# The Role of Process and Outcome Accountability Claims for Shaping AI Developers' Perceived Accountability

Jan-Hendrik Schmidt
*Technical University of Darmstadt*, schmidt@ise.tu-darmstadt.de

Sebastian Clemens Bartsch
*Technical University of Darmstadt*, bartsch@ise.tu-darmstadt.de

Martin Adam
*Technical University of Darmstadt*, adam@ise.tu-darmstadt.de

# The Role of Process and Outcome Accountability Claims for Shaping AI Developers' Perceived Accountability

*Completed Research Paper*

**Jan-Hendrik Schmidt**
Technical University of Darmstadt
Hochschulstr. 1,
64289 Darmstadt, Germany
schmidt@ise.tu-darmstadt.de

**Sebastian Clemens Bartsch**
Technical University of Darmstadt
Hochschulstraße 1,
64289 Darmstadt, Germany
bartsch@ise.tu-darmstadt.de

**Martin Adam**
Technical University of Darmstadt
Hochschulstraße 1,
64289 Darmstadt, Germany
adam@ise.tu-darmstadt.de

## Abstract

*As accountability becomes increasingly important for developers of artificial intelligence (AI)-based systems, governance mechanisms such as AI principles or audits are often criticized for not sufficiently influencing AI developers. Therefore, we examine how visualized arguments in user interfaces (UIs) of integrated development environments (IDEs) can increase AI developers' perceived accountability. Combining construal level theory and Toulmin's model of argumentation, four UI design artifacts were developed, each containing a claim of process or outcome accountability with or without monitoring and evaluation tools that act as claim-supporting data. Results of an online experiment with 164 AI developers show that claiming process accountability increases AI developers' perceived accountability more than claiming outcome accountability, both without supporting data. However, when supporting data are available, both claims increase AI developers' perceived accountability comparably effectively. The study's results highlight the theoretical and practical usefulness of visualized arguments in UIs of IDEs to promote AI developers' accountability.*

**Keywords:** Artificial Intelligence, AI-based Systems, Perceived Accountability, Process Accountability, Outcome Accountability, Toulmin's Model of Argumentation, Construal Level Theory

## Introduction

Accountability is often considered a central component in the governance of artificial intelligence (AI)-based systems to clarify who should be responsible and liable for such systems, their behavior, and their outcomes (Wieringa, 2020). Often, accountability is assigned to AI developers, who in turn frequently struggle with accountability gaps between the "designer's control and algorithm's behavior" (Mittelstadt et al., 2016, p. 11), creating an additional burden in the context of developing AI-based systems (e.g., Benlian, 2022). Accountability gaps are a growing concern for organizations providing AI-based systems, as they cannot ensure control over their AI-based systems nor predict their behavior (Benbya et al., 2021; Raji et al., 2020). Accordingly, more and more organizations address accountability gaps by enforcing governance mechanisms such as AI principles and audits to hold their AI developers accountable (Jobin et al., 2019).

Despite their recent popularity, AI principles and audits are often considered ineffective in influencing AI developers in their daily work (Raji et al., 2020). AI principles can be understood as organizational guidelines for AI developers when developing AI-based systems (Mittelstadt, 2019), for example, by providing design recommendations or specifying data quality requirements. However, they are often accused of being too abstract and vague to serve as evaluation measures for AI developers' actions and decisions (Mittelstadt, 2019; Raji et al., 2020). AI audits, in turn, aim at collecting and analyzing data on the outcomes of AI-based systems (e.g., recommendations or decisions) to assess whether they comply with organizational policies or legal and industry requirements to protect their users' interests and rights (Brown et al., 2021; Raji et al., 2020). However, AI audits still lack legally prescribed and standardized requirements for their components or implementations, resulting in various industry- and company-specific solutions (Raji et al., 2020). Moreover, they often target the technical implementation of AI principles, thereby overlooking the application contexts of AI-based systems and other challenges AI developers face (Brown et al., 2021). As a result, conducting AI audits remains abstract and their results intangible, making them less suitable for holding AI developers accountable for the outcomes of AI-based systems.

Given these drawbacks of existing governance mechanisms of AI-based systems, it is worth investigating alternative ways to raise awareness among AI developers about their accountability for AI-based systems. Prior information systems (IS) research has investigated the effectiveness of user interface (UI) design artifacts in increasing IS users' perceived accountability (e.g., Adam, 2022; Vance et al., 2015). However, these studies have not examined how UI design artifacts influence AI developers and how arguments they visualize should be constructed to inform AI developers that they will be held accountable. In particular, the question arises whether the use of UI design artifacts, for example, in integrated development environments (IDEs) stating the accountability of AI developers (i.e., *claim*) is sufficient to raise their awareness of their accountability or whether additional tools need to be integrated into UI design artifacts that provide evidence of AI developers' accountability (i.e., *claim-supporting data*). Further, previous IS research has understood accountability for AI-based systems largely as a unitary, indivisible construct, ignoring temporal or activity-based distinctions (e.g., Benbya et al., 2021; Martin, 2019). However, such an approach is not sufficient to address accountability gaps in AI-based systems, especially given the amount of time that elapses between AI-based systems' development processes and the appearance of their outcomes, both of which require accountability (e.g., Ågerfalk et al., 2021; Mikalef et al., 2022). Accordingly, it is important to examine how AI developers are affected when they are held accountable for their actions during development processes of AI-based systems (*process accountability*), where the outcomes tend to remain irrelevant, or when they are held accountable for the outcomes of their developed AI-based systems (*outcome accountability*), regardless of the development processes that led to those outcomes (Lerner & Tetlock, 1999). Against this background, this study investigates how AI developers perceive their accountability in response to UI design artifacts in IDEs that visualize arguments of their process versus outcome accountability for AI-based systems (i.e., claims). At the same time, this study examines whether presented data in corresponding UI design artifacts in the form of evaluation and monitoring tools of AI-based systems' development processes or outcomes support the effects of such claims of process versus outcome accountability on AI developers' perceived accountability (i.e., claim-supporting data). These two research objectives lead to our research question:

*RQ: How do visualized arguments (i.e., claims) for process versus outcome accountability affect AI developers' perceived accountability? And how is this effect influenced by the absence versus presence of tools supporting the argument (i.e., claim-supporting data)?*

To answer this research question, we combine construal level theory (CLT) (Trope & Liberman, 2010) with Toulmin's model of argumentation (1958). We conducted a 2 (claim: process accountability vs. outcome accountability) × 2 (claim-supporting data: present vs. absent) online experiment with between subject-design with 164 AI developers. The results show that our UI design artifacts implemented in an IDE increase AI developers' perceived accountability in both cases of visualized arguments for process and outcome accountability. However, UI design artifacts that contain only a process accountability claim increase AI developers' perceived accountability more than UI design artifacts that contain only an outcome accountability claim. Yet, when claim-supporting data are present, both types of accountability claims are comparably effective in increasing AI developers' perceived accountability.

Our study makes three important contributions to IS research on accountability for AI-based systems (e.g., Benbya et al., 2021; Mikalef et al., 2022), especially concerning AI developers' accountability (e.g., Martin, 2019; van den Broek et al., 2021). First, we go beyond previous IS research on accountability, which focused significantly on users' accountability perceptions (e.g., Adam, 2022; Vance et al., 2013, 2015), and examine the accountability perceptions of AI developers as essential stakeholders in developing, operating and maintaining AI-based systems. Second, we introduce the conceptual distinction between process and outcome accountability for AI-based systems, which provides a more nuanced view compared to the conceptualization of accountability in previous IS research (e.g., Adam, 2022; Martin, 2019; Schmidt et al., 2023). This is particularly crucial in light of the different goals of accountability as a means of governance for development processes and outcomes of AI-based systems (e.g., Mikalef et al., 2022). Third, we shed light on the relevance of the structure and components of visualized accountability arguments. We show that accountability is perceived particularly effectively when AI developers are not only made aware of their accountability but are also presented with evidence for their accountability. From a practical perspective, our study shows that UI design artifacts incorporated in IDEs provide organizations with straightforward and easy-to-understand solutions to increase AI developers' perceived accountability for both the development processes and emerging outcomes of AI-based systems.

## Theoretical Background and Related Literature

### *Accountability of AI Developers: Process and Outcome Accountability*

Accountability is often described as a governance mechanism in organizational relationships that involves an individual's obligation to explain and justify his or her actions to an audience upon request. The audience has the right to evaluate the explanations and justifications and impose positive or negative consequences depending on its judgment (Bovens, 2010; Lerner & Tetlock, 1999; Vance et al., 2015). Due to the increasing autonomy, learnability, and inscrutability of AI-based systems (Berente et al., 2021), accountability for such systems has gained a lot of traction recently in legislative and regulatory initiatives (e.g., European Commission, 2021) as well as IS research (e.g., Schmidt et al., 2023; Wieringa, 2020). More specifically, IS research primarily explores who should be held accountable for AI-based systems, including their development, behavior, and outcomes (e.g., recommendations or decisions), and consequently be able to explain and justify them (Wieringa, 2020). In this regard, an emphasis of IS research is also on how accountability for AI-based systems' development processes and outcomes can be ensured, for which governance mechanisms such as AI principles and audits are being explored (e.g., Jobin et al., 2019; Mittelstadt, 2019; Raji et al., 2020).

Previous IS research also explored how accountability perceptions of users can be increased, for which UI design artifacts have been shown to be appropriate (e.g., Adam, 2022; Vance et al., 2015). The UI design artifacts employed in these studies utilize accountability mechanisms identified by social psychology, such as *expectation of evaluation* and *awareness of monitoring* (Griffith, 1993; Lerner & Tetlock, 1999). While expectation of evaluation refers to the situations of justification and evaluation of one's actions within accountability relationships (Lerner & Tetlock, 1999), awareness of monitoring describes individuals' perceptions that their actions are being monitored for this purpose (Griffith, 1993). However, users of AI-based systems are not the only stakeholders involved in the development, operation, and usage of AI-based systems and whose accountability perceptions are therefore relevant. IS research emphasizes not least the accountability of AI developers, whose actions and decisions during the development and design of AI-based systems have substantial effects on the behavior and outcomes of those very systems (Martin, 2019; Mikalef et al., 2022; Wieringa, 2020). Accordingly, accountability for AI-based systems can only be thoroughly explored if AI developers' perceptions of accountability are also considered.

In examining AI developers' perceptions of accountability, accountability for AI-based systems should not be seen as a unitary, indivisible construct, as previous IS research has predominantly done: Although IS research has conceptually addressed various lifecycle stages of AI-based systems, such as their development processes and the emergence of their outcomes during use (e.g., Martin, 2019; Mikalef et al., 2022; Wieringa, 2020), there is a lack of empirical studies that take such differentiation into account. Previous IS research on users' accountability perceptions also has not distinguished whether the actions for which accountability is assumed occurred during decision-making processes or were the cause of a particular outcome (e.g., Adam, 2022; Vance et al., 2013, 2015). However, previous social psychological research on accountability has indicated that people respond differently to being held accountable for processes or outcomes (Lerner & Tetlock, 1999): When a person is held accountable for processes, it increases his or her time and effort to compare potential courses of action before making a decision (Doney & Armstrong, 1995). In contrast, when a person is held accountable for outcomes, it leads to a stronger commitment to a prior course of action even after a particular decision has been made (Simonson & Staw, 1992).

Accordingly, IS research on accountability for AI-based systems should also differentiate between process accountability and outcome accountability, not least because the development processes of AI-based systems and the emergence of their outcomes represent distinct lifecycle elements that may be far apart in time (e.g., van den Broek et al., 2021). This is also particularly important because AI developers may perceive the importance of accountability in more or less concrete terms depending on whether they are faced with process or outcome accountability, and therefore need to explain and justify the development processes of their AI-based systems taking place at present or explain and justify the outcomes of those systems when they have emerged after the systems have been in operation for some time.

## *Construal Level Theory and AI Developers' Perceived Accountability*

CLT refers to the cognitive processing of a goal (e.g., justifying one's actions) and related messages (e.g., information about how one must justify his or her actions) by depicting that a more concrete representation of an event or object results from a closer psychological distance from it (Trope & Liberman, 2010). Within CLT, *psychological distance* describes an individual's "subjective experience that something is close or far away" (Trope & Liberman, 2010, p. 440) concerning, for example, time (i.e., temporal distance) and likelihood (i.e., hypothetical distance). For instance, AI developers may think about an upcoming meeting in which they must justify their actions more concretely (e.g., which people will attend the meeting), while in general they may perceive accountability for AI-based systems more abstractly (e.g., why taking accountability for AI-based systems is important in the first place). Accordingly, the form of accountability assigned to AI developers (i.e., accountability for the development processes or accountability for the outcomes of AI-based systems) elicits different temporal distances according to CLT. AI developers perceive currently ongoing development processes of AI-based systems with less temporal distances than outcomes of AI-based systems that might occur sometime in the future. For example, as part of development processes, AI developers may decide which algorithms or training datasets to use or which capabilities to grant AI-based systems (e.g., Zhang et al., 2020). For the outcomes of AI-based systems, however, AI developers can usually only monitor their systems during operation for the occurrence of undesirable results such as discriminatory behavior or model drift, or try to take precautions through documented specifications of the intended use case (e.g., Benbya et al., 2021).

Previous IS research has investigated the effects of temporal distances on IS users' decision or choice behavior by altering their levels of abstraction and, thus, their construal levels (e.g., Adam et al., 2020; Wendt et al., 2022). For example, information associated with lower temporal distance (e.g., by being highly topical or representing real-time data) compared to information associated with high temporal distance has been shown to elicit higher IS reuse intentions among users and lead to more vivid perceptions, positively affecting users' choice behavior (Adam et al., 2020; Wendt et al., 2022). Accordingly, it can be assumed that process accountability signaling UI design artifacts for IDEs, which represent lower temporal distance given currently ongoing development processes, elicit more concrete mental construals compared to outcome accountability signaling UI design artifacts. However, as accountability for AI-based systems often remains vague or implicit (Martin, 2019; Mikalef et al., 2022), the question arises of which arguments are appropriate for conveying to AI developers that they are accountable for the development processes or outcomes of AI-based systems.

## *Toulmin's Model of Argumentation*

Toulmin (1958) established a model of argumentation according to which arguments of daily communication mainly consist of the three central components *claim*, *data*, and *warrant*, that are needed to compose a "skeleton of a pattern for analyzing arguments" (Toulmin, 2003, p. 92). While a claim describes an assertion or conclusion presented for general acceptance (Ye & Johnson, 1995), data represent the evidence used to support the claim (VerLinden, 1998). Warrants are propositions linking data and claims, often taking on an incidental and explanatory function (Toulmin, 1958). They aim to explicitly establish the legitimacy of the step from data to claim against the background of larger contexts whose legitimacy is presupposed (Toulmin, 2003). Figure 1 shows an example argument with the relationships between these three components. However, warrants are often only implicitly referred to without being explicitly expressed like data (Toulmin, 2003). Therefore, our study also does not explicitly present warrants in its UI design artifacts and focuses exclusively on the two components claim and claim-supporting data. Accordingly, our visualized arguments consist of only a claim alone or a combination of a claim and claim-supporting data. Other combinations, such as claim-supporting data without a claim, could also be investigated but rarely occur in everyday communication (Kim & Benbasat, 2006), which is why we excluded them from our study.

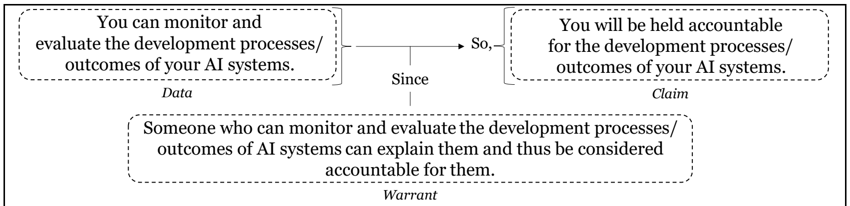



**Figure 1. Structure of Communication Arguments**

Previous IS research has shown that arguments have a more substantial effect when they consist of a claim and claim-supporting data. More specifically, the combination of claim and claim-supporting data has been shown to increase customers' trust in an online store, which has not been achieved with claims alone (e.g., Kim & Benbasat, 2006; Schneider et al., 2017). Furthermore, the persuasiveness of arguments has also been shown to depend on their contextual orientation, for example, when different motivational orientations were addressed (e.g., Schneider et al., 2017). In the context of AI-based systems' governance, the claim is often made that AI developers should be held accountable for these systems (e.g., Benbya et al., 2021; Martin, 2019). However, it is also emphasized that AI-based systems are moving away from their stakeholders, such as AI developers, as their actions and outcomes become more opaque and more automated (Berente et al., 2021), thereby impairing accountability. Therefore, claims of AI developers' accountability for AI-based systems' development processes or outcomes can only be persuasive if AI developers can monitor and evaluate their development processes or outcomes. The presence of tools in IDEs that AI developers can use to monitor and evaluate the development processes and outcomes of AI-based systems could serve as data that supports corresponding claims of AI developers' process or outcome accountability. Such tools could, for example, enable AI developers to document key assumptions and decisions during development processes or to review the model performance of their AI-based systems in operation. This ultimately gives AI developers the ability to monitor and evaluate the development processes or outcomes of their AI-based systems for subsequent explanation and justification.

Accordingly, different psychological distances according to CLT and the structure of the respective arguments following Toulmin's model of argumentation must be considered if AI developers are to be persuaded of their process or outcome accountability for AI-based systems. Figure 2 shows the conceptual framework we relied on during the development of the UI design artifacts underlying our study.

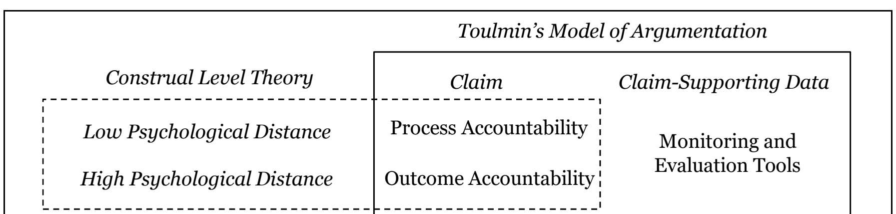



**Figure 2. Conceptual Framework of Accountability Arguments**

## Research Model and Hypotheses Development

Our research model explores how accountability arguments for development processes of AI-based systems versus outcomes of AI-based systems should be visualized as UI design artifacts for IDEs to increase AI developers' perceived accountability. Accordingly, we first address the effects of UI design artifacts, each consisting of a claim for process versus outcome accountability, on AI developers' perceived accountability, each of which is contrasted with a hanging control group devoid of UI design artifacts (H1A/ H1B). We then examine the extent to which the presence versus absence of claim-supporting data in claim-containing UI design artifacts influences their effects on AI developers' perceived accountability (H2A/H2B). Our third set of hypotheses explicitly compares the effects of UI design artifacts, each consisting of a claim for process versus outcome accountability (H3A). It further explores how claim-supporting data's presence versus absence alters the effects of such process versus outcome accountability claims on AI developers' perceived accountability (H3B). Figure 3 shows our corresponding research model.

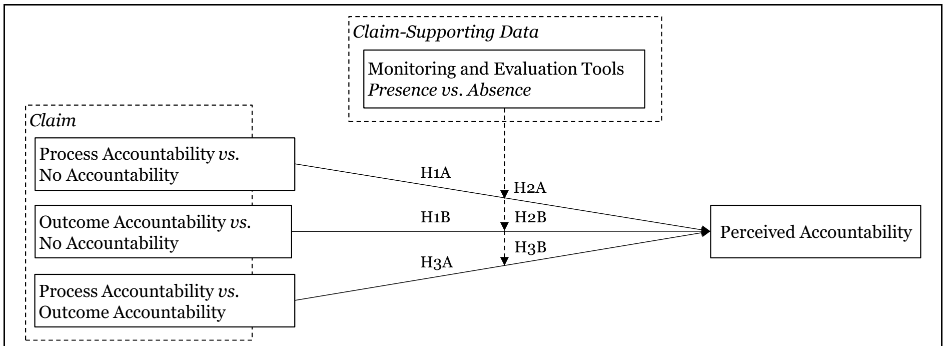



**Figure 3. Research Model**

When individuals are told that they are held accountable, they should feel the need to explain and justify their behavior and decisions to an audience. This perception increases the likelihood that individuals will systematically and deeply consider their behavior, which is also understood as *systematic processing* (Crano & Prislin, 2006). Systematic processing stimulates individuals to take greater account of the factual consequences of their actions and, therefore, to perceive correspondingly higher accountability for their actions (Vance et al., 2015). Accordingly, UI design artifacts for IDEs that visualize arguments of AI developers' process or outcome accountability and consist of corresponding claims should increase their perceived accountability. CLT supports this assumption, as accountability signaling claims should elicit a mental construal in AI developers, albeit perhaps abstract (Trope & Liberman, 2010), that can be assumed to affect AI developers' perceived accountability. We thus formulate our first set of hypotheses:

*H1A: UI design artifacts including (vs. excluding) a process accountability claim create higher perceived accountability among AI developers.*

*H1B: UI design artifacts including (vs. excluding) an outcome accountability claim create higher perceived accountability among AI developers.*

According to Toulmin's model of argumentation (1958), an argument persuades individuals more when its claim is supported by data (i.e., evidence of the claim). Prior IS research indicates that individuals are more likely to accept an expert system's conclusions when its explanations are consistent with Toulmin's model of argumentation (Ye & Johnson, 1995). Arguments that include both claim and claim-supporting data, and thus align with Toulmin's model of argumentation, have also been shown to be more effective in influencing consumers' beliefs (e.g., Kim & Benbasat, 2006; Schneider et al., 2017). Therefore, we assume that UI design artifacts for IDEs that visualize process versus outcome accountability arguments will be more persuasive if their claims are accompanied by claim-supporting data. However, to be persuasive, corresponding claim-supporting data should provide evidence of why AI developers will be held accountable, such as because they have access to tools to monitor and evaluate AI-based systems' development processes or outcomes. Consequently, claim-supporting data should positively moderate the relationship between process versus outcome accountability claims and AI developers' perceived accountability, resulting in our second set of hypotheses:

*H2A: UI design artifacts including a process accountability claim and claim-supporting data (vs. a process accountability claim only) create higher perceived accountability among AI developers.*

*H2B: UI design artifacts including an outcome accountability claim and claim-supporting data (vs. an outcome accountability claim only) create higher perceived accountability among AI developers.*

How do UI design artifacts for IDEs compare that include visualized arguments of process versus outcome accountability in their effects on AI developers' perceived accountability? CLT suggests that arguments for their process accountability affect AI developers' perceptions differently than arguments for their outcome accountability, as the development processes of AI-based systems elicit different temporal distances than the outcomes of AI-based systems. Process accountability focuses on development activities that are currently in progress, thereby referencing the present. Consequently, being held accountable for development processes induces a relatively closer temporal distance and, thus, a more concrete and low-level mental construal (Trope & Liberman, 2003). In contrast, accountability for AI-based systems' outcomes induces a more abstract and high-level mental construal, as AI-based systems' outcomes occur, if at all, only later (Trope & Liberman, 2003). Following CLT, concrete and low-level mental construals are more persuasive than abstract and high-level mental construals, as their reference points are more salient in people's minds. As a result, individuals are more likely to rely on concrete details to form their judgments and beliefs. In contrast, reference points of more abstract and higher-level mental construals are less salient and more likely to be forgotten or pushed out of memory (Trope & Liberman, 2010). Accordingly, UI design artifacts for IDEs representing visualized arguments of process accountability that include only a claim but no claim-supporting data should be more persuasive to AI developers than UI design artifacts for IDEs representing visualized arguments of outcome accountability that include only a claim but no claim-supporting data. Consequently, AI developers should perceive a higher level of accountability when presented with visualized arguments of their process accountability. Therefore, we posit the following hypothesis:

*H3A: UI design artifacts including a process accountability claim (vs. an outcome accountability claim) create higher perceived accountability among AI developers.*

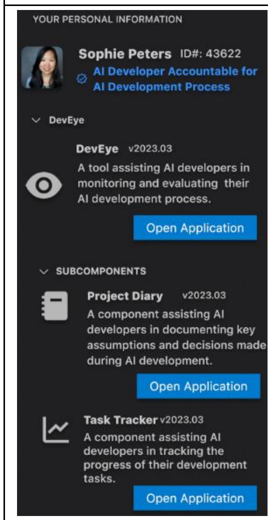
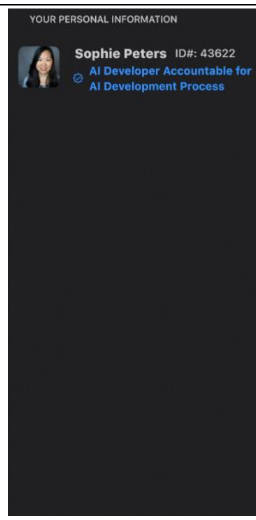
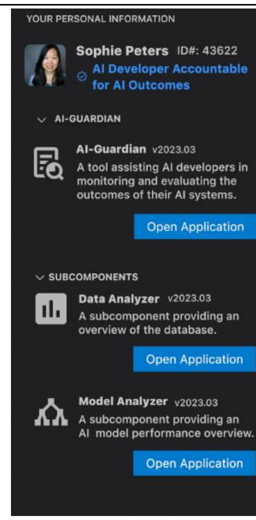
However, suppose that arguments visualized by UI design artifacts for IDEs that already include a claim of process or outcome accountability also include claim-supporting data that provide evidence for why AI developers can be held accountable. In this case, the claim of AI developers' accountability for either AI-based systems' development processes or outcomes receives matching related evidence, which, according to CLT, induces a *construal fit* (Hansen & Wänke, 2010). A construal fit raises AI developers' processing fluency (Hansen & Wänke, 2010), increasing their perceived validity of the corresponding claim and, thus, the extent to which they perceive their accountability. Accordingly, in the case of both process and outcome accountability, arguments that include claim and claim-supporting data will likely be more persuasive to AI developers than arguments that include only a claim. However, if arguments of process accountability that contain only a claim already elicit a concrete, low-level construal of their accountability in AI developers, a construal fit brought about by claim-supporting data can no longer strongly increase this effect. Accordingly, adding claim-supporting data should not strongly concretize AI developers' mental construals of process accountability. In contrast, in the case of outcome accountability, inducing a construal fit by adding claim-supporting data to arguments that already contain a claim may have remarkably positive effects on AI developers' perceived accountability. Suppose added claim-supporting data provide evidence why AI developers can be held accountable for the outcomes of an AI-based system, such as because they have tools to monitor and evaluate the AI-based system. In this case, a construal fit emerges as AI developers can form a more concrete idea of what will happen (i.e., justifying how the AI-based system works and how its outcomes are generated), resulting in a more concrete mental construal. Thus, adding claim-supporting data to an argument should enable AI developers to form a more concrete mental construal of their outcome accountability, which should consequently greatly increase the perceived accountability of AI developers. Accordingly, we derive our final hypothesis:

*H3B: UI design artifacts including an outcome accountability claim (vs. a process accountability claim) and claim-supporting data lead to a greater increase in perceived accountability among AI developers.*

## Research Methodology

### *Method and Experimental Design*

Our study employed scenario-based vignettes for an online experiment and thus followed established guidelines for experimental designs (e.g., Aguinis & Bradley, 2014). Participants were presented with vignettes of IDEs that contained UI design artifacts that visualized process or outcome accountability arguments with monitoring and evaluation tools present or absent. The objective of our online experiment was to examine and contrast the effects of process and outcome accountability arguments visualized as UI design artifacts for IDEs on AI developers' perceived accountability. The effects were also to be examined depending on whether claim-supporting data is present versus absent in the UI design artifacts. To achieve these objectives, we employed a 2 × 2 between-subjects design with an additional hanging control group. This study design enables us to perform both relative and absolute treatment comparisons. Furthermore, our design enables investigating the individual effects of the argument components in isolation, namely claim (process versus outcome accountability) and claim-supporting data (absent versus present), on AI developers' perceived accountability. The hanging control group represents the perceived accountability of AI developers who were not shown any arguments visualized as UI design artifacts (Schneider et al., 2017). Furthermore, as Toulmin's model of argumentation (1958) suggests that arguments that consist only of claim-supporting data without a claim are unlikely to produce the intended effect, we excluded this group. Table 1 shows our resulting five experimental conditions.

<table>
<tr><th colspan="2">Claim: Process Accountability</th><th colspan="2">Claim: Outcome Accountability</th><th rowspan="2"></th></tr>
<tr><th>Data: Present</th><th>Data: Absent</th><th>Data: Present</th><th>Data: Absent</th></tr>
<tr>
<td>
</td>
<td>
</td>
<td>
</td>
<td>
</td>
<td>Hanging Control Group<br><br>No UI Design Artifacts were shown</td>
</tr>
<tr><td colspan="5">Table 1. Five Experimental Conditions<br>(2 × 2 Between-Subject Design Plus Hanging Control Group)</td></tr>
</table>

To test our hypotheses, we created a scenario in the context of developing an AI-based system to calculate bank customers' creditworthiness. We chose this context because AI-based systems for automatic credit scoring of individuals represent a use case of AI-based systems that may be legally classified as high-risk in the future (European Commission, 2021). Additionally, developing such AI-based systems represents current and realistic real-world projects (Strich et al., 2021). The scenario instructed participants to imagine working as an AI developer in a software company and being tasked with either programming the AI-based system for credit scoring or managing its development process. Depending on the allocation, participants were communicated to be accountable for either the AI-based system's development process or its outcomes. Participants were also informed that their employer had provided a customized IDE, which they had to use in their tasks accordingly.

We developed the UI design artifacts for our customized IDE following previous IS studies on accountability (e.g., Adam et al., 2022; Vance et al., 2013, 2015). More specifically, we adapted two accountability mechanisms utilized within these studies to develop our UI design artifacts, namely expectation of evaluation and awareness of monitoring. Expectation of evaluation represented the claim, as participants were communicated to be held accountable for either the AI-based systems' development process or its outcomes. Awareness of monitoring represented the claim-supporting data, as participants were shown the possibility of using tools to explicitly monitor and evaluate either the development process or the outcomes of the developed AI-based system. Our study used these accountability mechanisms as independent variables. All UI design artifacts were consistent across participants and varied only in their accompanying icons and positioning in the customized IDE. Following previous IS research, we added labels to the customized IDE to increase recognition of the UI design artifacts in our experimental conditions (e.g., Adam, 2022; Vance et al., 2013, 2015). Each participant was randomly assigned to one experimental condition and rated their perceived accountability after seeing the customized IDE. Figure 4 shows an example of an IDE that contains a UI design artifact that visualizes process accountability arguments and includes both claim and claim-supporting data.

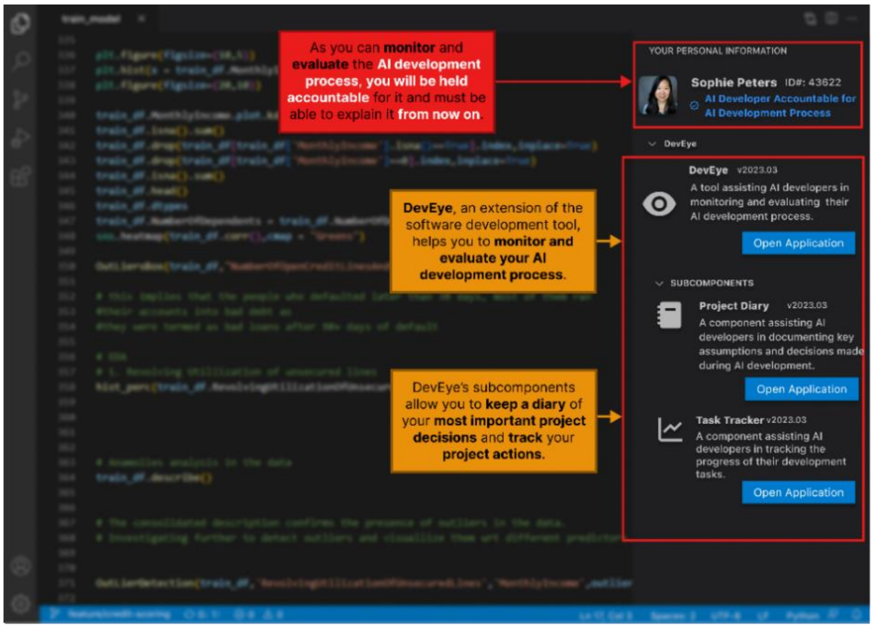



**Figure 4. Sample IDE with UI Design Artifacts**

### *Measurement of Variables*

We followed previous IS studies on accountability (e.g., Adam, 2022; Schmidt et al., 2023; Vance et al., 2015) by adapting and using existing scales to measure AI developers' *Perceived Accountability* (Hochwarter et al., 2005) to elicit AI developers' interpretation and experience of their accountability. The adaptations of the items we used were: "I would feel very accountable for my actions on the presented AI development project that I am working on.", "If the presented AI development project would not go as planned, I would expect to hear about it from my superiors.", and "I would often have to explain my actions in the presented AI development project." The items were measured on a 7-point Likert-type scale, ranging from 1 (*strongly disagree*) to 7 (*strongly agree*). In addition, items were included to assess the realism of the scenario (Lapierre, 2007), as well as demographic data (i.e., *Gender*, *Age*, and *AI Development Experience* in years). Further, we included *Organizational Trust* (Robinson, 1996) and *Impulsivity* (Pogarsky, 2004) as control variables. On the one hand, high levels of organizational trust could lead to more positive perceptions of UI design artifacts caused by a positive bias toward the organization; on the other hand, low levels of organizational trust could lead to more critical evaluations of UI design artifacts due to skepticism. In addition, AI developers with higher impulsivity might be affected by spontaneous reactions rather than careful deliberation regarding their responses to UI design artifacts.

## Analysis and Results

### *Data Collection, Sample Demographics, and Controls*

We conducted three explanatory interviews to ensure functionality and understanding of the scenario and our developed UI design artifacts for IDEs. The interviews, each about 45 minutes long, were conducted with participants who had professional experience managing or developing AI-based systems and were primarily aimed at ensuring that the developed UI design artifacts for IDEs were perceived as having practical relevance and that the scenario presented was evaluated as realistic. Based on the results of these interviews, we optimized our developed UI design artifacts for IDEs. We collected our final dataset via the crowdsourcing platform *Prolific* (prolific.co), known for high reliability and high-quality data (e.g., Palan & Schitter, 2018). Our participants had to be U.S. citizens to ensure a good understanding of the English language, have professional experience in programming AI-based systems, and be employed. In exchange for participating in our study, participants received monetary compensation. Our data collection resulted in 200 participants, of which we excluded 36 participants from the dataset based on failed attention checks (25) and the Relative Speed Index (RSI) criterion (11) to ensure high data quality. The RSI criterion relates the time required to complete the study for a single participant to the time required to complete the study for all participants, with data from participants who complete the experiment at least twice as fast as the typical participant not considered high quality (Leiner, 2019). Accordingly, our final dataset includes 164 participants, resulting in a valid response rate of 82.0%. The average age of our respondents was 36.4 years, of which 75.0% identified as male and 24.4% as female, 84.8% had a Bachelor's degree or higher, and on average, our study participants had 3.2 years of experience developing AI-based systems. A summary of our descriptive statistics is provided in Table 2. Using multiple one-way ANOVAs, we verified that the assignment of participants to the different experiment conditions was random. For the items gender ($F = 0.598$, $df = 4$, $p > 0.05$), age ($F = 2.166$, $df = 4$, $p > 0.05$), AI development experience ($F = 1.041$, $df = 4$, $p > 0.05$), organizational trust ($F = 0.72$, $df = 4$, $p > 0.05$), and impulsivity ($F = 0.236$, $df = 4$, $p > 0.05$) no significant differences were found between the experimental groups. Thus, we can conclude that these individual characteristics were randomly distributed among our experimental conditions.

| **Construct** | **Mean** | **SD** | **Min** | **Max** | **N** |
|---|---|---|---|---|---|
| AI Development Experience | 3.116 | 2.374 | 0.000 | 14.000 | 164 |
| Organizational Trust | 5.837 | 1.086 | 1.000 | 7.000 | 164 |
| Impulsivity | 2.693 | 1.424 | 1.000 | 7.000 | 164 |
| Perceived Accountability | 5.793 | 1.142 | 1.000 | 7.000 | 164 |
| Control Group | 4.914 | 1.139 | 1.000 | 7.000 | 35 |
| Outcome Accountability without Data | 5.609 | 1.039 | 3.000 | 7.000 | 29 |
| Outcome Accountability with Data | 6.115 | 0.746 | 4.000 | 7.000 | 32 |
| Process Accountability without Data | 6.141 | 0.687 | 4.333 | 7.000 | 35 |
| Process Accountability with Data | 6.200 | 0.696 | 4.333 | 7.000 | 33 |

**Table 2. Descriptive Statistics**

## *Measurement Validation*

We performed a principal component analysis with varimax rotation to measure our items' discriminant and convergent validity. Based on Kaiser's criterion, all components with Eigenvalues greater than 1 were extracted, with the resulting three components corresponding to the expected number of theoretical components (Hair et al., 2009). All items loaded above the threshold of 0.75 and thus can be attested strong convergence validity (Hair et al., 2009). High discriminant validity can also be attested as no item cross-loaded higher than 0.30 to a component to which it did not belong. By calculating Cronbach's α, we tested the internal reliability of our items. Since all our constructs exceeded the threshold of 0.80, we can assume high reliability (Streiner, 2003). Table 3 shows our discriminant and convergent validity measures.

| | **Range of Loadings** | **Mean** | **SD** | **α** | **PA** | **OT** | **I** |
|---|---|---|---|---|---|---|---|
| **PA** | 0.847 – 0.867 | 5.793 | 1.142 | 0.851 | **0.809** | | |
| **OT** | 0.865 – 0.909 | 5.837 | 1.181 | 0.909 | 0.190 | **0.885** | |
| **I** | 0.830 – 0.924 | 2.693 | 1.640 | 0.860 | 0.050 | -0.045 | **0.863** |

$p < 0{,}05$; α = Cronbach's alpha; SD = Standard deviation; square root of AVE in bolded cells
PA = Perceived Accountability; OT = Organizational Trust; I = Impulsivity

**Table 3. Internal Consistency and Convergent Validity**

We also performed Harman's single-factor test with principal axis factoring and restricting factors to analyze whether our results were affected by common method bias (CMB). As our results showed that a single factor accounted only for 12.7% and thus well below the critical value of 50%, it is unlikely that CMB occurred in our study or affected our results (Podsakoff et al., 2003). Finally, we evaluated the external validity of our results based on participants' assessments of the scenario's fidelity to reality, which was also collected on a 7-point Likert scale, ranging from 1 (*strongly disagree*) to 7 *(strongly agree*). Since the participants indicated high values and accordingly considered the scenario to be realistic ($\bar{x} = 5.649$, $\sigma = 1.131$), external validity can be assumed.

## Hypotheses Testing

Planned contrast analyses were conducted across all five groups, including the hanging control group, to examine the effects of process versus outcome accountability arguments visualized as UI design artifacts for IDEs on AI developers' perceived accountability. We performed hierarchical regressions to directly compare the effects of UI design artifacts visualizing arguments of process versus outcome accountability on AI developers' perceived accountability in both the presence and absence of claim-supporting data. We excluded the hanging control group from the hierarchical regressions as our objective was to compare the effects of UI design artifacts visualizing arguments of process versus outcome accountability against each other. As can be seen in the results of our planned contrast analyses in Table 4, UI design artifacts that visualize arguments of process accountability with only a claim increase AI developers' perceived accountability compared to the hanging control group ($\bar{x} = 6.141$ vs. $\bar{x} = 4.914$, $p < 0.001$), thereby supporting H1A. Similarly, H1B is supported as UI design artifacts that visualize outcome accountability arguments with only a claim also significantly increase AI developers' perceived accountability compared to the hanging control group ($\bar{x} = 5.609$ vs. $\bar{x} = 4.914$, $p < 0.01$). When AI developers face UI design artifacts that include claim-supporting data in addition to claims, similar results emerge. AI developers presented with UI design artifacts that visualize arguments for process accountability and contain a claim and claim-supporting data exhibit higher perceived accountability compared to AI developers presented with corresponding UI design artifacts that contain only a respective claim ($\bar{x} = 6.200$ vs. $\bar{x} = 6.141$, $p > 0.05$). Yet, H2A is not supported as the difference is insignificant. However, we find support for H2B, as AI developers' perceived accountability increased when presented with UI design artifacts that visualize outcome accountability arguments and include both claim and claim-supporting data, compared to corresponding UI design artifacts that include only a respective claim ($\bar{x} = 6.115$ vs. $\bar{x} = 5.609$, $p < 0.05$). We can therefore conclude that being presented with arguments containing claims of both process and outcome accountability increases AI developers' perceived accountability. While adding claim-supporting data to arguments of outcome accountability further significantly increases AI developers' perceived accountability, this effect does not occur for adding claim-supporting data to arguments of process accountability.

| **Condition** | **Mean (SD)** | **Mean Differences (Value of Contrast)** | | | |
|---|---|---|---|---|---|
| | | 2 | 3 | 4 | 5 |
| Control Group | 4.914 (1.139) | -1.227*** | -1.286*** | -0.695** | -1.201*** |
| Process Accountability<br>*Claim Only* | 6.141 (0.687) | | -0.059 | 0.532* | 0.027 |
| Process Accountability<br>*Claim* and *Claim-Supporting Data* | 6.200 (0.696) | | | 0.591** | 0.085 |
| Outcome Accountability<br>*Claim Only* | 5.609 (1.039) | | | | -0.505* |
| Outcome Accountability<br>*Claim* and *Claim-Supporting Data* | 6.115 (0.746) | | | | |

Significance: * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$

**Table 4. Results of the Planned Contrast Analysis**

In the previous step, we examined the effects of visualized arguments of process and outcome accountability, each consisting of either only a claim or also additional claim-supporting data, on AI developers' perceived accountability compared to a hanging control group. Figure 5 compares the effects of

UI design artifacts for IDEs visualizing process and outcome accountability arguments without the hanging control group. Most strikingly, UI design artifacts that visualize process accountability arguments with and without claim-supporting data elicit high perceptions of accountability among AI developers. Also, including claim-supporting data in UI design artifacts increases AI developers' perceived accountability in the case of both visualized arguments for process and outcome accountability. However, the effect of claim-supporting data in the case of outcome accountability arguments is stronger and only in this case significant. It raises AI developers' perceived accountability almost to the level it holds in the case of process accountability arguments. Therefore, while a claim of process accountability alone is sufficient to increase AI developers' perceived accountability strongly, adding claim-supporting data is advisable regarding arguments concerning the accountability for AI-based systems' outcomes.

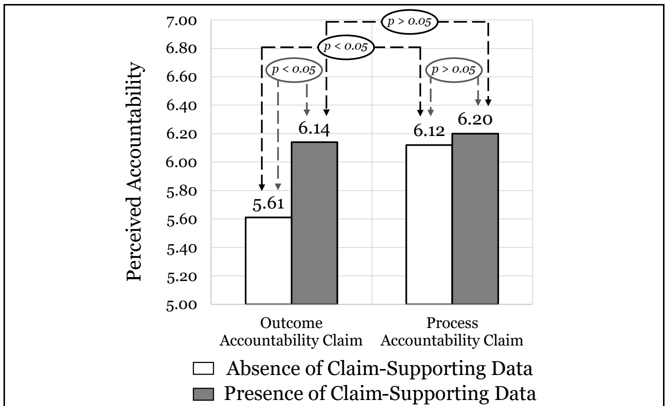



**Figure 5. Comparison of Process vs. Outcome Accountability UI Design Artifacts in IDEs**

| | **Model I** | | **Model II** | |
|---|---|---|---|---|
| | Coefficient | Std. Error | Coefficient | Std. Error |
| Intercept | 4.041 | 0.576 | 3.856 | 0.576 |
| **Manipulation** | | | | |
| Claim[a] | 0.286* | 0.134 | 0.544** | 0.192 |
| Data[b] | 0.193 | 0.139 | 0.461* | 0.193 |
| Claim x Data | - | - | -0.528* | 0.266 |
| **Controls** | | | | |
| Age | 0.001 | 0.007 | 0.000 | 0.007 |
| Gender | 0.068 | 0.153 | 0.092 | 0.151 |
| AI Development Experience | 0.032 | 0.030 | 0.031 | 0.029 |
| Organizational Trust | 0.265 | 0.067 | 0.274*** | 0.066 |
| Impulsivity | -0.019 | 0.048 | -0.023 | 0.266 |
| *$R^2$* | *0.197* | | *0.223* | |
| *Adjusted $R^2$* | *0.151* | | *0.171* | |
| *Residual Std. Error* | *0.755 (df = 121)* | | *0.746 (df = 120)* | |
| *F-statistic* | *4.249*** (df = 7; 121)* | | *4.301*** (df = 8; 120)* | |

* p < 0.05; ** p < 0.01; *** p < 0.001; n =129

[a] Claim was dummy coded with 0 = Outcome Accountability and 1 = Process Accountability

[b] Data was dummy coded with 0 = Absence of Data and 1 = Presence of Data

**Table 5. Hierarchical Regression Results**

To examine H3A and H3B, we performed a hierarchical regression on the data from the experimental conditions of process and outcome accountability, the results of which can be seen in Table 5. We first included the two main effects, claim and claim-supporting data (Model I), before adding the interaction effect (Model II). Both models are highly significant ($p < 0.001$). Model I reveals a statistical effect of claim on AI developers' perceived accountability in the case of process accountability versus outcome accountability, both without claim-supporting data ($\beta = 0.286$, $p < 0.05$), thus supporting H3A. Accordingly, in the absence of claim-supporting data, UI design artifacts for IDEs that visualize arguments of process accountability increase AI developers' perceived accountability more strongly than UI design artifacts for IDEs that visualize arguments of outcome accountability. Model II shows that when an interaction effect is included, UI design artifacts for IDEs that visualize process accountability arguments and consist only of a claim still lead to stronger perceived accountability by AI developers than UI design artifacts for IDEs that visualize outcome accountability arguments and consist only of a claim ($\beta = 0.544$, $p < 0.01$). However, this tendency changes when data is present to support corresponding claims. As can be seen from the negative interaction effect ($\beta = -0.528$, $p < 0.05$), the effects of claim and claim-supporting data are dependent on the presence of each other in the case of outcome accountability. In this, we find support for H3B. Outcome accountability arguments thus benefit particularly strongly from the presence of claim-supporting data. Organizational trust as the only significant control variable suggests that AI developers with high trust in their organization are also receptive to its arguments regarding their accountability for AI-based systems ($\beta = 0.274$, $p < 0.001$). Additionally, the significant effect of the presence of claim-supporting data suggests that the presence of monitoring and evaluation tools has the potential to increase AI developers' perceived accountability, regardless of the specific arguments of process or outcome accountability ($\beta = 0.461$, $p < 0.05$).

## Discussion

AI developers' accountability for development processes and outcomes of AI-based systems is gaining increasing attention. However, governance mechanisms such as AI principles and audits are often considered ineffective, as they remain too abstract and vague to reach AI developers in their daily work. Accordingly, it is important to identify alternative ways to raise awareness among AI developers about their accountability for AI-based systems. Therefore, this study raised the research question of whether UI design artifacts for IDEs that visualize arguments of process versus outcome accountability can increase AI developers' perceived accountability based on CLT and Toulmin's model of argumentation. Four UI design artifacts for IDEs visualizing arguments of AI developers' accountability were examined in an online experiment, of which two UI design artifacts each contained a claim of process or outcome accountability, with claim-supporting data absent or present once each. Our study's empirical results show that when UI design artifacts contain only a claim, AI developers' perceived accountability is increased more when these claims visualize process accountability arguments than outcome accountability arguments. However, when claim-supporting data is present, both visualized arguments increase AI developers' perceived accountability comparably effectively, with outcome accountability (vs. process accountability) especially benefiting from the presence of claim-supporting data. It should be noted that the increase in perceived accountability in the case of visualized process accountability is not statistically significant, which may be due to AI developers' more concrete mental construals of development processes compared to outcomes of AI-based systems.

### *Contributions to Theory and Practice*

Our study makes three important contributions to IS research on accountability for AI-based systems (e.g., Benbya et al., 2021; Mikalef et al., 2022), especially concerning AI developers' perceived accountability (e.g., Martin, 2019; Schmidt et al., 2023). First, our study sheds light on the accountability perceptions of AI developers as key stakeholders in the development and deployment of AI-based systems. Previous IS research on accountability predominantly explored what accountability for AI-based systems should entail and who should assume accountability for AI-based systems (e.g., Ågerfalk et al., 2021; Benbya et al., 2021; Martin, 2019; Wieringa, 2020). Additionally, accountability perceptions and responses of users of AI-based systems have already been examined (e.g., Adam, 2022; Adam et al., 2022). Whereas previous IS research on the accountability perceptions of AI developers has so far only focused on the effects of intrapersonal incongruences of accountability perceptions on AI developers' job satisfaction and role ambiguity (Schmidt et al., 2023), we focused on how AI developers' accountability perceptions can be increased as such. Our study thus further emphasizes the relevance of accountability perceptions of AI developers, who are essential for the development and deployment of AI-based systems and thus for various lifecycle stages of AI-based systems. Emphasizing the accountability perceptions of AI developers and investigating how to increase them is important because it provides a new theoretical perspective on accountability issues of AI-based systems that can be used to ensure that assigned accountability for AI-based systems is accepted and reaches AI developers in their daily work. Notably, our UI design artifacts for IDEs can increase AI developers' perceived accountability without explicitly stating AI principles or conducting extensive AI audits. Further, our developed UI design artifacts for IDEs do not impose any additional workload on AI developers and can increase their accountability perceptions in seconds.

Second, we introduce a more nuanced view of AI accountability by distinguishing between process and outcome accountability as visualized claims in UI design artifacts for IDEs. Previous IS research has treated accountability largely as a unitary, indivisible construct, neglecting to differentiate between temporal or activity-based dimensions (e.g., Adam, 2022; Schmidt et al., 2023; Vance et al., 2013, 2015). Our distinction between process and outcome accountability provides a more precise understanding of accountability in terms of activities and events, as well as the points in time for which accountability is required. Such a distinction is particularly important in the context of AI-based systems, as AI developers are not the only stakeholders involved in the lifecycle of AI-based systems. Furthermore, many events in the lifecycle of AI-based systems can only be anticipated to a limited extent during their development, which should be taken into account when assigning and communicating different types of accountability.

Third, we complement existing IS research on AI developers' perceived accountability by combining CLT with Toulmin's model of argumentation to investigate the effects of different visualized accountability claims. Whereas previous IS research has predominantly focused on the effects of arguments (e.g., Kim & Benbasat, 2006; Schneider et al., 2017) or used Toulmin's model of argumentation to analyze written information (e.g., Cummings & Dennis, 2018), we highlight the effects of different argumentation structures and components leading to different mental construals. The results of our study emphasize the relevance of an adequate argumentation layout (i.e., claim and claim-supporting data) if communicated accountability is consequently to be perceived by reducing psychological distance. This is particularly important given the different stages in the lifecycle of an AI-based system, as these are accompanied by different demands for accountability, which need to be communicated accordingly.

Beyond these contributions to IS research, our study also offers important practical implications, especially for IS control and governance (Saunders et al., 2020). First, while organizations seek to ensure accountability through governance mechanisms such as AI principles and audits, they must also ensure that their AI developers perceive their accountability. Accordingly, organizations need to understand that accountability cannot simply be assigned, but that these assignments must be effectively communicated, including the objective and lifecycle stage for which accountability is assigned. Neglecting effective ways to communicate accountability argumentatively can result in accountability gaps persisting despite organizational efforts, as accountability attributions fall short and accountability for the development processes or outcomes of AI-based systems is not recognized.

Second, our study results suggest that AI developers particularly perceive high accountability for the development processes of AI-based systems when they are confronted with a corresponding claim, and for the outcomes of AI-based systems when they are provided with data such as tools to meet their accountability obligations that support the claim by serving as evidence. Thus, our results highlight the need for organizations to choose different communication strategies depending on the desired accountability. In light of these findings, we recommend that organizations not only signal to their AI developers that they will be held accountable but also be transparent about the reasons for assigning accountability and explain how AI developers can meet the associated accountability obligations.

Third, our results show that using UI design artifacts can enhance AI developers' perceived accountability with minimal efforts or investments, unlike AI principles or audits. In addition, UI design artifacts such as those employed in this study can be implemented in various IDEs for developing and deploying AI-based systems, with customization options to meet specific accountability requirements. Moreover, multiple UI design artifacts can complement each other in a given IDE and potentially build on common data to support accountability claims. Thus, our study helps organizations clearly communicate the specific aspects of AI-based systems their AI developers will be held accountable for, thereby addressing accountability gaps in AI-based systems. Therefore, our study's results can help organizations develop a culture of accountability that encourages AI developers to take responsibility for their work and ensure that their AI-based systems' development processes and outcomes meet ethical and legal requirements.

### *Limitations and Directions for Future Research*

As with any other research study, ours is not without limitations. In light of increasing legislative and regulatory requirements and private initiatives, we believe that these limitations open up fruitful avenues for future research. First, we conducted an online experiment that examined AI developers' perceived accountability in the context of a scenario, but not their actual behavior. Accordingly, future research could implement comparable UI design artifacts in actual IDEs as part of a field study and investigate how AI developers respond when facing UI design artifacts that emphasize their process or outcome accountability, thus examining how organizations can employ governance mechanisms to hold AI developers accountable. In doing so, future research might particularly focus on any (un)ethical behavior of AI developers to address potential harmful societal impacts of AI-based systems already at an early stage. Second, our UI design artifacts for IDEs represented only two accountability mechanisms (e.g., Adam, 2022; Vance et al., 2013, 2015). Future research could take this as an opportunity to examine other accountability mechanisms (i.e., social presence and identifiability) for potential effects on AI developers' perceived accountability, including potential technical challenges that could hinder the development of appropriate IDEs. Third, we focused on accountability for the development processes or outcomes of AI-based systems to keep our UI design artifacts for IDEs clear and concise. Future research could explore the effects of UI design artifacts as accountability attributions for other aspects of the lifecycle of AI-based systems, such as their operation or use, or addressing other stakeholders, such as managers of organizations that provide AI-based systems. Likewise, investigating combinations of different accountability attributions in different contexts based on this study's results is promising. Fourth, our participants were U.S. citizens, so future research may want to examine potential cultural differences in AI developers' perceptions of accountability. Lastly, our scenario included a financial setting. Future research could investigate whether the contexts of AI-based systems affect AI developers' perceived accountability, both for its development process and outcomes.

## Acknowledgments

The authors gratefully acknowledge the funding support by the German Research Foundation (DFG) as part of the project "Accountable Artificial Intelligence-based Systems" (project number 471168026). In addition, the authors would like to thank the anonymous reviewers and associate editor for their valuable comments and suggestions to improve this paper.

## References

Adam, M. (2022). Accountability-Based User Interface Design Artifacts and Their Implications for User Acceptance of AI-Enabled Services. 2022 European Conference on Information Systems, Timişoara, Romania.

Adam, M., Roethke, K., & Benlian, A. (2022). Human Versus Automated Sales Agents: How and Why Customer Responses Shift Across Sales Stages. *Information Systems Research*, *forthcoming*, 1-21. **https://doi.org/10.1287/isre.2022.1171**

Adam, M., Werner, D., Wendt, C., & Benlian, A. (2020). Containing COVID-19 Through Physical Distancing: The Impact of Real-Time Crowding Information. *European Journal of Information Systems*, *29*(5), 595-607. **https://doi.org/10.1080/0960085X.2020.1814681**

Ågerfalk, P., Conboy, K., Crowston, K., Eriksson Lundström, J., Jarvenpaa, S., Mikalef, P., & Ram, S. (2021). Artificial Intelligence in Information Systems: State of the Art and Research Roadmap. *Communications of the Association for Information Systems*, *50*(1), 420-438. **https://doi.org/10.17705/1CAIS.05017**

Aguinis, H., & Bradley, K. J. (2014). Best Practice Recommendations for Designing and Implementing Experimental Vignette Methodology Studies. *Organizational Research Methods*, *17*(4), 351-371. **https://doi.org/10.1177/1094428114547952**

Benbya, H., Pachidi, S., & Jarvenpaa, S. (2021). Special Issue Editorial: Artificial Intelligence in Organizations: Implications for Information Systems Research. *Journal of the Association for Information Systems*, *22*(2), 281-252. **https://doi.org/10.17705/1jais.00662**

Benlian, A. (2022). Sprint Zeal or Sprint Fatigue? The Benefits and Burdens of Agile ISD Practices Use for Developer Well-Being. *Information Systems Research*, *33*(2), 557-578. **https://doi.org/10.1287/isre.2021.1069**

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**

Bovens, M. (2010). Two Concepts of Accountability: Accountability as a Virtue and as a Mechanism. *West European Politics*, *33*(5), 946-967. **https://doi.org/10.1080/01402382.2010.486119**

Brown, S., Davidovic, J., & Hasan, A. (2021). The Algorithm Audit: Scoring the Algorithms That Score Us. *Big Data & Society*, *8*(1). **https://doi.org/10.1177/2053951720983865**

Crano, W. D., & Prislin, R. (2006). Attitudes and Persuasion. *Annual Review of Psychology*, *57*, 345-374. **https://doi.org/10.1146/annurev.psych.57.102904.190034**

Cummings, J., & Dennis, A. R. (2018). Virtual First Impressions Matter: The Effect of Enterprise Social Networking Sites on Impression Formation in Virtual Teams. *MIS Quarterly*, *42*(3), 697-718. **https://doi.org/10.25300/MISQ/2018/13202**

Doney, P. M., & Armstrong, G. M. (1995). Effects of Accountability on Symbolic Information Search and Information Analysis by Organizational Buyers. *Journal of the Academy of Marketing Science*, *24*(1), 57-65. **https://doi.org/10.1177/0092070396024001O**

Proposal for a Regulation of the European Parliament and the Council Laying down Harmonised Rules on Artificial Intelligence (Artificial Intelligence Act) and Amending Certain Union Legislative Acts, (2021). https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELLAR:e0649735-a372-11eb-9585-01aa75ed71a1

Griffith, T. L. (1993). Monitoring and Performance: A Comparison of Computer and Supervisor Monitoring. *Journal of Applied Social Psychology*, *23*(7), 549-572. **https://doi.org/10.1111/j.1559-1816.1993.tb01103.x**

Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2009). *Multivariate Data Analysis* (7th ed., Vol. 24). Upper Saddle River NJ: Prentice Hall.

Hansen, J., & Wänke, M. (2010). Truth From Language and Truth From Fit: The Impact of Linguistic Concreteness and Level of Construal on Subjective Truth. *Personality and Social Psychology Bulletin*, *36*(11), 1576-1588. **https://doi.org/10.1177/0146167210386238**

Hochwarter, W. A., Perrewé, P. L., Hall, A. T., & Ferris, G. R. (2005). Negative Affectivity as a Moderator of the Form and Magnitude of the Relationship Between Felt Accountability and Job Tension. *Journal of Organizational Behavior: The International Journal of Industrial, Occupational and Organizational Psychology and Behavior*, *26*(5), 517-534. **https://doi.org/10.1002/job.324**

Jobin, A., Ienca, M., & Vayena, E. (2019). The Global Landscape of AI Ethics Guidelines. *Nature Machine Intelligence*, *1*(9), 389-399. **https://doi.org/10.1038/s42256-019-0088-2**

Kim, D., & Benbasat, I. (2006). The Effects of Trust-Assuring Arguments on Consumer Trust in Internet Stores: Application of Toulmin's Model of Argumentation. *Information Systems Research*, *17*(3), 286-300. **https://doi.org/10.1287/isre.1060.0093**

Lapierre, L. M. (2007). Supervisor Trustworthiness and Subordinates' Willingness to Provide Extra-Role Efforts 1. *Journal of Applied Social Psychology*, *37*(2), 272-297. **https://doi.org/10.1111/j.0021-9029.2007.00160.x**

Leiner, D. J. (2019). Too Fast, Too Straight, Too Weird: Non-Reactive Indicators for Meaningless Data in Internet Surveys. *Survey Research Methods*, *13*(3), 229-248. **https://doi.org/10.18148/srm/2019.v13i3.7403**

Lerner, J. S., & Tetlock, P. E. (1999). Accounting for the Effects of Accountability. *Psychological Bulletin*, *125*(2), 255-275. **https://doi.org/10.1037/0033-2909.125.2.255**

Martin, K. (2019). Designing Ethical Algorithms. *MIS Quarterly Executive*, *18*(2), 129-142. **https://doi.org/10.2139/ssrn.3056692**

Mikalef, P., Conboy, K., Lundström, J. E., & Popovič, A. (2022). Thinking Responsibly About Responsible AI and 'the Dark Side' of AI. *European Journal of Information Systems*, *31*(3), 257-268. **https://doi.org/10.1080/0960085X.2022.2026621**

Mittelstadt, B. (2019). Principles Alone Cannot Guarantee Ethical AI. *Nature Machine Intelligence*, *1*(11), 501-507. **https://doi.org/10.1038/s42256-019-0114-4**

Mittelstadt, B. D., Allo, P., Taddeo, M., Wachter, S., & Floridi, L. (2016). The Ethics of Algorithms: Mapping the Debate. *Big Data & Society*, *3*(2), 1-21. **https://doi.org/10.1177/2053951716679679**

Palan, S., & Schitter, C. (2018). Prolific.ac—A Subject Pool for Online Experiments. *Journal of Behavioral and Experimental Finance*, *17*, 22-27. **https://doi.org/10.1016/j.jbef.2017.12.004**

Podsakoff, P. M., MacKenzie, S. B., Lee, J.-Y., & Podsakoff, N. P. (2003). Common Method Biases in Behavioral Research: A Critical Review of the Literature and Recommended Remedies. *Journal of Applied Psychology*, *88*(5), 879-903. **https://doi.org/10.1037/0021-9010.88.5.879**

Pogarsky, G. (2004). Projected Offending and Contemporaneous Rule-violation: Implications for Heterotypic Continuity. *Criminology*, *42*(1), 111-136. **https://doi.org/10.1111/j.1745-9125.2004.tb00515.x**

Raji, I. D., Smart, A., White, R. N., Mitchell, M., Gebru, T., Hutchinson, B., Smith-Loud, J., Theron, D., & Barnes, P. (2020). Closing the AI Accountability Gap: Defining an End-to-End Framework for Internal Algorithmic Auditing. 2020 Conference on Fairness, Accountability, and Transparency, Barcelona, Spain.

Robinson, S. L. (1996). Trust and Breach of the Psychological Contract. *Administrative Science Quarterly*, *41*, 574-599. **https://doi.org/10.2307/2393868**

Saunders, C., Benlian, A., Henfridsson, O., & Wiener, M. (2020). IS Control & Governance. In A. Bush & A. Rai (Eds.): MIS Quarterly Research Curations.

Schmidt, J.-H., Bartsch, S. C., Adam, M., & Benlian, A. (2023). Accountability Incongruence and Its Effects on AI Developers' Job Satisfaction. 2023 European Conference on Information Systems, Kristiansand, Norway.

Schneider, D., Lins, S., Grupp, T., Benlian, A., & Sunyaev, A. (2017). *Nudging Users Into Online Verification: The Case of Carsharing Platforms* 2017 International Conference on Information Systems, Seoul, South Korea.

Simonson, I., & Staw, B. M. (1992). Deescalation Strategies: A Comparison of Techniques for Reducing Commitment to Losing Courses of Action. *Journal of Applied Psychology*, *77*(4), 419-426. **https://doi.org/10.1037/0021-9010.77.4.419**

Streiner, D. L. (2003). Starting at the Beginning: An Introduction to Coefficient Alpha and Internal Consistency. *Journal of Personality Assessment*, *80*(1), 99-103. **https://doi.org/10.1207/S15327752JPA8001_18**

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 Identity. *Journal of the Association for Information Systems*, *22*(2), 304-324. **https://doi.org/10.17705/1jais.00663**

Toulmin, S. (1958). *The Uses of Argument* (1st ed.). Cambridge University Press.

Toulmin, S. (2003). *The Uses of Argument* (2nd ed.). Cambridge University Press.

Trope, Y., & Liberman, N. (2003). Temporal Construal. *Psychological Review*, *110*(3), 403-421. **https://doi.org/10.1037/0033-295X.110.3.403**

Trope, Y., & Liberman, N. (2010). Construal-Level Theory of Psychological Distance. *Psychological Review*, *117*(2), 440. **https://doi.org/10.1037/a0018963**

van den Broek, E., Sergeeva, A., & Huysman, M. (2021). When the Machine Meets the Expert: An Ethnography of Developing AI for Hiring. *MIS Quarterly*, *45*(3), 1557-1580. **https://doi.org/10.25300/MISQ/2021/16559**

Vance, A., Lowry, P. B., & Eggett, D. (2013). Using Accountability to Reduce Access Policy Violations in Information Systems. *Journal of Management Information Systems*, *29*(4), 263-290. **https://doi.org/10.2753/MIS0742-1222290410**

Vance, A., Lowry, P. B., & Eggett, D. (2015). Increasing Accountability Through User-Interface Design Artifacts. *MIS Quarterly*, *39*(2), 345-366. **https://doi.org/10.25300/MISQ/2015/39.2.04**

VerLinden, J. G. (1998). Verifiers in Everyday Argument: An Expansion of the Toulmin Model. Eighty Fourth Annual Meeting of the National Communication Association.

Wendt, C., Werner, D., Adam, M., & Benlian, A. (2022). Influencing Crowding at Locations with Decision Support Systems: The Role of Information Timeliness and Location Recommendations. *Decision Support Systems*, *160*. **https://doi.org/10.1016/j.dss.2022.113817**

Wieringa, M. (2020). What to Account for When Accounting for Algorithms: A Systematic Literature Review on Algorithmic Accountability. *2020 ACM Conference on Fairness, Accountability, and Transparency* 2020 Conference on Fairness, Accountability, and Transparency, Barcelona, Spain.

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**

Zhang, Z., Hummel, J. T., Nandhakumar, J., & Waardenburg, L. (2020). Addressing the Key Challenges of Developing Machine Learning AI Systems for Knowledge-Intensive Work. *MIS Quarterly Executive*, *19*(4), 221-238. **https://doi.org/10.17705/2msqe.00035**