features were developed, and their reusability was evaluated by practitioners and rated positively. The instantiated design principles in the first design cycle led to an XUI with the two design principles for cosmetic and functional customization. A between-subject experiment was used to examine how much customization features influence the perceived explanation quality. The participants in the group rated all constructs that were operationalized for this purpose significantly better if they had access to customization features. A second experimental study examined how dimensions relevant to XAI and XUI, such as perceived interactivity or perceived customization, influence user engagement. It has been shown here that these two constructs predict user engagement and have a significantly positive influence. The same effect was also found for user engagement from satisfaction. The article thus presents the complex insights generated during the DSR project.

References

Adadi, A., & Berrada, M. (2018). Peeking Inside the Black-Box: A Survey on Explainable Artificial Intelligence (XAI). *IEEE Access, 6*, 52138-52160. https://doi.org/10.1109/ACCESS.2018.2870052

Adam, M. T. P., Gregor, S., Hevner, A., & Morana, S. (2021). Design Science Research Modes in Human-Computer Interaction Projects. *AIS Transactions on Human-Computer Interaction, 13*(1), 1-11. https://doi.org/10.17705/1thci.00139

Ammar, N., & Shaban-Nejad, A. (2020). Explainable Artificial Intelligence Recommendation System by Leveraging the Semantics of Adverse Childhood Experiences: Proof-of-Concept Prototype Development. *JMIR Medical Informatics, 8*(11), Article e18752. https://doi.org/10.2196/18752

Anderson, J. C., & Gerbig, D. W. (1988). Structural equation modeling in practice: A review and recommended two-step approach. *Psychological Bulletin, 103*(3), 411-423. https://doi.org/10.1307/0033-2909.103.3.411

Al-Natour, S., Benbasat, I., & Cenfetelli, R. T. (2022). Designing Caring and Informative Decision Aids to Increase Trust and Enhance the Interaction Atmosphere. *AIS Transactions on Human-Computer Interaction, 14*(1), 1-29. https://doi.org/10.17705/1thci.00159

Alicioglu, G., & Sun, B. (2022). A survey of visual analytics for Explainable Artificial Intelligence methods. *Computer & Graphics, 102*, 505-520. https://doi.org/10.1016/j.cag.2021.09.002

Alsallakh, B., Hanbury, A., Hauser, H., Miksch, S., & Rauber, A. (2014). Visual Methods for Analyzing Probabilistic Classification Data. *IEEE Transactions on Visualization and Computer Graphics, 20*(12), 1703-1712. https://doi.org/10.1109/TVCG.2014.2346660

Alufaisan, Y., Marusich, L. R., Bakdash, J. Z., Zhou, Y., & Kantarcioglu, M. (2021). Does Explainable Artificial Intelligence Improve Human Decision-Making? *Proceedings of the AAAI Conference on Artificial Intelligence, 35*(8), 6618-6626. https://doi.org/10.1609/aaai.v35i8.16819

Arapakis, I., Lalmas, M., & Valkanas, G. 82014). Understanding Within-Content Engagement through Pattern Analysis of Mouse Gestures. *Proceedings of the 23rd ACM International Conference on Information and Knowledge Management*, 1439-1448. https://doi.org/10.1145/2661829.2661909

Arrieta A. B., Diaz-Rodrigeuz, N., Del Ser, J., Bennetot, A., Tabik, S., Barbado, A., Garcia, S., Gil-Lopez, S., Molina, D., Benjamins, R., Chatila, R., & Herrera, F. (2020). Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI. *Information Fusion, 58*, 82-115. https://doi.org/10.1016/j.inffus.2019.12.012

Bailey, R., Wise, K., & Bolls, P. (2009). How Avatar Customizability Affects Children's Arousal and Subjective Presence During Jung Food-Sponsored Online Video Games. *CyberPsychology & Behavior, 12*(3), 277-283. https://doi.org/10.1089/cpb.2008.0292

Bandura, A., & Schunk, D. H. (1981). Cultivating competence, self-efficacy, and intrinsic interest through proximal self-motivation. *Journal of Personality and Social Psychology, 41*(3), 586-598. https://doi.org/10.1037/0022-3514.41.3.586

Barda, A. J., Horvat, C. M., & Hochheiser, H. (2020). A qualitative research framework for the design of user-centered displays of explanations for machine learning model predictions in healthcare. *BMC Medical Informatics and Decision Making, 20*, Article 257. https://doi.org/10.1186/s12911-020-01276-x

Baskerville, R., Baiyere, A., Gregor, S., Hevner, A., & Rossi, M. (2018). Design Science Research Contributions: Finding a Balance between Artifact and Theory. *Journal of the Association for Information Systems, 19*(5), 358-376. https://doi.org/10.17705/1jais.00495

Baskerville, R., & Pries-Heje, J. (2010). Explanatory Design Theory. *Business Information Systems & Engineering, 2*, 271-282. https://doi.org/10.1007/s12599-010-0118-4

Bau, D., Zhou, B., Kosla, A., Oliva, A. (2017). Network Dissection: Quantifying Interpretability of Deep Visual Representations. *2017 IEEE Conference on Computer Vision and Pattern Recognition*, 3319-3327. https://doi.org/10.1109/CVPR.2017.354

Bolin, M., Webber, M., Rha, P., Wilson, T., & Miller, R. C. (2005). Automation and customization of rendered web pages. *Proceedings of the 18th annual ACM symposium on User Interface software and technology*, 163-172. https://doi.org/10.1145/1095034.1095062

Bradley, C., Wu, D., Tang, H., Singh, I., Wydant, K., Capps, B., Wong, K., Agostinelli, F., Irvin, M., & Srivastava, B. (2022). Explainable Artificial Intelligence (XAI) User Interface Design for Solving a Rubik's Cube.

Stephanidis, C., Antona, M., Ntoa, S., Salvendy, G. (Eds.), Communications in Computer and Information Science: vol. 16655. *International Conference on Human-Computer Interaction* (pp. 605-612). Springer, Cham. https://doi.org/10.1007/978-3-031-19682-9_76

vom Brocke, J., Simons, A., Riemer, K., Niehaves, B., Plattfaut, R., & Cleven, A. (2015). Standing on the Shoulders of Giants: Challenges and Recommendations of Literature Search in Information Systems Research. *Communications of the Association for Information Systems, 37*, 205-224. https://doi.org/10.17705/1CAIS.03709

vom Brocke, & Maedche, A. (2019). The DSR grid: six core dimensions for effectively planning and communicating design science research projects. *Electronic Markets, 29*, 379-385. https://doi.org/10.1007/s12525-019-00358-7

vom Brocke, Winter, R., Hevner, A., & Maedche, A. (2020). Special Issue Editorial – Accumulation and Evolution of Design Knowledge in Design Science Research: A Journey Through Time and Space. *Journal of Association for Information Systems, 21*(3), 520-544. https://doi.org/10.17705/1jais.00611

Bradley, C., Wu, D., Tang, H., Singh, I., Wydant, K., Capps, B., Wong, K., Agostinelli, F., Irvin, M., & Srivastava, B. (2022). Explainable Artificial Intelligence (XAI) User Interface Design for Solving Rubik's Cube. In Stephanidis, C., Antona, M., Ntoa, S., Salvendy, G. (Eds.) HCI International 2022 – Late Breaking Posters. HCII 2022. *Communications in Computer and Information Science*, vol. 1655. Springer, Cham. https://doi.org/10.1007/978-3-031-19682-9_76

Bunde, E. (2021). AI-Assisted and Explainable Hate Speech Detection for Social Media Moderators – A Design Science Approach. *Proceedings of the 54th Hawaii International Conference on System Sciences*, 1264-1273. http://hdl.handle.net/10125/70766

Bunde, E., Eisenhardt, D., Sonntag, D., Profitlich, H.-J., & Meske, C. (2023). Giving DIAnA More TIME – Guidance for the Design of XAI-Based Medical Decision Support Systems. Gerber, A., Baskerville, R. (Eds.), Lecture Notes in Computer Science: vol. 13873, *Design Science Research for a New Society: Society 5.0* (pp. 107-122). https://doi.org/10.1007/978-3-031-32808-4_7

Bunde, E., Kühl, N., & Meske, C. (2022). Fake or Credible? Towards Designing Services to Support Users' Credibility Assessment of News Content. In *Proceedings of the 55th Hawaii International Conference on System Sciences 2022*, pp. 1883-1892. http://hdl.handle.net/10125/79569

Burkolter, D., Weyers, B., Kluge, A., & Luther, W. (2014). Customization of user interfaces to reduce errors and enhance user acceptance. *Applied Ergonomics, 45*, 346-353. https://doi.org/10.1016/j.apergo.2013.04.017

Chandra Kruse, L., Sandeep, P., & Seidel, S. (2022). How Designers Use Design Principles: Design Behaviors and Application Modes. *Journal of the Association for Information Systems, 23*(5), 1235-1270. https://doi.org/10.17705/1jais.00759

Chromik, M., & Butz, A. (2021). Human-XAI Interaction: A Review and Design Principles for Explanation User Interfaces. In Ardito, C., Lanzilotti, R., Malizia, A., Petrie, H., Piccinno, A., Desolda, G., Inkpen, K. (Eds.), Lecture Notes in Computer Science: vol. 12933 (pp. 619-640). *Human-Computer Interaction – INTERACT 2021*. Springer, Cham. https://doi.org/10.1007/978-3-030-85616-8_36

Chung, K.-H., & Chin, J.-I. (2008). The Relationship among e-Retailing Attributes, e-Satisfaction and e-Loyalty. *Management Review: An International Journal, 3*(1), 23-45.

Constantin, M. G., Redi, M., Zen, G., & Ionescu, B. (2019). Computational Understanding of Visual Interestingness Beyond Semantics: Literature Survey and Analysis of Covariates. *ACM Computing Surveys, 52*(2), Article 25. https://doi.org/10.1145/3301299

Cramer, H., Evers, V., Ramlal, S., van Someren, M., Rutledge, L., Stash, N., Aroyo, L., & Wielinga, B. (2008). The effects of transparency on trust in and acceptance of a content-based art recommender. *User Modeling and User-Adapted Interaction, 18*, 455-496. https://doi.org/10.1007/s11257-008-9051-3

Darejeh, A., & Salim, S. S. (2016). Gamification Solutions to Enhance Software User Engagement – A Systematic Review. *International Journal of Human-Computer Interaction, 32*(8), 613-642. https://doi.org/10.1080/10447318.2016.1183330

Dikmen, M., & Burns, C. (2020). Abstraction Hierarchy Based Explainable Artificial Intelligence. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 64*(1), 319-323. https://doi.org/10.1177/1071181320641073

Dudley, J. J., & Kristensson, P. O. (2018). A Review of User Interface Design for Interactive Machine Learning. *ACM Transactions on Interactive Intelligent Systems, 8*(2), Article 8. https://doi.org/10.1145/3185517

Ferreira, J. J., & Monteiro, M. S. (2020). What Are People Doing About XAI User Experience? A Survey on AI Explainability Research in Practice. In Marcus, A., Rosenzweig, E. (Eds.), Lecture Notes in Computer Science: vol. 12201. *Design, User Experience, and Usability. Design for Contemporary Interactive Environments* (pp. 56.73). Springer, Cham. https://doi.org/10.1007/978-3-030-49760-6_4

Fink, L. 82022). Why and How Online Experiments Can Benefit Information Systems Research. *Journal of the Association for Information Systems, 23*(6), 1333-1346. https://doi.org/10.17705/1jais.00787

Freitas, A. A. (2006). Are We Really Discovering "Interesting" Knowledge From Data?. 1-6.

Fukazawa, Y., Mara, M., Onogi, M., & Ueno, H. (2009). Automatic mobile menu customization based on user operation history. *Proceedings of the 11th International Conference on Human-Computer Interaction with Mobile Devices and Services*, Article 50. https://doi.org/10.1145/1613858.1613921

Füßl, A., Nissen, V., & Heringklee, S. H. (2023). Interactive Machine Learning of Knowledge Graph-Based Explainable Process Analysis. In Ruiz, M., Soffer, P. (Eds.), Lecture Notes in Business Information Processing: vol. 482. *Advanced Information Systems Engineering Workshops* (pp. 112-124). Springer, Cham. https://doi.org/10.1007/978-3-031-34985-0_12

Gregor, S., Chandra Kruse, L., & Seidel, S. (2020). Research Perspectives: The Anatomy of a Design Principle. *Journal of the Association for Information Systems, 21*(6), 1622-1652. https://doi.org/10.17705/1jais.00649

Gregor, S., & Hevner, A. R. (2013). Positioning and Presenting Design Science Research for Maximum Impact. *MIS Quarterly, 37*(2), 337-355. https://doi.org/10.25300/MISQ/2013/37.2.01

Gregor, S., & Jones, D. (2007). The Anatomy of a Design Theory. *Journal of the Association for Information Systems, 8*(2), 312-335. https://doi.org/10.17705/1jais.00129

Gunning, D., & Aha, D. (2019). DARPA's Explainable Artificial Intelligence (XAI) Program. *AI Magazine, 40*(2), 44-58. https://doi.org/10.1609/aimag.v40i2

Gunning, D., Stefik, M., Choi, J., Miller, T., Stumpf, S., & Yang, G-Z. (2019). XAI – Explainable Artificial Intelligence. *Science Robotics, 4*(37), eaay7120. https://doi.org/10.1126/scirobotics.aay7120

Ha, T., Sah, Y. J., Park, Y., & Lee, S. (2022). Examining the effects of power status of an explainable artificial intelligence system on users' perceptions. *Behavior & Information Technology, 41*(5), 946-958. https://doi.org/10.1080/0144929X.2020.1846789

Habers, M., van den Bosch, K., & Meyer, J.-J. (2010). Design and Evaluation of Explainable BDI Agents. *2010 IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology*, 125-132. https://doi.org/10.1109/WI-IAT.2010.115

Hair, J. F., Hult, G. T. M., Ringle, C M., & Sarstedt, M. (2022). *A Primer on Partial Least Squares Structural Equation Modeling (PLS-SEM)* (3 ed.). Thousand Oaks, CA. Sage.

Haque, A. K. M. B., Islam, A. K. M. N., & Mikalef, P. (2023). Explainable Artificial Intelligence (XAI) from a user perspective: A synthesis of prior literature and problematizing avenues for future research. *Technological Forecasting and Social Change, 186*(Part A), Article 122120. https://doi.org/10.1016/j.techfore.2022.122120

Hauser, D. J., Ellsworth, P. C., & Gonzalez, R. (2018). Are Manipulation Checks Necessary?. *Frontiers in Psychology, 9*, 1-10. https://doi.org/10.3389/fpsyg.2018.00998

Henseler, J., Ringle, C. M., & Sarstedt, M. (2015). A New Criterion for Assessing Discriminant Validity in Variance-based Structural Equation Modeling. *Journal of the Academy of Marketing Science, 43*(1), 115-135. https://doi.org/10.1007/s11747-014-0403-8

Hevner, A. R., March, S. T., Park, J., & Ram, S. (2004). Design science in information systems research. *MIS Quarterly, 28*(1), 75-105.

Hoffman, R. R., Mueller, S. T., Klein, G., & Litman, J. (2018). Metrics for Explainable AI: Challenges and Prospects. *Technical Report, DARPA Explainable AI Program*. https://doi.org/10.48550/arXiv.1812.04608

Holzinger, A. (2016). Interactive machine learning for health informatics: when do we need the human-in-the-loop?. *Brain Informatics, 3*, 119-131. https://doi.org/10.1007/s40708-016-0042-6

Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal, 6*(1), 1-55. https://doi.org/10.1080/10705519909540118

Hui, S. L. T., & See, S. L. (2015). Enhancing User Experience Through Customisation of UI Design. *Procedia Manufacturing, 3*, 1932-1937. https://doi.org/10.1016/j.promfg.2015.07.237

Huysmans, J., Dejaeger, K., Mues, C., Vanthienen, & Baesens, B. (2011). An empirical evaluation of the comprehensibility of decision table, tree and rule based predictive models. *Decision Support Systems, 51*, 141-154. https://doi.org/10.1016/j.dss.2010.12.003

Ivari, J. (2020). Editorial: A Critical Look at Theories in Design Science Research. *Journal of the Association for Information Systems, 21*(3), 502-519. https://doi.org/10.17705/1jais.00610

Ivari, J., Hansen, M. R. P., & Haj-Bolouri, A. (2018). A Framework for Light Reusability Evaluation of Design Principles in Design Science Research. Chatterjee, S., Dutta, K., Sundarraj, R. P. (Eds.), Lecture Notes in Computer Science: vol. 10844. *DESRIST International Conference on Design Science Research in Information Systems and Technology* (pp. 1-16). Springer, Cham. https://doi.org/10.1007/978-3-319-91800-6

Ivari, J., Hansen, M. R. P., & Haj-Bolouri, A. (2021). A proposal for minimum reusability evaluation of design principles. *European Journal of Information Systems, 30*(3), 286-303. https://doi.org/10.1080/0960085X.2020.1793697

Jorritsma, W., Cnossen, F., & van Ooijen, P. M. A. (2015). Adaptive support for user interface customization: a study in radiology. *International Journal of Human-Computer Studies, 77*, 1-9. https://doi.org/10.1016/j.ijhcs.2014.12.008

Kaggle (2023a). *IBM HR Analytics Employee Attrition & Performance* [Data set]. https://www.kaggle.com/datasets/pavansubhasht/ibm-hr-analytics-attrition-dataset

Kaggle (2023b). *Salary Prediction Classification* [Data set]. https://www.kaggle.com/datasets/ayessa/salary-prediction-classification

Kalyanaraman, S., & Sundar, S. S. (2006). The Psychological Appeal of Personalized Content in Web Portals: Does Customization Affect Attitudes and Behavior? *Journal of Communication, 56*(1), 110-132. https://doi.org/10.1111/j.1460-2466.2006.00006.x

Kang, H., & Kim, H. K. (2020). My avatar and the affirmed self: Psychological and persuasive implications of avatar customization. *Computer in Human Behavior, 112*, Article 106446. https://doi.org/10.1016/j.chb.2020.106446

Kang, Y. J., & Lee, W. J. (2015). Self-customization of online service environments by users and its effect on their continuance intention. *Service Business, 9*, 321-342. https://doi.org/10.1007/s11628-014-0229-y

Kang, H., & Lou, C. (2022). AI agency vs. Human agency: understanding human-AI interactions on TikTok and their implications for user engagement. *Journal of Computer-Mediated Communication, 27*(5), Article zmac014. https://doi.org/10.1093/jcmc/zmac014

Kang, H., & Sundar, S. S. (2013). Depleted egos and affirmed selves: The two faces of customization. *Computers in Human Behavior, 29*(6), 2273-2280. https://doi.org/1016/j.chb.2013.05.018

Karahanna, E., Benbasat, I., Bapna, R., & Rai, A. (2018). Opportunities and Challenges for Different Types of Online Experiments. *MIS Quarterly, 42*(4), iii-x.

Karnowski, V., Kümpel, A. S., Leonhard, L., & Leiner, D. J. (2017). From incidental news exposure to news engagement. How perceptions of the news post and news usage patterns influence engagement with news articles encountered on Facebook. *Computers in Human Behavior, 76*, 42-50. https://doi.org/10.1016/j.chb.2017.06.041

Katerina, T., Nicolaos, P., & Charalampos, Y. (2014). Mouse Tracking for Web Marketing: Enhancing User Experience in Web Application Software by Measuring Self-Efficacy and Hesitation Levels. *International Journal on Strategic Innovative Marketing, 1*, 233-247. https://doi.org/10.15556/IJSIM

Kaur, H., Nori, H., Jenkins, S., Caruana, R., Wallach, H., & Vaighan, J. W. (2020). Interpreting Interpretability: Understanding Data Scientists' Use of Interpretability Tools for Machine Learning. *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, 1-14. https://doi.org/10.1145/3313831.3376219

Keneni, B. M., Kaur, D., Bataineh, A. A., Devabhaktuni, V. K., Javaid, A. Y., Zaientz, J. D., & Marinier, R. P. (2019). Evolving Rule-Based Explainable Artificial Intelligence for Unmanned Aerial Vehicles. *IEEE Access, 7*, 17001-17016. https://doi.org/10.1109/ACCESS.2019.2893141

Kim, B., Rudin, C., & Shah, J. (2014). The Bayesian case model: a generative approach for case-based reasoning and prototype classification. *Proceedings of the 27th International Conference on Neural Information Processing Systems, 2*, 1952-1960. https://doi.org/10.48550/arXiv.1503.01161

Kleinsmith, A., & Gillies, M. (2013). Customizing by doing for responsive video game characters. *International Journal of Human-Computer Studies, 71*(7-8), 775-784. https://doi.org/10.1016/j.ijhcs.2013.03.005

Knapic, S., Malhi, A., Saluja, R., & Främling, K. (2021). Explainable Artificial Intelligence for Human Decision Support System in the Medical Domain. *Machine Learning & Knowledge Extraction, 3*, 740-770. https://doi.org/10.3390/make303037

Ku, O., Hou, C.-C., & Chen, S. Y. (2016). Incorporating customization and personalization into game-based learning: A cognitive style perspective. *Computers in Human Behavior, 65*, 359-368. https://doi.org/10.1016/j.chb.2016.08.040

Kuechler, W., & Vaishnavi, V. (2012). A Framework for Theory Development in Design Science Research: Multiple Perspectives. *Journal of the Association for Information Systems, 13*(6), 395-423. https://doi.org/10.17705/1jais/vol13/iss6/3

Langer, M., Oster, D., Speith, T., Hermanns, H., Kästner, L., Schmidt, E., Sesing, A., & Baum, K. (2021). What do we want from Explainable Artificial Intelligence (XAI)? – A stakeholder perspective on XAI and conceptual model guiding interdisciplinary XAI research. *Artificial Intelligence, 296*, Article 103473. https://doi.orG/10.1016/j.artint.2021.103473

Leichtmann, B., Humer, C., Hinterreiter, A., Streit, M., & Mara, M. (2023). Effects of Explainable Artificial Intelligence on trust and human behavior in a high-risk decision task. *Computers in Human Behavior, 139*, Article 107539. https://doi.org/10.1016/j.chb.2022.107539

Li, M., & Gregor, S. (2011). Outcomes of effective explanations: Empowering citizens through online advice. *Decision Support Systems, 52*(1), 119-132. https://doi.org/10.1016/j.dss.2011.06.001

Li, W., Mao, Y., & Zhou, L. (2021). The Impact of Interactivity on User Satisfaction in Digital Social Reading: Social Presence as a Mediator. *International Journal of Human-Computer Interaction, 37*(17), 1636-1647. https://doi.org/10.1080/10447318.2021.1898850

Lim, B. Y., Dey, A. K., & Avrahami, D. (2009). Why and why not explanations improve the intelligibility of context-aware intelligent systems. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 2119-2128. https://doi.org/10.1145/1518701.1519023

Lipton, Z. C. (2018). The mythos of model interpretability. *Communications of the ACM, 61*(10), 36-43. https://doi.org/10.1145/3233231

Lopez-Giraldez, F., & Townsend, J. P. (2011). PhyDesign: an online application for profiling phylogenetic informativeness. *BMC Evolutionary Biology, 11*, Article 152. https://doi.org/10.1186/1471-2148-11-152

Macias, J. A., & Paterno, F. (2008). Customization of Web applications through an intelligent environment exploiting logical interface descriptions. *Interacting with Computers, 20*(1), 29-47. https://doi.org/10.1016/j.intcom.2007.07.007

Marathe, S., & Sundar, S. S. (2011). What drives customization?: control or identify? *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 781-790. https://doi.org/10.1145/1978942.1979056

Masrek, M. N., Razali, M. H., Ramli, I., & Andromeda, T. (2018). User Engagement and Satisfaction: The Case of Web Digital Library. *International Journal of Engineering & Technology, 7*(4.15), 19-24. https://doi.org/10.14419/ijet.v7i4.1521364

McCay-Peet, L., Toms, E. G., & Kelloway, E. K. (2015). Examination of relationship among serendipity, the environment, and individual differences. *Information Processing & Management, 51*(4), 391-412. https://doi.org/10.1016/j.ipm.2015.02.004

Meske, C., & Bunde, E. (2023). Design Principles for User Interfaces in AI-Based Decision Support Systems: The Case of Explainable Hate Speech Detection. *Information Systems Frontiers*, *25*, 743-773 https://doi.org/10.1007/s10796-021-10234-5

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. https://doi.org/10.1080/10580530.2020.1849465

Meth, H., Mueller, B., & Maedche, A. (2015). Designing a Requirement Mining System. *Journal of the Association for Information Systems, 16*(9), 799-837. https://doi.org/10.17705/1jais.00408

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

Minh, D., Wang, H. X., Li, Y. F., & Nguyen, T. N. (2022). Explainable artificial intelligence: a comprehensive review. *Artificial Intelligence Review, 55*, 3503-3568. https://doi.org/10.1007/s10462-021-10088-y

Mohseni, S., Zarei, N., & Ragan, E. D. (2021). A Multidisciplinary Survey and Framework for Design and Evaluation of Explainable AI Systems. *ACM Transactions on Interactive Intelligent Systems, 11*(3-4), Article 24. https://doi.org/10.1145/3387166

Naiseh, M., Al-Thani, D., Jiang, N., & Ali, R. (2023). How the different explanation classes impact trust calibration: The case of clinical decision support systems. *International Journal of Human-Computer Studies, 169*, Article 102941. https://doi.org/10.1016/j.ijhcs.2022.102941

Niu, X., & Al-Doulat, A. (2021). LuckyFind: Leveraging Surprise to Improve User Satisfaction and Inspire Curiosity in a Recommender System. *Proceedings of the 2021 Conference on Human Information Interaction and Retrieval*, 163-172. https://doi.org/10.1145/3406522.3446017

Noori, W. E., & Albahri, A. S. (2023). Towards Trustworthy Myopia Detection: Integration Methodology of Deep Learning Approach, XAI Visualization, and User Interface System. *Applied Data Science and Analysis, 2023*, 1-15. https://doi.org/10.58496/ADSA/2023/001

O'Brien, H. L., & Toms, E. G. (2008). What is user engagement? A conceptual framework for defining user engagement with technology. *Journal of the American Society for Information Science and Technology, 59*(6), 938-955. https://doi.org/10.1002/asi.20801

Oh, J., & Sundar, S. S. (2015). How Does Interactivity Persuade? An Experimental Test of Interactivity on Cognitive Absorption, Elaboration, and Attitudes. *Journal of Communication, 65*(2), 213-236. https://doi.org/10.1111/jcom.12147

Oppenheimer, D. M., Meyvis, T., & Davidenko, N. (2009). Instructional manipulation checks: Detecting satisficing to increase statistical power. *Journal of Experimental Social Psychology, 45*(4), 867-872. https://doi.org/10.1016/j.jesp.2009.09.009

Panigutti, C., Beretta, A., Fadda, D., Giannotti, F., Pedreschi, D., Perotti, A., & Rinzivillo, S. (2023). Co-design of human-centered, explainable AI for clinical decision support. *ACM Transactions on Interactive Intelligent Systems*, 1-34. https://doi.org/10.1145/3587271

Peer, E., Rothschild, D., Gordon, A., Evernden, Z., & Damer, E. (2022). Data quality of platforms and panels for online behavioral research. *Behavior Research Methods, 54*, 1643-1662. https://doi.org/10.3758/s13428-021-01694-3

Peffers, K., Tuunanen, T., Rothenberger, M. A., & Chatterjee, S. (2007). A Design Science Research Methodology for Information Systems Research. *Journal of Management Information Systems, 24*(3), 45-77. https://doi.org/10.2753/MIS0742-1222240302

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., & Cournapeau (2011). Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research, 12*, 2825-2830.

Poursabzi-Sangdeh, F., Goldstein, D. G., Hofman, J. M., Vaughan, J. W. W., & Wallach, H. (2021). Manipulating and Measuring Model Interpretability. *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems,* 1-52. https://doi.org/10.1145/3411764.3445315

Ribeiro, M. T., Singh, S., & Guestrin, C. (2016). "Why Should I Trust You?": Explaining the Predictions of Any Classifier. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining,* 1135-1144). https://doi.org/10.1145/2939672.2939778

Ribeiro, M. T., Singh, S., & Guestrin, C. (2018). Anchors: High-Precision Model-Agnostic Explanations. *Proceedings of the AAAI Conference on Artificial Intelligence, 32*(1), 1527-1535. https://doi.org/10.1609/aaai.v32i1.11491

Rivera, D. (2005). The effect of content customization on learnability and perceived workload. *Extended Abstracts on Human Factors in Computing Systems*, 1749-1752. https://doi.org/10.1145/1056808.1057013

Schoonderwoerd, T. A. J., Jorritsma, W., Neerincx, M. A., & van den Bosch, K. (2021). Human-centered XAI: Developing design patterns for explanations of clinical decision support systems. *International Journal of Human-Computer Studies, 154*. Article 102684. https://doi.org/10.1016/j.ijhcs.2021.102684

Schreiber, J. B. (2008). Core reporting practices in structural equation modeling. *Research in Social and Administrative Pharmacy, 4*(2), 83-97. https://doi.org/10.1016/j.sapharm.2007.04.003

Seidel, S., Chandra Kruse, L., Szekely, N., Gau, M., & Stieger, D. (2018). Design principles for sensemaking support systems in environmental sustainability transformations. *European Journal of Information Systems, 27*(2), 221-247. https://doi.org/10.1057/s41303-017-0039-0

Senoner, J., Netland, T., & Feuerriegel, S. (2022). Using Explainable Artificial Intelligence to Improve Process Quality: Evidence from Semiconductor Manufacturing. *Management Science, 68*(8), 5704-5723. https://doi.org/10.1287/mnsc.2021.4190

Sheng, H., & Joginapelly, T. (2012). Effects of Web Atmospheric Cues on Users' Emotional Responses in E-Commerce. *AIS Transactions on Human-Computer Interaction, 4*(1), 1-24. https://aisel.aisnet.org/thci/vol4/iss1/1

Shin, H., Bunosso, I., & Levine, L. R. (2022). The influence of chatbot humour on consumer evaluations of services. *International Journal of Consumer Studies*, 1-18. https://doi.org/10.1111/ijcs.12849

Sipior, J. C., Lombardi, D. R., & Gabryelczyk, R. (2021). AI Recruiting Tools at ShipIt2Me.com. *Communications of the Association for Information Systems, 48*(39), 443-455. https://doi.org/10.7705/1CAIS.04839

Song, X., Yu, A. S. L., Kellum, J. A., Waitman, L. R., Matheny, M. E., Simpson, S. Q., Hu, Y., & Liu, M. (2020). Cross-site transportability of an explainable artificial intelligence model for acute kidney injury prediction. *Nature Communications, 11*(5668), 1-12. https://doi.org/10.1038/s41467-020-19551-w

Sonnenberg, C., & vom Brocke, J. (2012). Evaluation in the Science of the Artificial – Reconsidering the Build-Evaluate Pattern in Design Science Research. In Peffers, K., Rothenberger, M., Kuechler, B. (Eds.), Lecture Notes in Computer Science: vol. 7286. *Design Science Research in Information Systems. Advances in Theory and Practice*. DESRIST 2012 (pp. 381-397). Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-29863-9_28

Spinner, T., Schlegel, U., Schäfer, H., & El-Assady, M. (2020). explAIner: A Visual Analytics Framework for Interactive and Explainable Machine Learning. *IEEE Transactions on Visualization and Computer Graphics, 26*(1), 1064-1074. https://doi.org/10.1109/TVCG.2019.2934629

Sundar, S. S. (2020). Rise of Machine Agency: A Framework for Studying the Psychology of Human-AI Interaction (HAII). *Journal of Computer-Mediated Communication, 25*(1), 74-88. https://doi.org/10.1093/jcmc/zmz026

Sundar, S. S., Jia, H., Waddell, T. F., & Huang, Y. (2015). Toward a Theory of Interactive Media Effects (TIME) – Four Models of Explaining How Interface Features Affect User Psychology. In Sundar, S. S. (Eds.), *The Handbook of the Psychology of Communication Technology* (1st ed., pp. 47-86). John Wiley & Sons, Inc. https://doi.org/10.1002/9781118426456.ch3

Sundar, S. S., Jia, H., Bellur, S., Oh, J., & Kim, H.-S. (2022). News Informatics: Engaging Individuals with Data-Rich News Content through Interactivity in Source, Medium, and Message. *Proceedings of the 2022 Conference on Human Factors in Computing Systems*, Article 317. https://doi.org/10.1445/3491102.32502207

Sundar, S. S., & Marathe, S. S. (2010). Personalization versus Customization: The Importance of Agency, Privacy, and Power Usage. *Human Communication Research, 36*(3), 298-322. https://doi.org/10.1111/j.1468-2958.2010.01377.x

Sundar, S. S., Oh, J., Bellur, S., Jia, H., & Kim., H.-S. (2012). Interactivity as self-expression: a field experiment with customization and blogging. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems 2012*, 395-404. https://doi.org/10.1145/2207676.2207731

Teng, C.-I. (2010). Customization, immersion satisfaction, and online gamer loyalty. *Computers in Human Behavior, 26*(6), 1547-1554. https://doi.org/10.1016/j.chb.2010.05.029

Theodosiou, M., Katsikea, E., Samiee, S., & Makri, K. (2019). A Comparison of Formative Versus Reflective Approaches for the Measurement of Electronic Service Quality. *Journal of Interactive Marketing, 47*, 53-67. https://doi.org/10.1016/j.intmar.2019.03.004

Venable, J., Pries-Heje, J., & Baskerville, R. (2016). FEDS: a Framework for Evaluation in Design Science Research. *European Journal of Information Systems, 25*, 77-89. https://doi.org/10.1057/ejis.2014.36

Vilone, G., & Longo, L. (2021). Notions of explainability and evaluation approaches for explainable artificial intelligence. *Information Fusion, 76*, 89-106. https://doi.org/10.1016/j.inffus.2021.05.009

Vilone, G., & Longo, L. (2022). A Novel Human-Centred Evaluation Approach and and Argument-Based Method for Explainable Artificial Intelligence. In Maglogiannis, I., Iliadis, L., Macintyre, J., Cortez, P. (Eds.), IFIP

Advances in Information and Communication Technology: vol. 646 (pp. 447-460). *Artificial Intelligence Applications and Innovations*. Springer, Cham. https://doi.org/10.1007/978-3-031-08333-4_36

Vitale, F., Chen, J., Odom, W., & McGrenere, J. (2020). Data Dashboard: Exploring Centralization and Customization in Personal Data Curation. *Proceedings of the 2020 ACM Designing Interactive Systems Conference*, 311-326. https://doi.org/10.1145/3357236.339547

Wang, R., & Sundar, S. S. (2020). How does Parallax Scrolling Influence User Experience? A Test of TIME (Thory of Interactive Media Effects). International *Journal of Human-Computer Interaction, 34*(6), 533-543. https://doi.org/10.1080/10447318.2017.1373457

van der Waa, J., Nieuwburg, E., Cremers, A., & Neerincx, M. (2021). Evaluating XAI: A comparison of rule-based and example-based explanations. *Artificial Intelligence, 291*, Article 103404. https://doi.org/10.1016/j.artint.2020.103404

Webster, J., & Watson, R. T. (2002). Analyzing the Past to Prepare for the Future: Writing a Literature Review. *MIS Quarterly, 26*(2), xiii-xxiii.

Wells, L., & Bednarz, T. (2021). Explainable AI and Reinforcement Learning – A Systematic Review of Current Approaches and Trends. *Frontiers in Artificial Intelligence, 4*, Article 550030. https://doi.org/10.3389/frai.2021.550030

Zacharias, J., von Zahn, M., Chen, J., & Hinz, O. (2022). Designing a feature selection method based on explainable artificial intelligence. *Electronic Markets*, *32*, 2159-2184. https://doi.org/10.1007/s12525-022-00608-1

Zhang, W., & Lim, B. Y. (2022). Towards Relatable Explainable AI with Perceptual Process. *Conference on Human Factors in Computer Systems*, Article 181. https://doi.org/10.1145/3491102.3501826

Zimmermann, R., Mora, D., Cirqueira, D., Helfert, M., Bezbradica, M., Werth, D., Weitzl, W. J., Riedl, R., & Auinger, A. (2022). Enhancing brick-and-mortar store shopping experience with an augmented reality shopping assistant application using personalized recommendations and explainable artificial intelligence. *Journal of Research in Interactive Marketing*, 1-26. https://doi.org/10.1108/JRIM-09-2021-0237

Appendix

Appendix A: Demographic Data for the Participants of the Experimental Study in Design Cycle 1

Characteristic/ Question	Baseline		Customization	
	N	%	N	%
Gender				
Female	28	31.1%	37	41.1%
Male	60	66.7%	52	57.8%
Other	2	2.2%	1	1.1%
Age				
18 – 29 years	34	37.8%	37	41.1%
30 – 39 years	23	25.6%	33	36.7%
40 – 49 years	25	27.7%	10	11.1%
50 – 59 years	6	6.7%	9	10.0%
> 60 years	2	2.2%	1	1.1%
Location				
Asia	0	0%	1	1.1%
Africa	18	20.0%	11	12.2%
Europe	57	63.4%	59	65.6%
North America	12	13.3%	16	17.8%
South America	3	3.3%	3	3.3%
Australia/ Oceania	0	0%	0	0%
Antarctica	0	0%	0	0%
Education				
High School	17	18.9%	22	24.5%
Bachelor's Degree	38	42.2%	45	50.0%
Master's Degree	29	32.3%	16	17.8%
Ph.D. or higher	3	3.3%	5	5.5%
Trade School	2	2.2%	1	1.1%
Other	1	1.1%	1	1.1%

Table A1. Overview of demographic characteristics ($N_{Baseline}$ = 90; $N_{Customization}$ = 90).

Appendix B: Demographic Data for the Participants of the Experimental Study in Design Cycle 2

Characteristic/ Question	N	Percentage
Gender		
Female	102	47.4%
Male	113	52.6%
Other	0	0.0%
Age		
18-29 years	101	47.0%
30-39 years	59	27.4%
40-49 years	44	20.5%
50-59 years	11	5.1%
> 60 years	0	0.0%
Location		
Asia	57	26.6%

Africa	88	40.9%
Europe	59	27.4%
North America	5	2.3%
South America	3	1.4%
Australia/ Oceania	3	1.4%
Antarctica	0	
Education		
High School	113	52.6%
Bachelor's Degree	89	41.4%
Master's Degree	13	6.0%
Ph.D. or higher	0	0.0%
Trade School	0	0.0%
Other	0	0.0%

Table B1. Overview of demographic characteristics (N = 215).

Appendix C: Screenshots for the XUI in Design Cycle 1

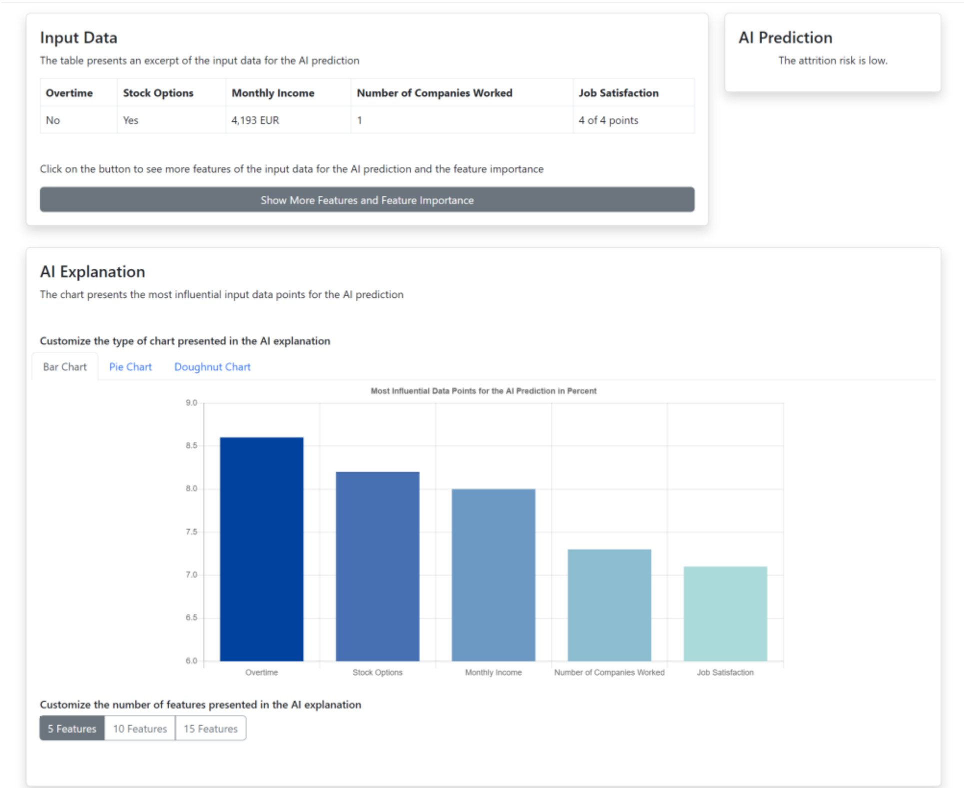


Figure C1. XUI with customization features with bar chart and five features.

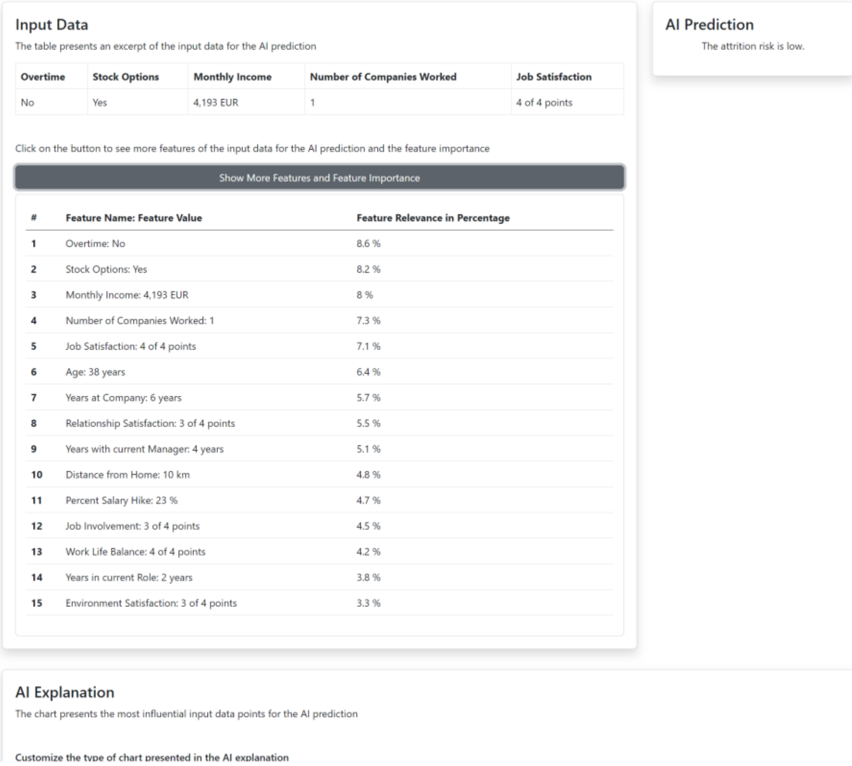


Figure C2. Overview of the feature relevance.

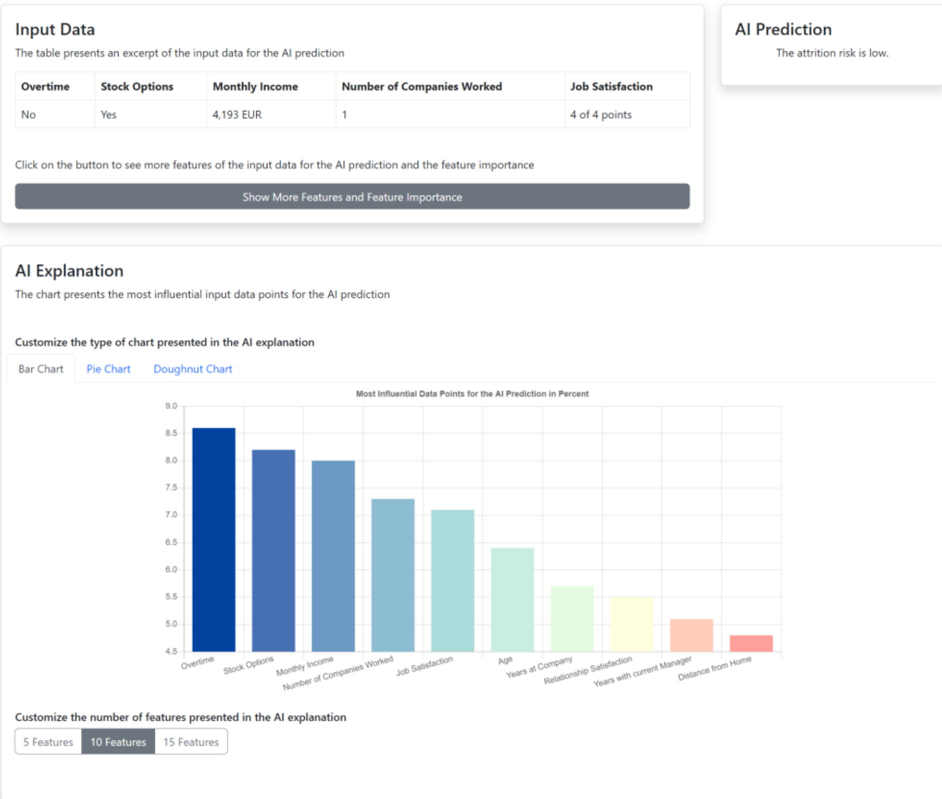


Figure C3. XUI with customization features with bar chart and 10 features.

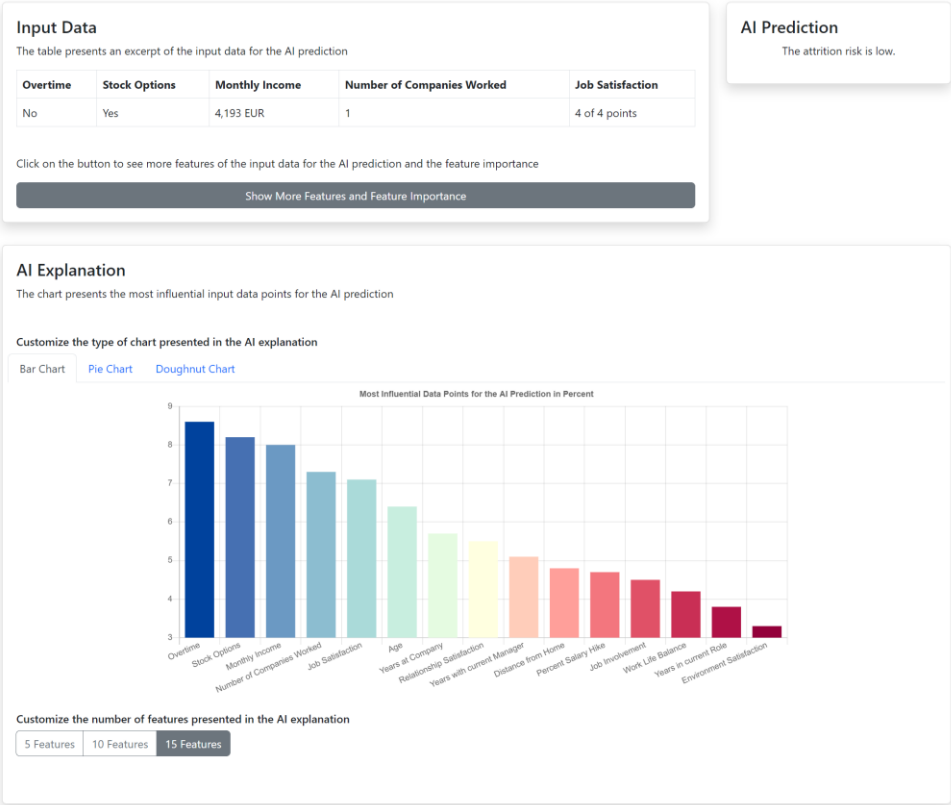


Figure C4. XUI with customization features with bar chart and 15 features.

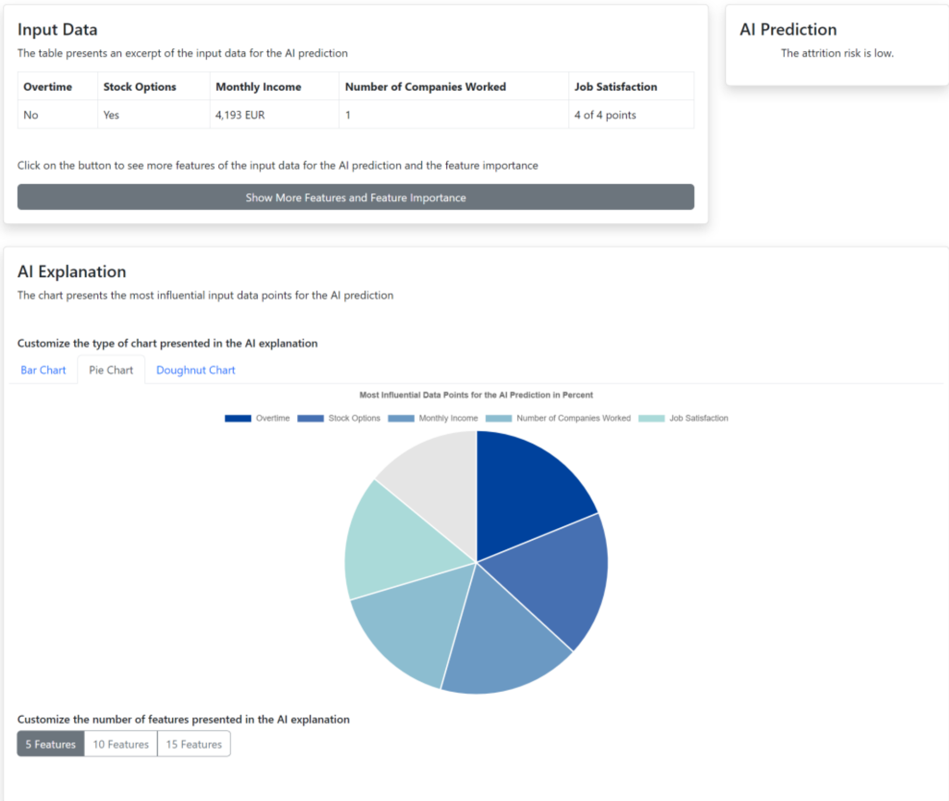


Figure C5. XUI with customization features with pie chart and five features.

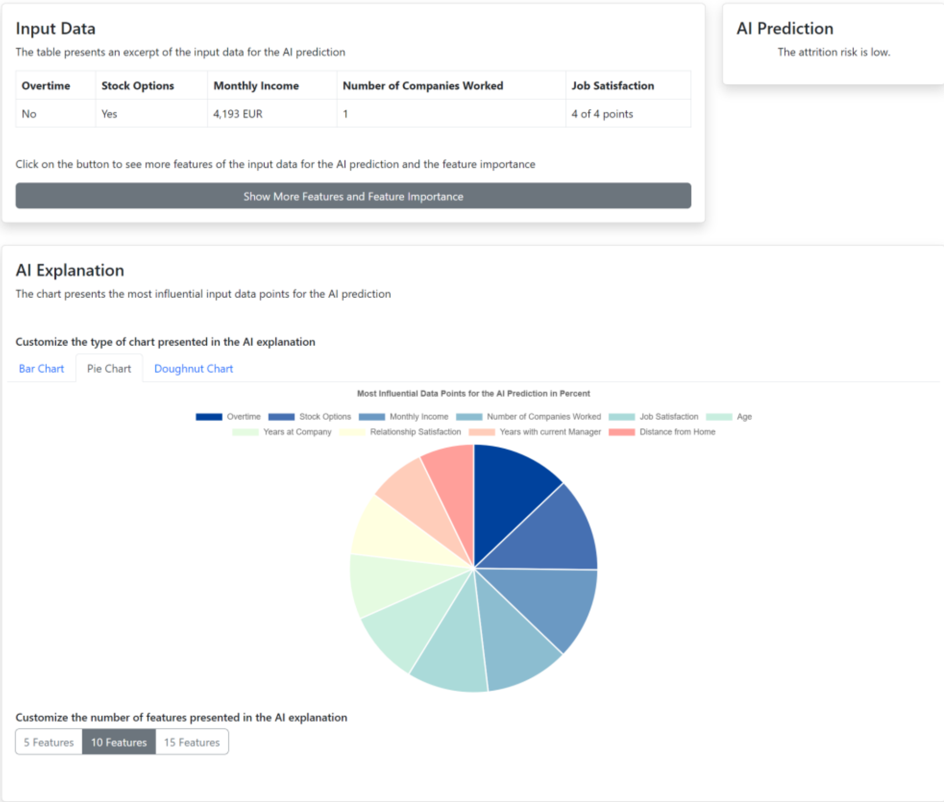


Figure C6. XUI with customization features with pie chart and 10 features.

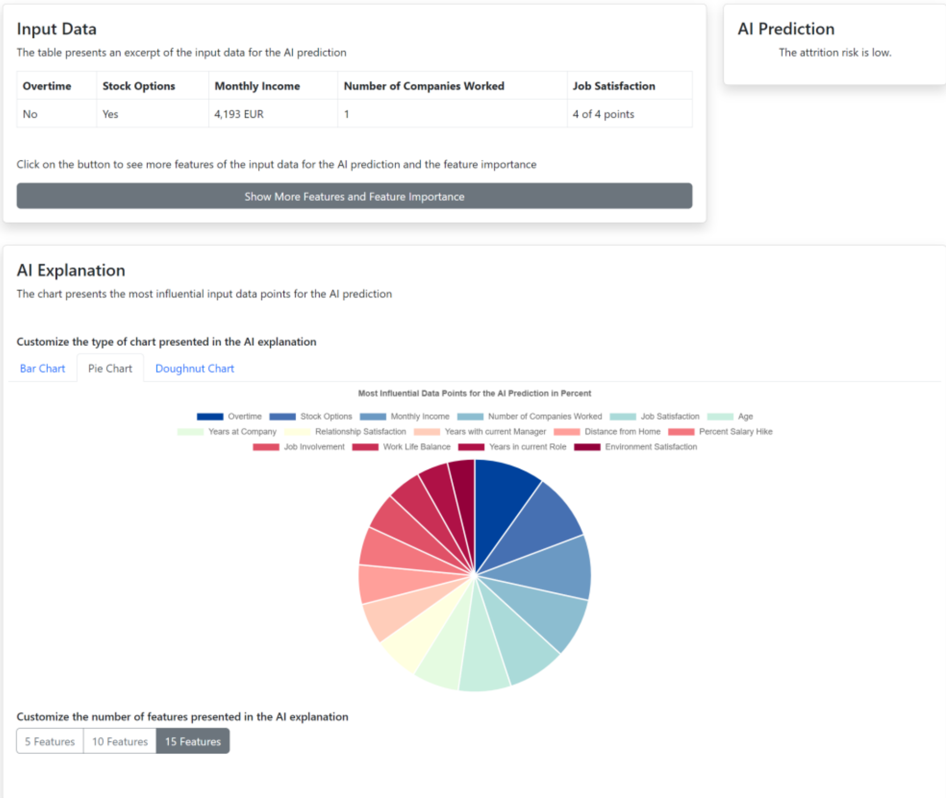


Figure C7. XUI with customization features with pie chart and 15 features.

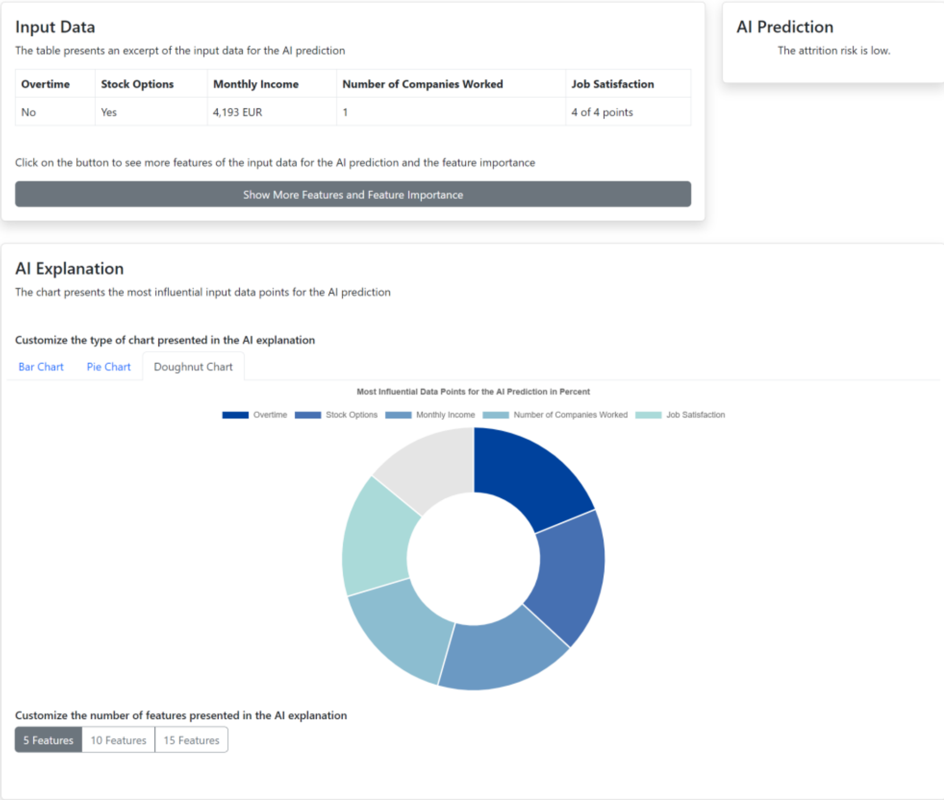


Figure C8. XUI with customization features with doughnut chart and five features.

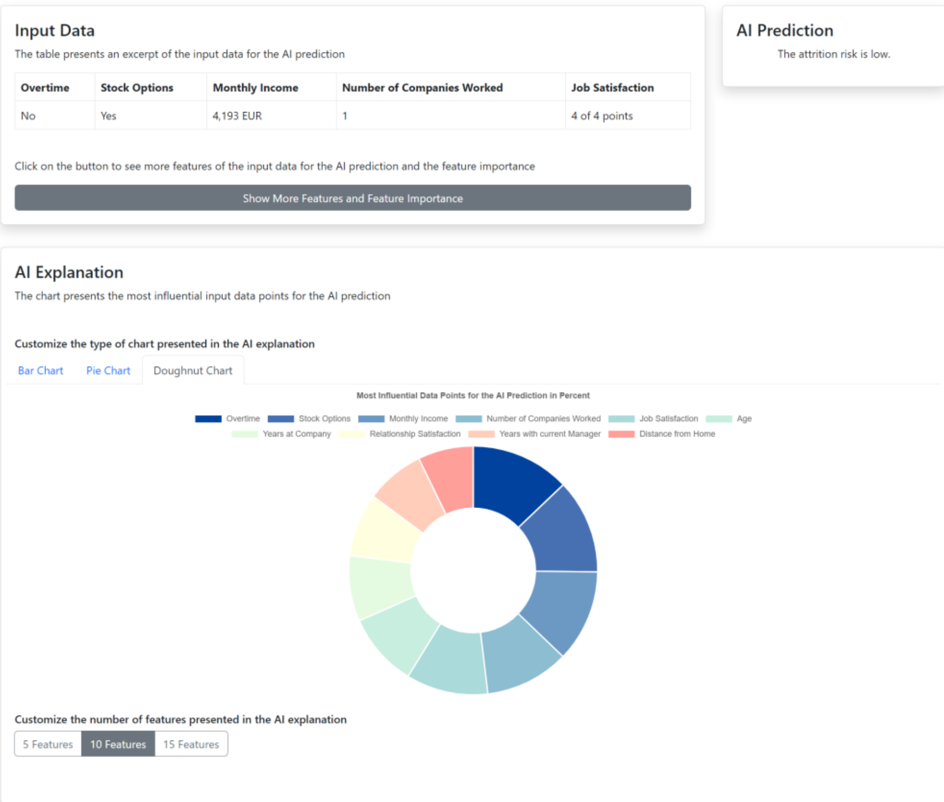


Figure C9. XUI with customization features with doughnut chart and 10 features.

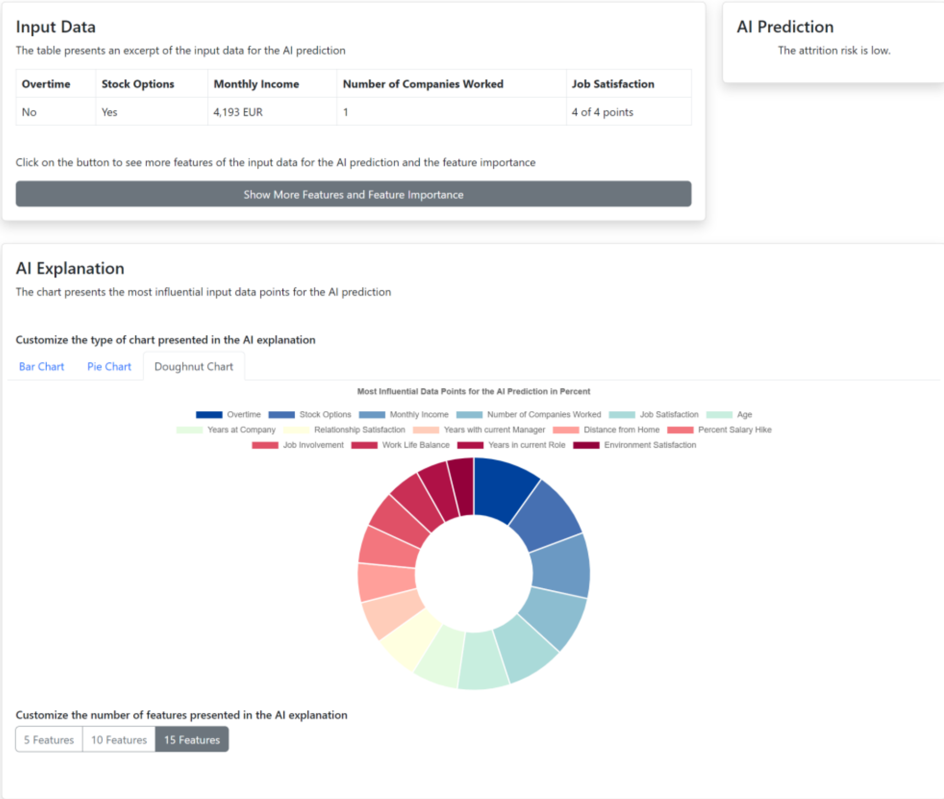


Figure C10. XUI with customization features with doughnut chart and five features.

Appendix D: Constructs, Items and Scales of the Post-Survey in Design Cycle 1

XConstructs and Items	Scales and Sources
Interestingness	Seven-point Likert scale ranging from Strongly disagree (1) to Strongly agree (7) (Shin et al., 2022)
The explanation is interesting	
The explanation is entertaining.	
The explanation is exciting.	
Informativeness	Seven-point Likert scale ranging from Strongly disagree (1) to Strongly agree (7) (Al-Natour et al., 2022)
The explanation seems knowledgeable.	
The explanation educates me.	
The explanation communicates a lot of information to me.	
The explanation is overall informative.	
Interactivity	Seven-point Likert scale ranging from Strongly disagree (1) to Strongly agree (7) (Sheng & Joginapelly, 2012)
The explanation is interactive.	
The explanation is engaging.	
The explanation is easy to navigate.	
It is easy to find my way through the explanation.	
The explanation provides immediate feedback.	
The explanation provides information I am looking for quickly.	
Explanation Satisfaction Scale	Seven-point Likert scale ranging from Strongly disagree (1) to Strongly agree (7) (Hoffman et al., 2018)
From the explanation, I understand how the artificial intelligence works.	
This explanation of how the artificial intelligence works is satisfying.	
This explanation of how the artificial intelligence works has sufficient detail.	

This explanation of how the artificial intelligence works seems complete.	
This explanation of how the artificial intelligence works tells me how to use it.	
This explanation of how the artificial intelligence works is useful to my goals.	
This explanation of the artificial intelligence shows me how accurate the system is.	
This explanation lets me judge when I should trust and not trust the artificial intelligence.	
Customization	Seven-point Likert scale ranging from Strongly disagree (1) to Strongly agree (7) (Theodosiou et al., 2019)
The explanation enables me to choose the right information for me.	
The explanation makes me feel that I am a unique user.	
The explanation considers my specific needs.	

Table D1. Constructs and items used in experimental study 1.

Appendix E: Screenshots for the XUI in Design Cycle 2

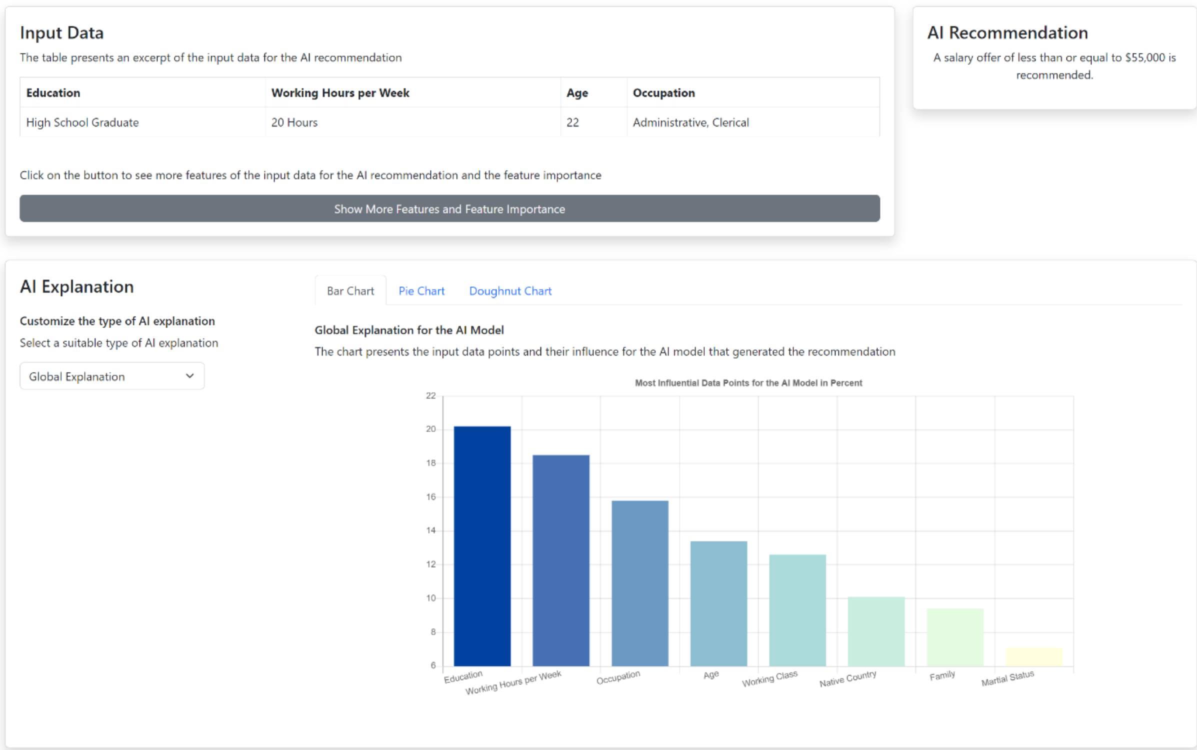


Figure E1. XUI with customization features with global explanation as bar chart.

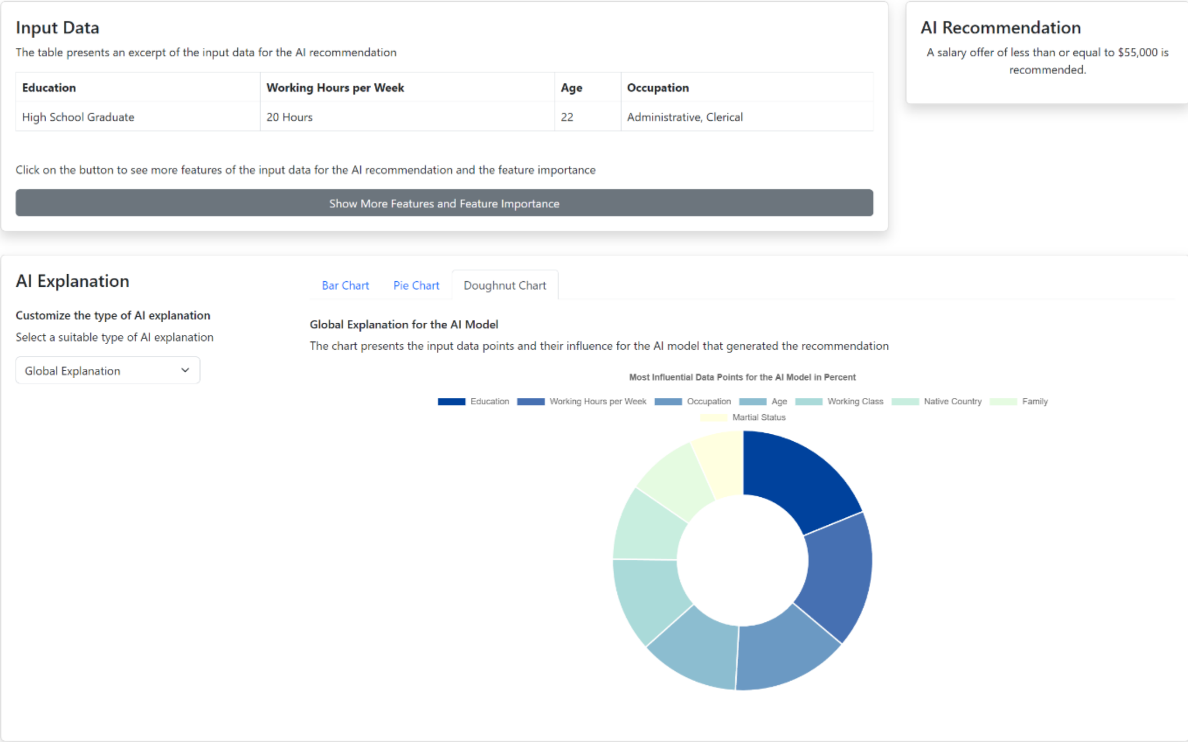


Figure E2. XUI with customization features with global explanation as doughnut chart.

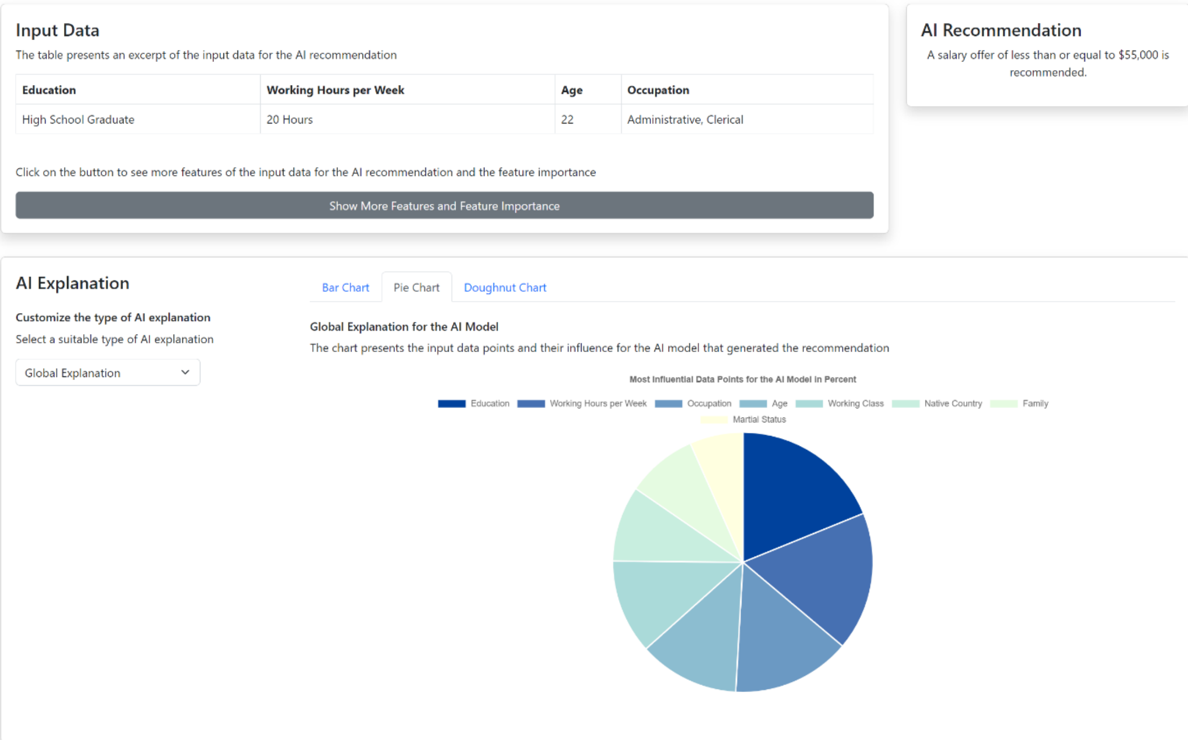


Figure E3. XUI with customization features with global explanation as pie chart.

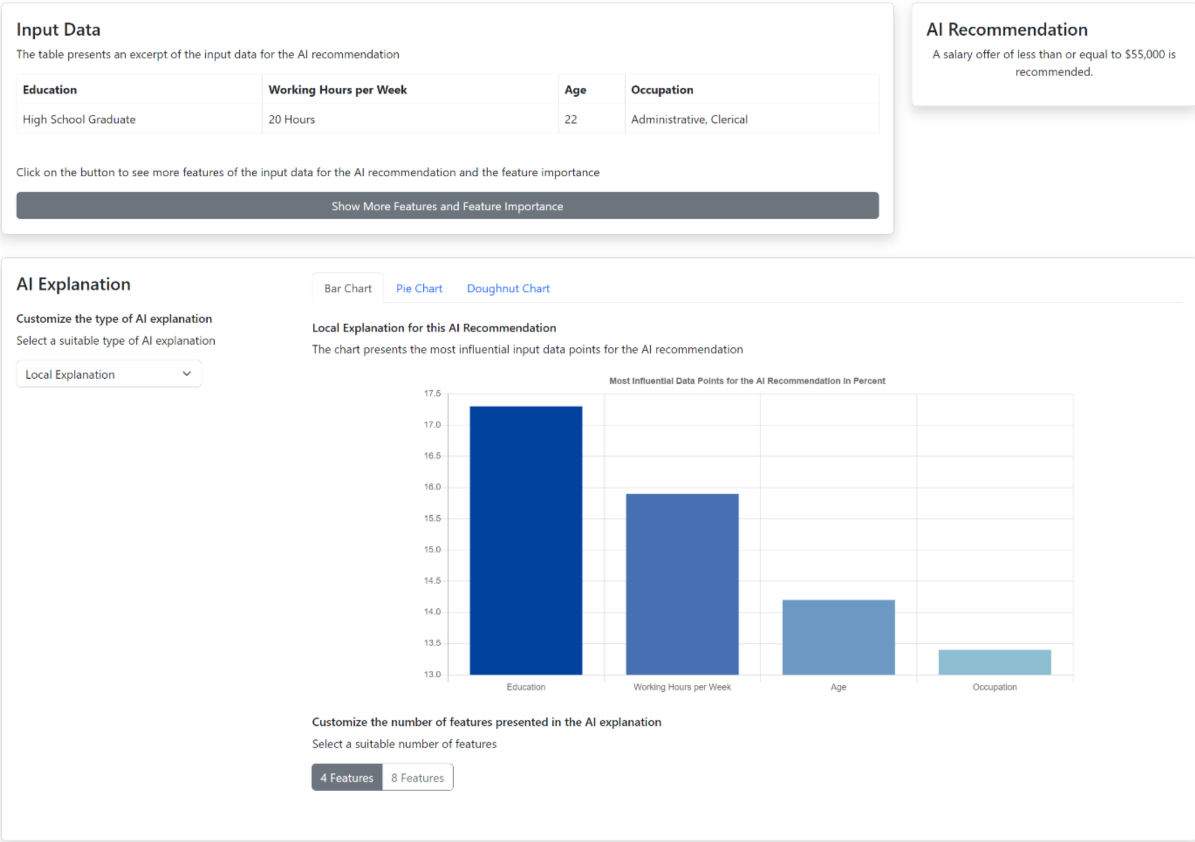


Figure E4. XUI with customization features with local explanation as bar chart and 4 features.

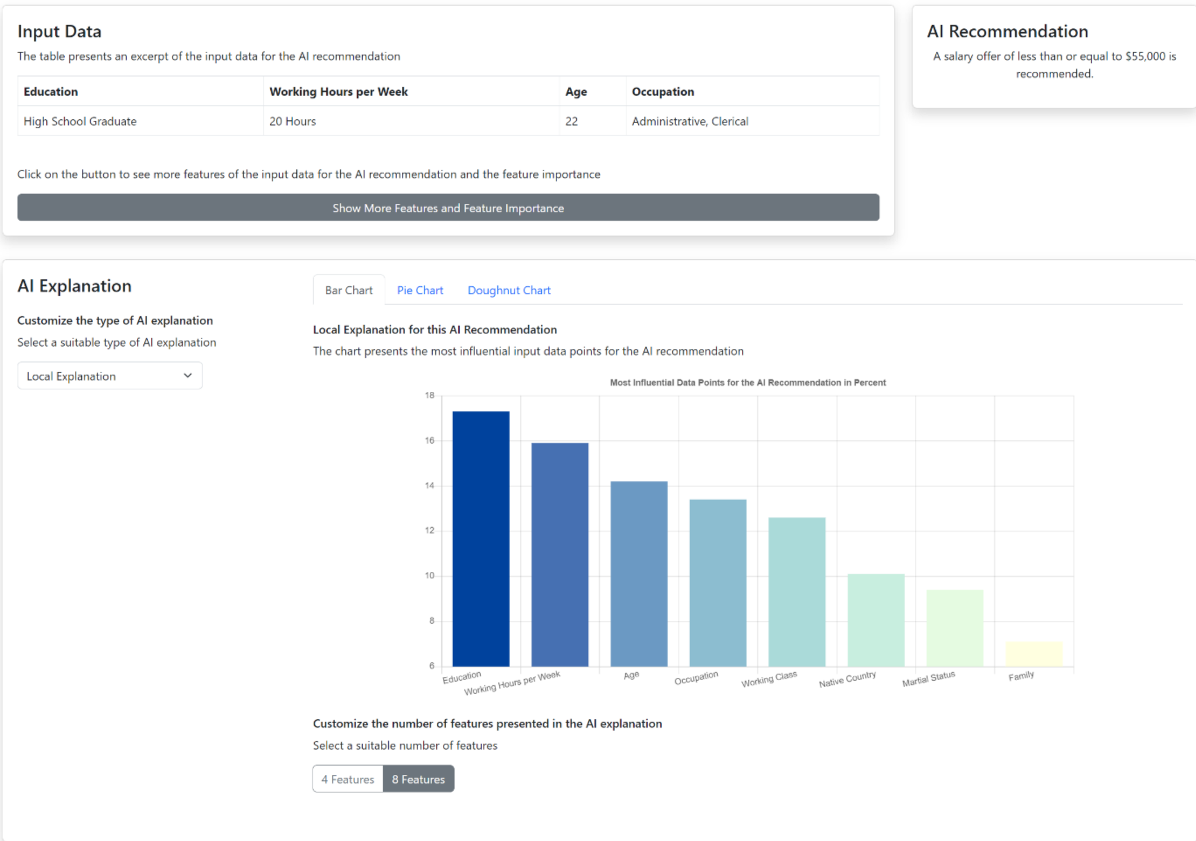


Figure E5. XUI with customization features with local explanation as bar chart and 8 features.

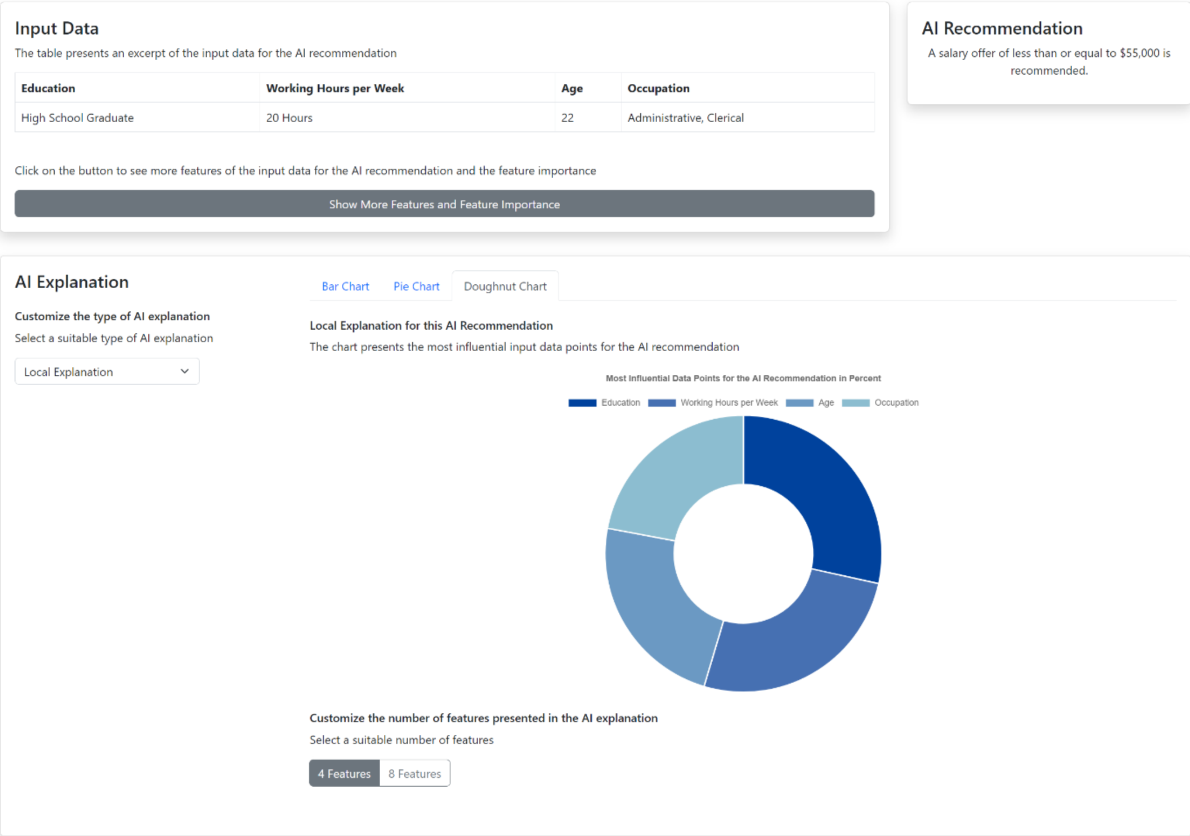


Figure E6. XUI with customization features with local explanation as doughnut chart and 4 features.

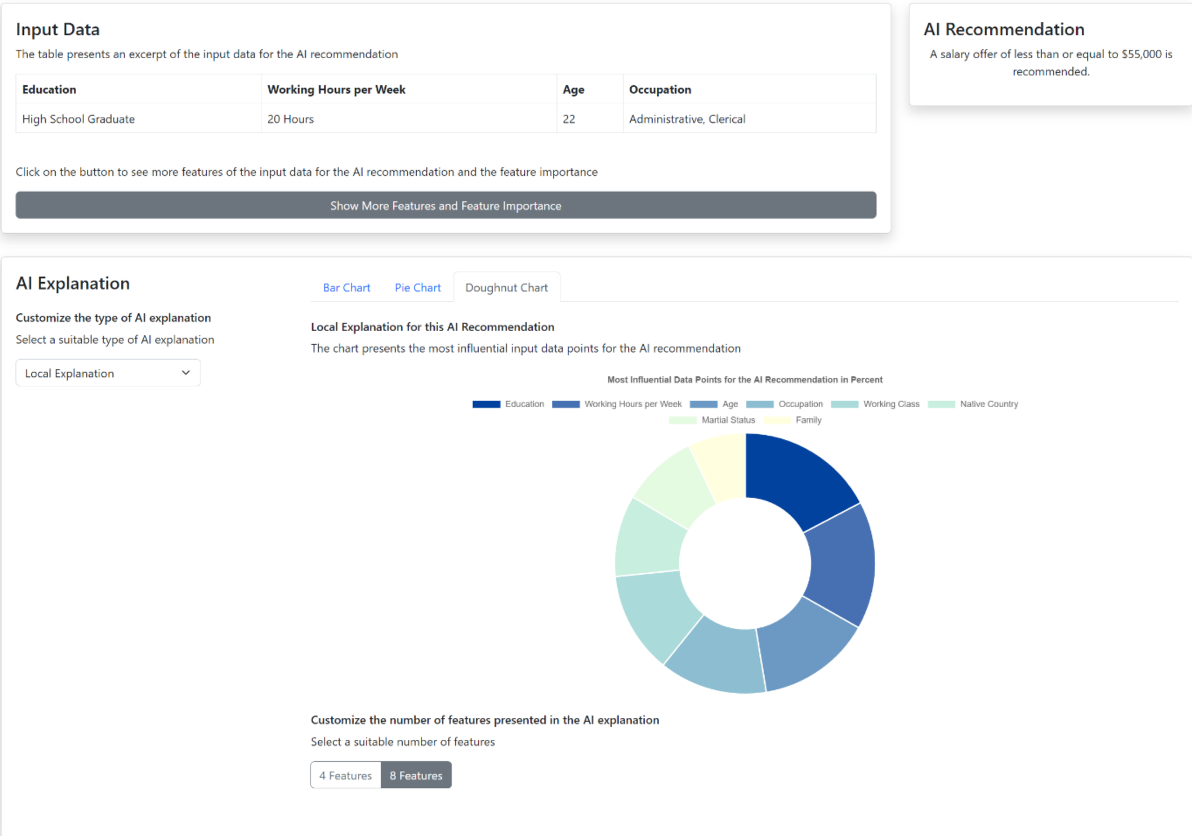


Figure E7. XUI with customization features with local explanation as doughnut chart and 8 features.

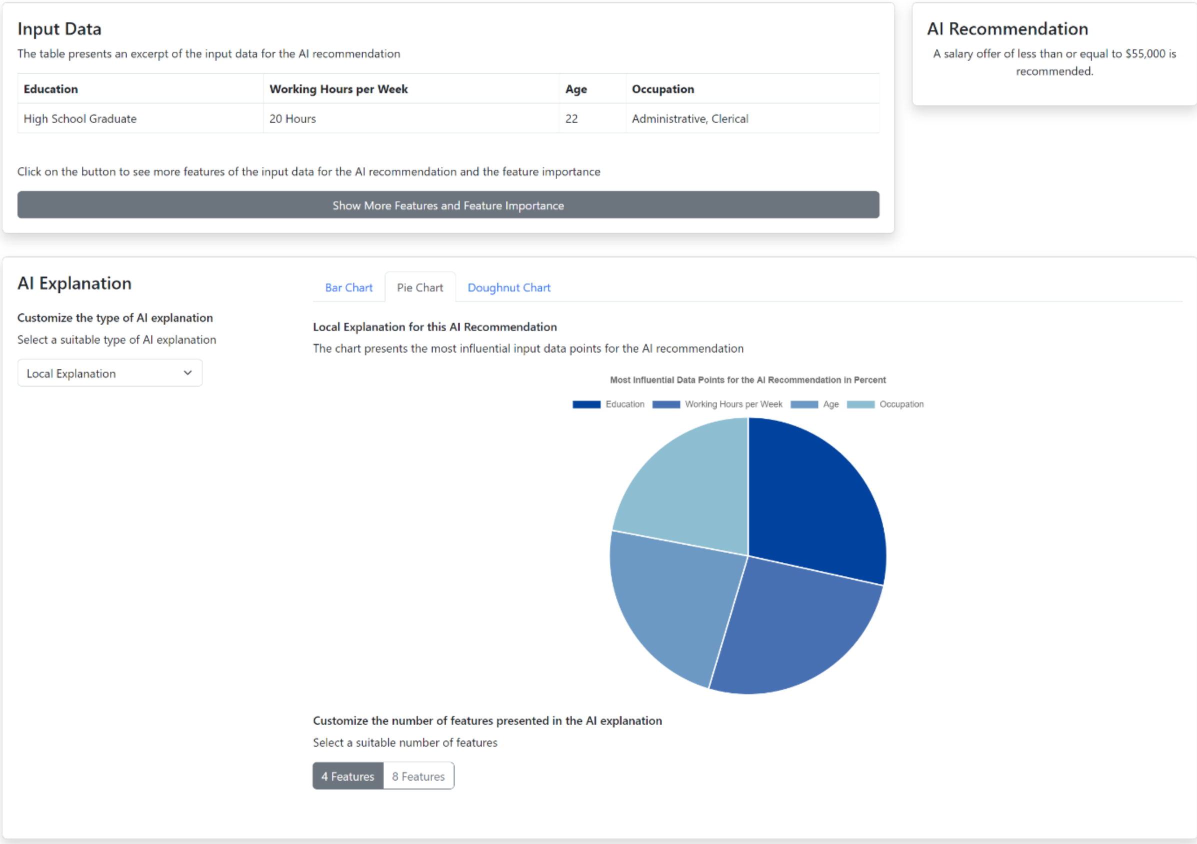


Figure E7. XUI with customization features with local explanation as pie chart and 4 features.

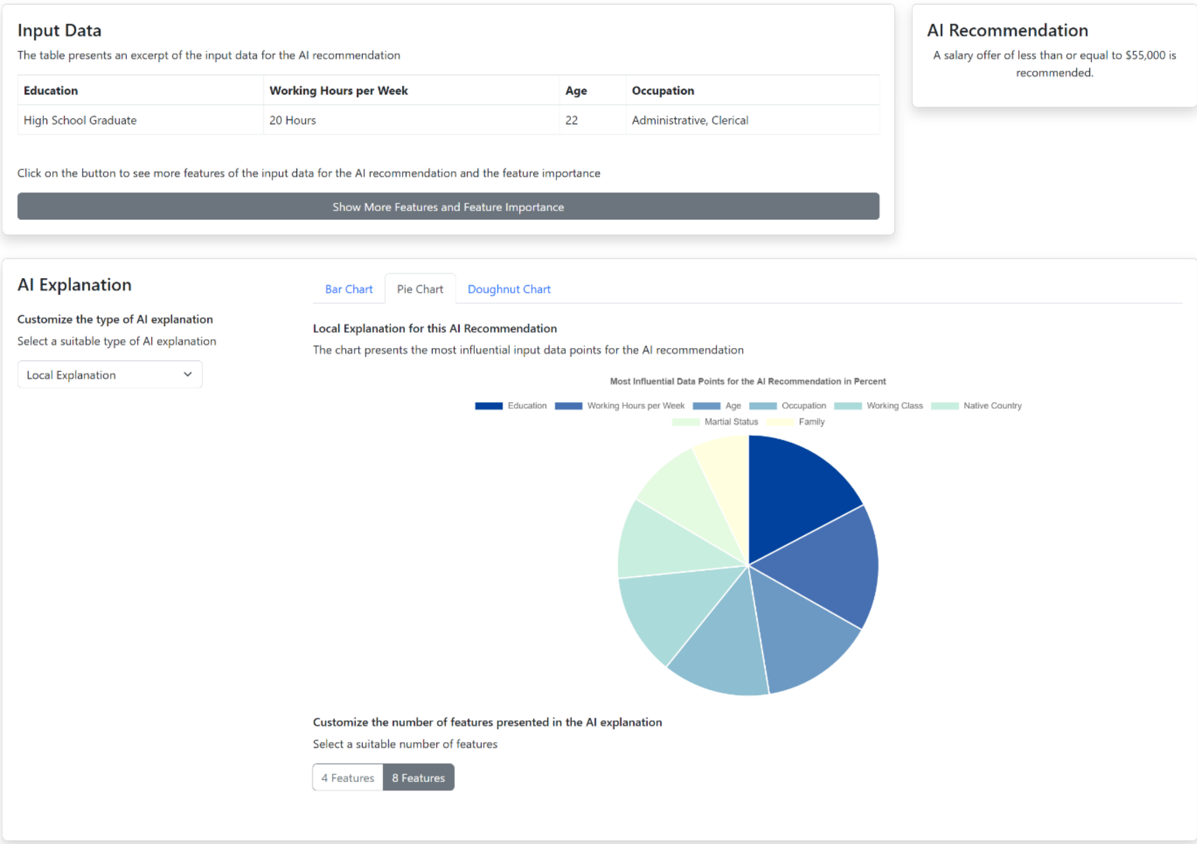


Figure E8. XUI with customization features with local explanation as pie chart and 8 features.

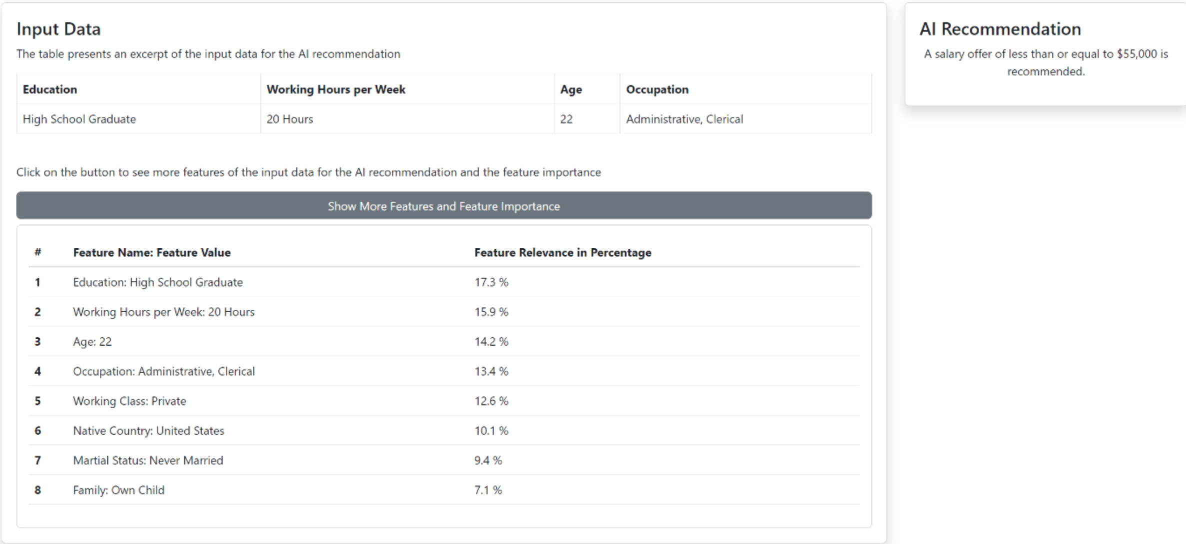


Figure E9. XUI with customization features – upper half of the XUI.

Appendix F: Constructs, Items and Scales of the Post-Survey in Design Cycle 2

Constructs and Items	Scales and Sources
Perceived Interactivity	
The explanation is interactive.	Seven-point Likert scale ranging from Strongly disagree (1) to Strongly agree (7) (Sheng & Joginapelly, 2012)
The explanation is engaging.	
The explanation is easy to navigate.	
It is easy to find my way through the explanation.	
The explanation provides immediate feedback.	
The explanation provides information I am looking for quickly.	
Perceived Interestingness	
The explanation is interesting	Seven-point Likert scale ranging from Strongly disagree (1) to Strongly agree (7)
The explanation is entertaining.	
The explanation is exciting.	
Perceived Customization	
The explanation enables me to choose the right information for me.	Seven-point Likert scale ranging from Strongly disagree (1) to Strongly agree (7) (Theodosiou et al., 2019)
The explanation makes me feel that I am a unique user.	
The explanation considers my specific needs.	
User Engagement	
Time appeared to go by very quickly when I was browsing the user interface.	Seven-point Likert scale ranging from Strongly disagree (1) to Strongly agree (7) (Wang & Sundar, 2017)
I lost track of time when I was browsing the user interface.	
While browsing the user interface, I was able to block out most other distractions.	
While browsing the user interface, I was absorbed in what I was doing.	
While browsing the user interface, I was immersed in the task that I was performing.	
I had fun interacting with the user interface.	
Satisfaction	
I enjoy using the explanation user interface.	

My choice to use such an explanation user interface would be a wise one.	Seven-point Likert scale ranging from Strongly disagree (1) to Strongly agree (7) (Li et al., 2021)
Interacting with the information in the explanation user interface is a pleasant experience.	
Overall, my feeling of the explanation user interface is satisfactory.	

Table F1. Constructs and items used in experimental study 2.

Appendix G: Demographic Data for the Participants of the Reusability Evaluation in Design Cycle 2

Characteristic/ Question	N	Percentage
Gender		
Female	20	32.8%
Male	41	67.2%
Other	0	0.0%
Age		
18-29 years old	42	68.9%
30-39 years old	13	21.3%
40-49 years old	3	4.9%
50-59 years old	2	3.3%
Over 60 years old	1	1.6%
Location		
Asia	15	24.6%
Africa	30	49.2%
Europe	10	16.4%
North America	1	1.6%
South America	2	3.3%
Australia/ Oceania	3	4.9%
Antarctica	0	0.0%
Education		
High School	39	63.9%
Bachelor’s Degree	16	26.3%
Master’s Degree	5	8.2%
Ph.D. or higher	1	1.6%
Trade School	0	0.0%
Other	0	0.0%
Experience with User Interface Design		
Web Apps	38	62.3%
Mobile Apps	6	9.8%
Cross-Platform Apps	13	21.3%
Desktop Apps	4	6.6%
Would you use the design principles for a suitable project?		
Yes	52	85.2%
No	9	14.8%
Would you recommend the design principles to colleagues?		
Yes	53	86.9%
No	8	13.1%

Table G1. Overview of demographic characteristics (N = 61).

Appendix H: Constructs, Items and Scales of the Reusability Evaluation in Design Cycle 2

Constructs and Items
Accessibility
The design principles are easy for me to understand
The design principles are easy for me to comprehend
The design principles are intelligible to me
Importance
In my view design principles for customization features in explanation user interfaces address a real problem in my professional practice
In my view design principles for customization features in explanation user interfaces address an important – acute or foreseeable – problem in my professional practice
Novelty and Insightfulness
I find that the design principles convey new ideas to me
I find the design principles insightful to my own practice
Actability and Appropriate Guidance
I think that the design principles can realistically be carried out in practice
I think that the design principles can easily be carried out in practice
I find that the design principles provide sufficient guidance for designing customization features in explanation user interfaces
I find that the design principles provide sufficient direction for designing customization features in explanation user interfaces
I find that the design principles are not restrictive when designing customization features in explanation user interfaces
I find that the design principles provide me with sufficient design freedom when designing customization features in explanation user interfaces
Effectiveness
I believe that the design principles can help design customization features in explanation user interfaces in practice
I find the design principles useful for designing customization features in explanation user interfaces in practice

Table H1. Constructs and items used in the reusability evaluation of the design principles.