Computer Supported Cooperative Work (CSCW)
https://doi.org/10.1007/s10606-024-09499-6

ECSCW CONTRIBUTION

Trusting Intelligent Automation in Expert Work: Accounting Practitioners' Experiences and Perceptions

Saara Ala-Luopa*[1], Thomas Olsson[1], Kaisa Väänänen[1], Maria Hartikainen[1] & Jouko Makkonen[1]

*[1] *Faculty of Information Technology and Communication Sciences, Tampere University, Tampere, Finland (E-mail: saara.ala-luopa@tuni.fi)*

Accepted: 5 April 2024

Abstract. AI-based applications are increasingly used in knowledge-intensive expert work, which has led to a discussion regarding their trustworthiness, i.e., to which degree these applications are ethical and reliable. While trust in technology is an important aspect of using and accepting novel information systems, little is known about domain experts' trust in machine learning systems in their work. To provide a real-life, empirical perspective on the topic, this study reports findings from an interview study of accounting practitioners' ($N=9$) trust in intelligent automation in their work. The findings underline the holistic nature of trust, suggesting that contextual and social aspects, such as participatory design practices, shape domain experts' trust in intelligent automation. For instance, the participants emphasize their contribution to product development and open communication with the system developers. In addition, the findings shed light on the characteristics of domain experts as technology users, such as the necessity of situation-specific expert knowledge when evaluating the systems' reliability. Thus, our findings suggest that trust in intelligent automation manifests at different levels, both in human-AI interaction and interpersonal communication and collaboration. This research contributes to the existing literature on trust in technology, especially AI-powered applications, by providing insights into trust in intelligent automation in expert work.

Keywords: Trust, Intelligent automation, Expert work, Interview study

1 Introduction

Recent developments in AI are expected to increase productivity and efficiency in domain experts' work by transferring repetitive tasks from humans to algorithms, supporting experts' decision-making processes, and contributing to increased accuracy and enhanced performance (Brynjolfsson and McAfee 2011, 2017; Jarrahi 2018). In this context, expert work refers to, e.g., non-routine tasks, applying cognitive skills and abstract knowledge held by an individual in an autonomous

Published online: 30 April 2024

or authoritarian position in a specific domain (Pakarinen and Huising 2023). Experts, compared to laypeople, are understood to demonstrate superior performance whose cumulative experience makes decision-making less effortful and more intuitive (Hoffman 1998; Ericsson 2014). While novel AI-based applications are increasingly developed to support expert work, trust in technology has been considered a significant antecedent supporting the use and acceptance of novel technology (McKnight et al. 2011; de Visser et al. 2018; Siau and Wang 2018; Gefen et al. 2003; Pavlou and Gefen 2004). In Computer-Supported Cooperative Work (CSCW) and Human–Computer Interaction (HCI) research, trust in technology is typically defined as the trustor's "attitude or willingness to accept vulnerability and uncertainty in a situation that aims to achieve a certain goal, performed by the trustee without an ability to monitor or control this party" (Lee and See 2004; Mayer et al. 1995; Vereschak et al. 2021).

Previous research on trust in technology has extensively explored trust in different contexts, such as online trust or e-trust (Corritore et al. 2001, 2003; Pennanen et al. 2007), trust in e-commerce (Pavlou 2003; McKnight et al. 2002), trust in recommendation agents (Benbasat and Wang 2005, Wang and Benbasat 2008; Komiak and Benbasat 2006) and trust in global virtual teams (Jarvenpaa et al. 1998). However, trust is dynamic and ever-changing, which motivates the need to reconsider trust in novel AI-based applications, including machine learning, especially in the context of expert work. We justify this with the following arguments. First, novel AI-based applications are learning systems that develop through interaction after being deployed into the target context. This might result in unpredictable outcomes and opaqueness when compared to the system's initial and intended functionality, which can undermine trust. Second, machine-learning models might not be explainable or understandable to their users (or even developers). This has raised discussion on trustworthy AI, referring to responsible, reliable, and ethical AI advancement, which is expected to necessitate the involvement of people using AI or affected by it (Jacovi et al. 2021; Zicari et al. 2021; Ashoori and Weisz 2019). So far, this remains scarce: Recent research on AI design and development shows that the end- users' perspective is mainly excluded from the processes (Hartikainen et al. 2022). In addition, domain expert users' perspective on trust in technology (and AI) has remained undefined in the field of CSCW (Vereschak et al. 2021), and when situated in a real-life context (Lockey et al. 2021). We approach the topic from a socio-technical perspective in which implementing a system in CSCW includes both technical, organizational, and social aspects, such as employee needs and respect for local practices (Bullinger-Hoffman et al. 2021; Mumford 2000). This motivates our third argument which maybe the most important in the context of CSCW: well-known challenges associated with technology design and development in a work-life context are related to the lack of mutual understanding between industry practitioners and domain experts. For instance, articulation

of work aims to enhance awareness by creating visibility to the work practices and the task at hand (Suchman 1995; Schmidt and Bannon 1992). Understanding domain expert users' perceptions of trust in intelligent technology in their work increases the probability of developing appropriate, usable, and acceptable AI applications for human-AI collaboration in expert work.

In this article, we explore accounting practitioners' (henceforth, *accountants*) trust in intelligent automation (IA) in their work. Accounting as an AI-application context offers an interesting area to explore the integration of AI-based technologies and modern knowledge work: accounting has been considered a particularly amenable domain for AI automation with repetitive and routine tasks and precisely defined data processing (Frey and Osborne 2017; Leitner-Hanetseder et al. 2021). Through semi-structured interviews ($N=9$), we collected data about accountants' experiences and perceptions of trust toward a machine learning (ML) system designed to automate invoice processing. To explore domain experts' willingness to trust IA, we ask: *What aspects contribute to the accountant practitioners' trust in intelligent automation?* In the target context, the accountants' work was monitored by an intelligent system that collects data about their invoice processing and, when reaching a certain level of confidence, makes suggestions for invoice automation. At the time of the study, most participants were gradually moving towards IA, but only one of them had actualized the system's full automatization benefits. The aim of the system is full automation, which processes invoices without human intervention. As an outcome, certain parts of the accountants' work are transferred to a machine-learning system, resulting in cost-efficiency and enhanced work practices. The studied system and its target context are elaborated in Section 3.

This study contributes to CSCW research by providing empirical insights into trust in IA in knowledge experts' work, specifically within the context of accounting. Exploring the impact of novel AI-based applications on expert work, we can reassess domain experts' trust in technology and how it should be approached given the inherent opaqueness and uncertainty of machine learning systems. By focusing on a specific user group and context, this study sheds light on the dynamic phenomenon of trust in IA, enhances our understanding of the collaborative partnership between humans and AI-embedded systems, and increases our understanding of appropriate, trustworthy, and acceptable AI design and development.

2 Related work

2.1 Trust in technology

Concerning research disciplines, our research builds on Computer-Supported Cooperative Work, Human–Computer Interaction, and Information Science.

Previous research on trust in technology has adapted a trust definition from organization and management science, which describes trust as the "willingness of a party to be vulnerable to the actions of another party based on the expectation that the other will perform a particular action important to the trustor, irrespective of the ability to monitor or control that party" (Mayer et al. 1995). Trust can also be perceived specifically from the perspective of automation, in which trust is defined as an "attitude, that an agent will help achieve an individual's goals in a situation characterized by uncertainty and vulnerability" (Lee and See 2004). Recent CSCW research exploring human-AI trust uses both definitions (Vereschak et al. 2021). Both descriptions have similar connotations: trust concerns a trustor's (entity that trusts) willingness for vulnerability, a certain attitude towards the trustee (entity that is trusted), and positive expectations regarding the likelihood of a favorable outcome. It is noteworthy that trust can be confused with related concepts, such as confidence (if vulnerability does not exist), reliance (the decision to follow someone's recommendation), and compliance (the decision to ask for a recommendation) (Vereschak et al. 2021). Therefore, it is important to understand the general aspects of trust, although novel AI-based applications might complement this existing understanding.

Previous research acknowledges certain characteristics that influence trust. In interpersonal trust, such characteristics are ability (competence), benevolence (intents and motivations), and integrity (acceptance) (Mayer et al. 1995), whereas their technical correspondence would be functionality, reliability, and helpfulness (Lankton and McKnight 2011; Muir and Moray 1996). *Functionality* is the degree to which the trustor anticipates the technology will have the functions or features needed to accomplish one's tasks. *Reliability* refers to an expectation that the technology will continually operate properly or will operate in a consistent, flawless manner, providing adequate and responsive help (*helpfulness*). Trust in automation, on the other hand, emphasizes *performance, process, and purpose* as general bases of trust. These factors inform the trustor about, e.g., what the automation does (performance), how the automation operates (process), and why the automation was developed (purpose) (Lee and Moray 1992; Lee and See 2004; Madhavan and Wiegmann 2007). It is noteworthy that trust does not always translate into behavior: even though trust exists, the trustor might not be willing to act on it (Vereschak et al. 2021). In our study, we are interested in the domain expert users' experiences and perceptions of trust and the aspects that influence their willingness to trust in IA. Previous research on trust in technology has not thoroughly explored domain experts' perspectives, but we would expect, e.g., reliability, process, and purpose to play a role in (domain experts') trust in AI, as these can be considered relatively essential in any technology use.

Despite the dynamic nature of trust, previous research on trust in automation considers interesting concepts that can be well-suited also for trust in novel AI-based applications. For instance, *appropriate trust* might be useful when

exploring trust in AI in expert work, referring to a situation in which the user's trust and the technology's capabilities are aligned (calibrated), and the user knows when to trust the technology, avoiding over-trusting and distrusting (Parasuraman and Riley 1997; Lee and See 2004). Recent research on trust in AI acknowledges that appropriate trust allows the users to apply their knowledge and improve the outcomes in situations where AI models may have limitations (Zhang et al. 2020a, b). A similar concept is *situational trust,* which emphasizes the context-dependent nature of trust, such as workload and risk regarding the trust decision, dividing trust antecedents into three different groups: the trustor, the trustee, and the context/environment. Situational trust underlines the holistic perspective on trust in technology (Hoff and Bashir 2015). This is especially important when considering the uncertainty of AI-based applications: machine-learning systems might produce different outcomes for different users, underlining the need for end-user control over these technologies. Both appropriate trust and situational trust are closely related to the CSCW concept of appropriation/tailorability which refers to individuals' adaptation to technology in their situation—technology is expected to support local practices (e.g., Orlikowski 1992, 1995; Dourish 2003). The gap between social requirements and technical feasibility is often argued to be the reason why IT systems are not supporting real work efficiently: representation of work, if created in the interest of introducing new information technologies, is unlikely to include aspects of the work considered beyond the reach of those technologies (Ackerman 2000; Suchman 1995). Such awareness is not possible to achieve without understanding the contextual and situational variance (Dourish and Bellotti 1992).

2.2 Trust in AI

Current research on trust in AI can be approached from the perspectives of emotional trust or cognitive trust, in which the former focuses on emotions/affects and the latter on calculated and rational trust decisions (Schoorman et al. 2007; Glikson and Woolley 2020). Cognitive trust is emphasized in initial trust (Gefen et al. 2003), and it may be based on second-hand knowledge, for instance, reputation, and guides the trustor's evaluation of the trustee's trustworthiness (e.g., McKnight et al. 1998). In this study, we focus on cognitive trust due to the context in which technology is embedded (expert work). Exploring cognitive trust in this context is interesting because domain experts are typically in a position in which they are expected to make rational decisions. In addition, complex cognitive tasks are also subject to change due to the recent development of AI (Saßmannshausen et al. 2021).

When researchers examine cognitive trust in AI, they measure it as a function of whether users are willing to take information or advice and act on it, as well as whether they see the technology as helpful, competent, or useful (Glikson and Woolley 2020). Especially usefulness can be considered important in the context

of expert work because tasks are typically conducted in fast-paced environments emphasizing the relevance of information and tools. However, it is interesting to what degree the usefulness can be evaluated (and when) because the benefits of AI applications might become visible in delay. When considering cognitive trust, it also reflects the trustworthiness of the human-AI interaction and collaboration (referring to the degree to which AI helps users to make rational decisions and the reliability of this collaboration). Trustworthy AI is a concept that has occurred on the ethical frameworks and guidelines around AI, promoting an idea that individuals, organizations, and societies will only ever be able to achieve the full potential of AI if trust can be established in its development, deployment, and use (HLEG AI 2019).

Glikson and Woolley (2020) conducted a review of users' cognitive and emotional trust in AI. According to this study, cognitive trust is based on perceptions of trustee reliance and competence, as well as whether they see the technology as helpful, competent, or useful. The results of this review reveal the important role of AI's transparency, reliability, task characteristics, and immediacy behaviors in developing cognitive trust. Transparency is the level to which the underlying operating rules and inner logic of the technology are apparent to the users. Reliability means exhibiting the same and expected behavior over time. In the case of AI, reliability is often difficult to assess, especially in the context of high machine intelligence, as learning from data can lead technology to exhibit different behavior, even if the underlying objective function remains the same. Task characteristics refer to the work the technology is performing, such as whether it deals with largely technical or interpersonal judgments. Immediacy behaviors refer to personalization, interactive or socially oriented abilities, or gestures intended to increase interpersonal closeness, such as proactivity and responsiveness. Glikson and Woolley (2020) underline that there is a growing need for research in real-life settings, such as organizations that are already using AI in their management or decision-making systems, and more field studies on the topic are needed.

Van der Werff et al. (2021) consider trust in AI to include micro and macro-level trust cues that provide information that influences trust decisions. Micro-level cues arise from information about the trustee or about the trustor, occurring from interpersonal interaction and evaluation, for instance, through experiences of interacting with technology. Macro-level cues occur in a wider contextual environment in which the AI is embedded, encompassing information that arises from, e.g., service providers, complementary technologies, and regulatory standards. This refers to a situation in which information about organizations, systems, or technology can shape trusting attitudes, decisions, and behavior regarding another trust referent (Bachmann 2001; Kosonen et al. 2008; Shapiro 1987). According to this study, both macro and micro-level cues affect users' trust in AI. At the macro level, the influencing trust cues are, for instance, organizational ability to build AI services, benevolence in creating services for customer

needs, and integrity regarding data privacy. Also, macro or organizational-level integrity was signaled through open communication, privacy protection, and fair information practices. On a micro-level, the trust cues arise from interaction with the technology, emphasizing the importance of ability, integrity, and propensity to trust. They also found that context-relevant knowledge and information asymmetry might play an important role in the development of trust propensity. Van der Werff et al. (2021) argue that AI cannot operate within a vacuum and should be studied in the context of trust, considering the trust-related cues that originate at more macro levels.

Vereschak et al. (2021) conducted a review of the methods to empirically investigate trust in AI-assisted decision-making. Their findings provide a practical perspective on studying human-AI trust in the field of CSCW and HCI both from quantitative and qualitative approaches. They explore, for instance, elements related to the participants' experience and expertise. The study findings show that previous research does not typically involve participants with prior experience with either the AI-embedded system or the task associated with it, although it is suggested. In addition, they discuss different aspects when elaborating on the task and how the task is integrated into the experiment. For instance, immediate feedback allows participants to dynamically update their level of trust. In most cases, the feedback is related to the participants' performance instead of system performance. They also underline a need to control participants' expectations, focusing on positive aspects. This could be done, for example, by providing instructions, error-free initial experience, and guaranteeing a minimum level of accuracy in system behavior (min. 60–70% accuracy depending on the context). They also found that most of the research included in the review evaluated trust as one of the multiple factors of users' experience instead of solely focusing on human-AI trust. It is noteworthy that most of their observations are based on studies conducted in laboratory settings, underlining a need for human-AI trust in real-life settings.

2.3 Domain expert users' trust in AI

While previous research provides interesting insight into trust in AI, there is a lack of studies that explore domain experts' perspectives regarding trust in real-life settings and during technology use. Few studies have made attempts to study domain experts' trust in AI with a qualitative approach. For instance, Bedué and Fritzsche (2022) conducted an interview study ($N = 12$) with company decision-makers (different domains) to explore their trust in AI, underlining the importance of trust in technology adoption. Focusing on the early stages of technology development, they found access to knowledge, transparency, explainability, certifications, and self-imposed standards and guidelines to be the main determinants of trust in AI. In addition to trust, these elements were found to reduce uncertainties and enhance AI adoption intentions. Access to knowledge refers to

new skills and more technical skill sets required when using AI. They argue that currently, this information is held by information technology providers, and, e.g., institutions lack knowledge on AI evaluation and development. Transparency and explainability cover both the internal functioning of a model as well as interpretability, e.g., verbal explanations or visualizations. Certifications are considered a signal of the company's integrity. For instance, self-imposed standards are perceived as crucial to fair and reasonable AI regulation. Bedué and Fritzsche (2022) emphasize that trust in AI needs to be understood in a wider context, underlining the collaboration of multiple institutes and alignment of AI with social norms and compliance with policies and standards.

Saßmannshausen et al. (2021) explored the antecedent variables on trust in AI within production management from three facets: AI characteristics, trustors' (domain experts) characteristics, and decision situation characteristics, underlining trust as essential to successful human-AI cooperation. A qualitative study was conducted with four expert interviews for hypothesis development. According to their findings, users' digital affinity or interest and pleasure in learning and using digital technologies was considered an important antecedent to increase experts' trust in AI. Exploring the AI characteristics, perceived ability (e.g., AI's ability or competence on task), and perceived comprehensibility (e.g., quality and plausibility of AI explanation) was found to positively affect trust. Considering the decision situation characteristics, the predictability and the error costs were found to indirectly increase trust in AI through perceived ability and/or perceived comprehensibility. Error costs refer to potential impact, consequences, and costs caused by wrong AI decisions, affecting the learned trust and helping users to understand the situation and AI capabilities. Based on their findings, they provide design guidelines for socially sustainable human-AI interaction in production management, such as designing explainable AI. For trust calibration, they suggest verifying the AI decision by another system or a human being. They also recommend introducing AI to digitally competent employees and even state that this should become a criterion for recruiting. Expert status as a trustor characteristic was not found to contribute to trust in their research.

Lockey et al. (2021) conducted a review of the antecedents of trust in AI, exploring vulnerabilities of key stakeholder groups in relation to AI systems, including domain experts, end-users, and societal perspectives. They argue that societal adoption of AI is recognized to depend on stakeholders' trust in AI, and understanding the trustor (e.g., domain expert) is particularly important in the context of AI as it influences the nature of the risks and vulnerabilities inherent in trusting an AI system. According to the findings, the key vulnerabilities faced by domain experts relate to professional knowledge, skills, identity, and reputation, for instance, loss of expert oversight or professional over-reliance. Transparency and explainability are considered as a trust challenge, emphasizing the experts' ability to understand, explain, and justify AI decisions to other

stakeholders. Domain experts need to remain accountable, e.g., for the accuracy and fairness of AI output and privacy, and thus, they are expected to be able to provide human oversight in the use of AI. The review also introduces reputational and legal risks, such as biased results or inappropriate data usage, influencing domain experts' trust in AI. They call for further research to understand what influences stakeholders or well-calibrated trust in AI systems because high trust might not always be appropriate. For instance, AI explanations might cause users to misplace trust in inaccurate AI outcomes, underlining a need for correct evaluation of a system's trustworthiness.

To this end, we can perceive trust to be important in user's AI adoption, human-AI cooperation, and societal AI adoption. Understanding the technology and the possibility of evaluating its trustworthiness, such as reliability and predictability (e.g., through feedback), is considered important for trust. When focusing on domain experts as technology users, trusting AI might require changes in their skillsets, e.g., regarding users' technical skills. In addition, contextual aspects, such as a given task or situation, can shape users' perceptions regarding trust. Trust in AI can also be influenced by certifications or standards or acknowledging potential negative consequences, e.g., reputational risks in error situations. To this end, it seems that trust in AI is holistic and sociotechnical by nature, expanding beyond human-AI interaction: although trust can be supported with technical means (e.g., explanations), social context also contributes to trust.

3 Context: studied professional and technological domain

The professionals in this study are accountants, whose main tasks include, e.g., their clients' financial administration, financial analysis and planning, tax preparations, and invoice processing. Accountants in our study work monthly, providing an overview of the financial activities at the end of each month. The accounting domain was selected for this study because it represents a knowledge-intensive work expected to be largely affected by novel AI technologies (Frey and Osborne 2017; Leitner-Hanetseder et al. 2021). In addition, accountant professionals themselves acknowledge that their profession undergoes significant disruptions, and to remain in their profession, they feel the need to adapt to the changes, new roles, and tasks emerging from human-AI collaboration (Asatiani et al. 2020; Leitner-Hanetseder et al. 2021).

The machine learning (ML) system we focused on for this study was developed by an external vendor to process invoices automatically. The aim is to automate mechanical and routine tasks in accounting, such as value-added tax (VAT) processing, cost allocations and accrued expenses. The system is tailored to support existing organizational processes and practices, and automated invoice processing can be integrated into an organization's existing

system or deployed in a separate interface that is developed around the system. The automation process is a three-fold process, starting from (1) automated learning. The system tracks the accountant's work and learns on a customer-specific basis how invoice processing is done, automatically creating client/company-specific rules. This is followed by (2) quality assurance. The system starts processing invoices, and the accountant monitors and evaluates the system outcomes. If the accountant corrects or makes changes to the invoices, the company-specific rules are modified based on these changes. The third step of the process is (3) intelligent automation. Those invoices that do not need any input from the accountant will be processed fully automatically without human intervention if the accountant permits this. Table 1 presents the automation levels according to the study participants.

The system deployment process is iterative. In most cases, the ML system is introduced to the accountants after the decision on the deployment has already been agreed upon with client company managers. There are multiple meetings with system developers and accountants: in the first meeting, the developers introduce the system and its practical functionalities, for instance, how IA works and what kind of clients are suitable (or not) to automate with the IA. A second meeting is organized after the accountants have used the system for ca. two months. This meeting focuses on fine adjustments in which developers aim to encourage accountants to use IA and help them overcome possible challenges or issues during the use. This is followed by regular monitoring of the use and a push for increasing automation. It is noteworthy that the deployment process emphasizes participatory/collaborative design practices in which the end-users are actively included in the design and development of the product (e.g., Auernhammer 2020).

Table 1 Overview of the study participants and their relations with the studied system. We label the participants according to their system use based on three categories: (a) indicates the system use of less than six months, (b) indicates system use of less than two years, and (c) indicates system use of more than two years

Participant ID	Use of system	Clients in the system	Level of intelligent automation
P1a	4 months	4	Level 1 (automated learning)
P2b	1 year	10	Level 2 (quality assurance)
P3c	2 years	6	Level 2 (quality assurance)
P4a	6 months	1	Level 1 (automated learning)
P5b	1 year	4	Level 2 (quality assurance)
P6c	3 years	10	Level 2 (quality assurance)
P7b	1 year	14	Level 3 (intelligent automation)
P8c	2 years	20	Level 2 (quality assurance)
P9c	2 years	1	Level 3 (intelligent automation)

4 Methodology

4.1 Study design and procedure

This study is qualitative research, utilizing semi-structured expert interviews. This method is useful for gathering information about the end-users' experiences and perceptions in qualitative HCI research (Blandford et al. 2016). The interviewed domain experts ($n=9$) are accountants using the specific ML system in their work for the purpose of IA. Despite the relatively small number of participants, they represent a homogenous user group working on similar tasks (invoice processing) and in comparable environments (accounting firms). In a qualitative study, four to twelve interviewees are considered sufficient (Kuzel 1992; Saunders 2012). The study approach is interpretive and considers the social and cultural context of the information system and the influence of the system on its context (Yang et al. 2020). The interviews were mainly conducted by the first author, except for three interviews, which were conducted by fourth (2) and fifth (1) authors. All interviews were recorded after asking for the participant's consent.

The key themes of the interviews included the description of the participants' current use of the system, their positive and negative experiences with IA, their perceptions of trust, and the perceived risks and pitfalls regarding the system. We intended to examine the topic as broadly as possible within the scope of this study. Therefore, the interview questions were formulated according to the previous research on trust in technology and trust in artificial intelligence (underlining, e.g., risks and vulnerabilities to understand trust as a theoretical concept) and discussing trust with the system developers (i.e., how they perceive trust in the system). Intentionally, we also included questions related to human-like trust instead of system-like trust: this was to explore if anthropomorphism had a role in trust formation. In addition, participants were shown a product demo to demonstrate their system use and use practices in-depth, allowing participants to discuss trust through product features and in an actionable manner. This was considered useful because understanding trust might be difficult to identify or verbalize among study participants who are not familiar with the theoretical perspective on trust. While the full list of interview questions can be found in the supplementary material, this article focuses on reporting the participants' experiences and perceptions regarding their trust in the system, as well as the current and future system use.

4.2 Study participants

The interviewees were recruited in collaboration with the developing company. After the first interviews, snowball sampling was utilized to avoid biased participant selection. Despite similarities in participants' professional status, they had diversity regarding the system use, e.g., the number of clients transferred to the system and the level of IA (Table 1). Surprisingly, this diversity was not found to affect

their experiences or perceptions regarding trust remarkably. The levels of IA in Table 1 are based on the process description elaborated by the developing company, expressing gradual transfer toward full automation. It is noteworthy that the accountants as study participants have a dual role: they represent both end users and the domain experts. Therefore, they can also be classed as both. For clarification, we focus on the term 'accountant' when we refer to the study participants when reporting the findings.

4.3 Data analysis

Thematic analysis was conducted to identify, examine, and record patterns of experiences and explain the phenomena (Cairns and Cox 2008). Firstly, all the transcriptions were carefully read and re-read, followed by descriptive, sentence-level open coding. These codes were summarized into 14 categories, including e.g. system benefits, uncertainties, control, social influence, and the notion of expertise. We identified three main themes from these codes: 1) experiences on technology trust, 2) perceptions of domain experts as trustees, and 3) the social context. The analysis was conducted by the first author utilizing Atlas.ti software and iteratively discussed, challenged, and refined by the co-authors, including early- mid-, and late-career researchers. Atlas.ti is extensively used in data analysis in qualitative research, particularly to label the raw data with descriptive codes and to recognize new viewpoints and connections between different research participants (Hwang 2008). Our research follows a constructivist-interpretive paradigm in which coding refers to an inductive interpretation (McDonald et al. 2019; Blandford et al. 2016).

5 Findings

This study aims to gain an overall understanding of trust in IA during its deployment process, as well as the social dynamics influencing the use and acceptance of IA in this socio-technical environment. Three main themes emerged from the analysis: 1) aspects of trust in IA, 2) domain experts trustor characteristics, and 3) social aspects of trust. These findings present how accountants' trust in IA is formed in the context of accounting and especially in specific tasks of invoice processing. We label the participants according to their system use based on three categories: (a) indicates the system use of less than six months, (b) indicates system use of less than two years, and (c) indicates system use of more than two years.

5.1 Personal experiences shaping trust in intelligent automation

5.1.1 Trust is strengthened by personal experiences of reliability and functionality

Expectations are an essential element of trust: the trusted party is expected to behave in a certain manner. This notion was visible among the study participants.

When asked participants' perceptions of trust, most of them underlined technology's reliability: technology is considered trustworthy if it operates correctly and with its intended function (P1a, P2b, P3c, P4a, P5b, P6c, P7b, P8c). P3c exemplified: "I can trust it, if it operates as it is intended to." Expectations guide the users' perceptions of trust, and how well technology meets these expectations affects the sense of reliability and, therefore, trust: "That it works as expected. Then I can trust it" (P4a). This is a common characteristic of trust in technology and does not seem to differ in the case of IA. In our study, the perception of functionality was visible in the participant's descriptions of the systems' capability to do what it is supposed to do, i.e., predict and automate invoice handling (P1a, P2b, P5b, P6c, P7b) as P5b underlined: "It can interpret the invoices correctly and with right sums, for instance, what comes to value added taxing."

The term intelligent in this study refers to the systems' learning capabilities, which set specific characteristics concerning the system's use and trust. Because learning systems behavior varies depending on the situation, this causes uncertainty in invoice processing among the study participants. Thus, they underline a need to have a personal verification of the system's reliable operation (P3c, P4a, P5b, P6c, P7b, P8c). P8c underlined the importance of having personal experience: "You have to see yourself that it works before you can trust it 100%." The evaluation was emphasized at the beginning of the deployment process among all the participants, regardless of how long they had used the system, and it was admitted to diminishing over time. P5b explained the process in the following way:

> Well, at least I evaluate the system in the beginning. When I see that it has done automatic invoice processing, I will go and check the invoices; what invoices has it done and are these invoices such that I would allow it to process.

The participants did not consider evaluating the system results very time-consuming. Instead, it was expected to have time-saving benefits regarding the overall invoice processing because the possible system errors were easier to fix immediately rather than wait for the monthly report. P3 has used the intelligent system already for two years but admitted to still double-checking the system results, underlining a need for spot checks: "Even though all [invoices] would go through the system with 100% confidence, I need to make spot checks, to make sure it is all correct."

It is noteworthy that the user interface includes an accuracy percentage, indicating the systems' confidence in automating the individual invoices. Whereas a few participants did not consider this percentage very important, stating that they would check the invoices anyway (P6c, P9c), most participants admitted paying attention to it (P2b, P3, P4a, P5b, P7b. For instance, P7b perceived it as an additional information source which was good to consider but not enough to gain a

complete assurance of the accurate results—system evaluation and monitoring were still needed: "If it approaches 100%, then the system monitoring and checking will decrease. But it depends on who is sending the invoices and from whom. You can't tell directly; the percentage is only one part of the invoice processing." Participants acknowledged that the low percentage indicated a need to pay more attention to the invoice and tried to improve it by training the system. Interestingly, participants' perceptions of the accuracy percentage reflected their hesitation towards system outcomes, as P3 exemplified: "Even though it [the percentage] is 100%, I trust it 90%." P8c admitted that even though the percentage was high, they would check the invoices anyway because of the 'routine.' This was acknowledged as situation-specific; for instance, few participants (P1a, P2b, P3c) considered automation more reliable with repetitive and predictable invoices. Some participants also admitted having uncertainty regarding the systems' learning process, for instance, how long the learning takes (P3c) or if the accountant loses visibility into a certain invoice after full automation (P1a).

5.1.2 Realistic expectations on system capabilities lead to appropriate trust

Although participants underlined the system reliability and functionality as an important antecedent for trust, they were also considerably tolerant towards system errors. The participants' responses underlined comparability to man-made outcomes. For instance, some participants indicated acceptance of system errors because 'both humans and machines make mistakes' (P1a, P6c, P7b). P6c underlined that the system outcomes need to be comparable to man-made, explaining that "machine needs to be able to perform as a human would, at least – if the outcomes are similar, it does not matter if the invoice processing was proceeded by an accountant or an intelligent system." P7b acknowledged the possible trade-off in the system use, explaining that system errors need to be situated in the overall picture: "Even though there might be errors, the system benefits should overcome the possible minor errors." When discussing experienced or imaginary error situations, the participants underlined a need to understand and explain its reasons and the possibility to fix it (P1a, P2b, P4a, P5b). In most cases, the error situations were expected to be solved together with the developers (P2b, P4a, P5b, P7b, P8c, P9c) as P2b exemplified:

> Firstly, I would figure out why the error happened. We would investigate the error and let [the developing company] know about it, and they would investigate and fix it – or even though I would fix it, they need to know about the error.

System errors might temporarily increase accountants' control over the automated invoice processing, for instance, by decreasing the level of automation (P5b, P6c) or by increasing the system monitoring (P2b, P4a, P6c, P8c, P9c).

Some participants (P3c, P6c, P7b, P8c) admitted that major or continuous system errors would violate their trust in the system. (P3c) explained: "If the system makes constant errors, for instance, interprets the invoices incorrectly and produces the wrong total – if there were many such cases, then, of course, it would diminish the trust."

Appropriate trust, not over-trusting or distrusting the system, was emphasized in our data. In general, participants were satisfied with the system and considered it useful and reliable. Still, there was a certain amount of hesitation in trusting the system completely. Participants (P1a, P3c, P4a, P5b, P7b, P8c) acknowledged that they should be cautious of not trusting the system excessively because invoice processing includes a lot of contextual variety and client-specific information. P4a underlined the situational nature of trust: "I do not always trust it 100% – sometimes I need to fix something later." This might align with their overall attitude toward assistive systems, as P7b exemplified: "I do not trust any of these kinds of products 100%, but I do trust these systems when I can supervise them." Trust was considered to increase gradually, and many participants acknowledged that trust in the system would increase during use (P1a, P2b, P3c, P4a, P5b, P6c, P8c, P9c). They acknowledged that the more they use the system, the more it learns and the easier it is for the accountants to evaluate the systems' outcomes and build trust in the system. P6c exemplified this, emphasizing the collaboration with the "we" pronoun: "The more I have used the system, the more I can trust it – I have seen that we have achieved good outcomes. That is how the trust is built." The IA system is developed to save accountants time in invoice processing, and few participants admit that recognizing this increased their trust in the system (P6c, P7b, P9c). Some participants (P3c, P7b, P8c) acknowledged that learning systems would take some time before the benefits become visible. "When you have the patience to push it in the right direction, eventually it will learn to produce correct outcomes" (P8c), further underlining the participant's realistic expectations towards the system and their patience as system users.

5.2 Domain experts as AI end-users

5.2.1 Accounting expertise guides the system use and a user's capacity to monitor the system's outcomes

The findings indicate that the participants set certain requirements for themselves as technology end-users. This might be explained by the systems' learning capabilities: the participants train the system through their own actions (invoice processing). Because they work with different clients, the system uses and outcomes are tailored and personalized, resulting in a feeling of accountability. The participants acknowledged having the responsibility of training the system (P3c, P4a, P6c, P7b, P8c, P9c) to use the system correctly without mistakes (P3c, P6c, P7b) and auditing and fixing the possible system errors (P3c, P4a, P5b, P6c, P9c). P7b exemplified this notion, emphasizing the divided responsibility between users

and system developers: "Ultimately, I have the responsibility. Of course, if the system does not work as it should, it is the developers' responsibility. But incorrect accounting is not their responsibility; that is on me." The system may also learn erroneous patterns or work practices. This can be caused by technical errors such as system updates or man-made mistakes, for instance, if an accountant has entered incorrect taxation into the invoice. Some participants also felt accountable for the system errors, which were considered to negatively affect client relationships (P1a, P2b, P3c, P5b) as P5b explained: I cannot tell the client that this mistake was made by a robot. The mistake is on me and reduces the clients' trust in me."

Expertise in accounting was considered essential in using the system (P1a, P2b, P4a, P7b, P8c), and it was expected to guide the evaluation of the system outcomes. P2b exemplified: "If you do not have professional knowledge and if you do not know what the correct outcome is, you will let all the system suggestions go through." According to P7b, the knowledge of accounting helps the user to guide their focus on essential information when using the system, such as "I need this particular information; where I can find and modify that information." Accounting expertise was considered important to correct the possible errors in intelligent invoice processing, as P4a stated: "You need to know what to fix." Participants also acknowledge the importance of tacit knowledge, the general know-how as an accountant that is not verbalized or visible in the invoice (P4a, P5b, P7b, P9c). Such examples are exceptions in taxation, customers' fringe benefits, or even misspellings in invoices. Participants acknowledged that the system functionality varies depending on the situation; for instance, predictable and repetitive invoices were considered easier to automate (P2b, P1a, P3c). This information was considered important when deciding the level of automation, such as what invoices are possible to automate and what required a situation or context-specific expert knowledge.

5.2.2 Trusting the system necessitates context- and client-specific expert knowledge

The operation of IA is based on the accountant's work as the ML system learns while accountants process the invoices. The more repetitive processes increase the system's confidence, and unless the accountants correct the automated invoices, the system suggests full automation. Nevertheless, some exceptions require specific expert knowledge because differences in taxation depend on the context where the clients operate (P1a, P2b, P7b, P9c): for instance, taxation in agriculture differs from construction business (P7b). Client-specific knowledge was considered essential among participants (P3c, P5b, P7b, P8c, P9c). P9c explained their experiences: "We know our clients well enough to assess if their invoices can or cannot be automated." This context-specific knowledge also affects the accountants' trust in the system, as P7b underlined: "Trust is not

related to the system, it is client specific – with some clients I have declined it [automation] because I know how this client works and I know their invoices cannot be automated." New clients were admitted requiring more monitoring in comparison to familiar clients, as P3c exemplified: "You cannot trust that it is correct even though the machine says so. To be sure, you should know your clients as well." This is related, for instance, to knowing the clients' products (P7b), needs for specific or unusual taxation (P2b), differences in the invoice cycle (P1a), practices regarding payment reminders (P4a), and client-specific requests (P5b). By processing the invoices manually, the accountants can gain cumulative insight into their client's operational environment, aiding in the accumulation of the accountant's knowledge. This knowledge guides system uses and creates a base for the accountant's evaluation of the system's functionality and reliability, i.e., how well they can trust the system to work as expected.

Interestingly, the requirement for client- and context-specific knowledge is reflected also in participants' interaction with their colleagues: for instance, during holiday periods, the participants might need to transfer their clients' invoice processing to their colleagues who were considered to lack this specific client knowledge. P9c pondered their level of trust in their colleagues versus the intelligent system and the systems' capacity to guide their colleagues appropriately:

> If there is a company that my colleague has never done and this colleague should take care of my clients during my holiday, then the system could suggest and recommend the correct invoice processing. In this case, I would probably trust the system more [than the colleague].

The quote above underlines the participants' personal experience and re-assurance of the system's usefulness and reliability concerning their clientele: even within the same organization, accountants work with different clients and with different taxation contexts. Also, some participants (P1a, P2b, P5b, P8c) admitted trusting the system in a similar manner as trusting their colleagues, with both having their benefits, as P5b exemplified: "The system remembers what it has been taught, but my colleague can forget something important – on the other hand, my colleague can also use their own brains in tricky situations."

5.3 Social aspects of trust

5.3.1 Collaborative deployment process increases end-users trust in intelligent automation

The interview data emphasized that the iterative nature of the development was considered positive among study participants. The early phases of system deployment were especially characterized by a collaborative partnership between the accountants and the developers, wherein the system underwent continual refinement and enhancement. The developers trained the accountants to use the system,

for instance, by educating them about the system's capabilities (P2b, P3c, P5b). However, this process was considered reciprocal, and accountants had an active role in informing the developers about their expert domain, thereby contributing to the developers' understanding of accounting (P2b, P3c, P5b, P8c). P5b explained: "They are not accountants, so they cannot see this from our perspective: what is important and what is not. But they also can tell us about intelligent automation to help us understand better what it can or cannot do." This was expected to benefit product development and to serve the accountants even better, as P8c underlined:

> If there is something special, like a specific taxation. I try to explain it to the developers in a way they understand it. That will help them to develop the product further, and then the product will serve us in the best possible way.

Nevertheless, some participants also expected the developers to have at least a bit of experience or expertise in accounting or first-hand knowledge of their work practices (P2b, P5b, P8c). P2b explained:

> In the long run, it is their strength [developers] that they know something about their clients – for instance, many IT professionals does not have a clue, how farms operate. We have given them suggestions on how to develop this and that. At least, they need to scratch the surface a bit.

A common view amongst the interviewees was that the developers have been willing to listen to the domain experts' suggestions and feedback regarding the system (P1a, P2b, P3c, P5b, P6c, P7b, P8c, P9c) and that they had an interest to develop the system further to serve the accountants' needs (P2b, P3c, P5b, P6c). P3c exemplified: "In my opinion, they have taken seriously the suggestions I, and others, have given them." This gave accountants a sense of a collaborative development process, as P6c explained: "It is clear that we have built this together. We have given suggestions, and they have always been very receptive and acted on our feedback." The participants acknowledged that developers have been very helpful and responded quickly to a possible problem situation (P1a, P2b, P3c, P4a, P8c) and the customer service has been good (P5b, P7b, P8c). Some participants underlined the importance of the developers' ability (e.g., how competent they are) and that they will fulfill their promises (P2b, P5b, P6c, P8c). P5b exemplified this attitude: "If we agree on something, and it is actualized, that increases trust – and if I have a feeling that this person [developer] knows what they are doing, that also increases trust." In addition, the developers' honesty about the systems' learning process was considered positive among the participants (P1a, P2b, P3c, P5b, P8c), as P2b commented: "They [developers] told us that it takes

some time before the system learns – we knew that there will be this learning period and it is part of the process." The positive experiences regarding the interaction with the developers were considered to increase participants' trust in the system (P3c, P5b, P6c, P7b, P8c, P9c) as P5b's comment above exemplifies. It seems that this attitude is emphasized at the beginning of the deployment process. For instance, P9c started using the system after it had already established its place in an organization. In their responses, the friendly interaction with the developers was not underlined. On the other hand, there might be a contrary situation: P3c considered that trust in the system is also built on the expectation that the developers want to improve the system: "If there is a developer who never asks the end-user about the product, that would decrease trust."

5.3.2 The relevance of social trust and personal experience of technology

It is noteworthy that most of the study participants consider themselves early adopters, having a genuine interest and receptive attitude to novel technologies and a willingness to learn about novel systems (P1a, P3c, P4a, P5b, P6c, P7b, P9c). The early adopters considered that some of their colleagues might be hesitant towards novel systems, for instance, because of fear (P5b) or established work practices and negative attitudes (P8c). A few participants also admitted that digitalization requires them to keep up with the novel technologies because "the change is inevitable" (P5b), and "you have to evolve; there really isn't another option" (P2b). P5b and P6c emphasized that real user stories or user references would probably increase hesitant users' trust in the system, further underlining the entanglement between social trust and trust in technology. The system providers also acknowledged this. For instance, they had organized a practice in which one of the accountants inside an organization was selected as an AI ambassador with the aim of encouraging and teaching their colleagues to use the system. P8c is such a corresponding person in their organization and explained this aim: "We have set the goal to expand the system use in our organization and spread the message that this is really a good and helpful tool – sometimes it is a bit challenging." According to P8c, the main challenges are related to negative attitudes toward the system, especially among the older generations who are not willing to change their established work practices. In addition, the time and patience that is required to gain the system benefits can sometimes be challenging to justify for their colleagues, especially those with more resistant attitudes.

6 Discussion

This study investigated accountants' trust in IA in their work, focusing on invoice processing. The study findings highlight the situational and context-dependent nature of trust in IA. For instance, a collaborative design process and open communication with the developers is considered to influence trust in IA. Most

interestingly, the findings acknowledge domain experts' specific trustor characteristics as AI end-users: the studied IA system develops and personalizes through the system use, and the accountants play a vital role in training the system by transferring their expert knowledge to the system, enabling the system to identify repetitive patterns, and suggesting automation possibilities. By leveraging their expert knowledge, accountants evaluate the system outcomes, influencing the iterative development and customization of IA. Next, we discuss the study findings considering prior research, starting from the social aspects, and continuing to human-AI interaction. Lastly, we cover the specific characteristics of domain experts as technology users.

6.1 Interpersonal and social trust shaping trust in intelligent automation

The concept of appropriate trust underlines the situational and context-dependent nature of trust, further emphasizing that trust in IA is not solely focused on technology and its characteristics but the whole sociotechnical system. This aligns with CSCW research in which the social and the technical elements should be designed in parallel because they influence each other (Bullinger-Hoffman et al. 2021). Our study findings underline the importance of a collaborative and iterative design process in relation to trust: domain experts are actively involved in the design and development of the product to ensure that the target activities and social context are properly understood by the system developers. They also feel respected in what they do and have a feeling of having a voice in the development and deployment process. Such articulation of the work is necessary to enhance understanding and awareness between industry practitioners and domain experts because it creates visibility to the work practices and the task at hand (Suchman 1995; Schmidt and Bannon 1992). This resulted in a more usable and useful end-product and a receptive and positive attitude towards the system – thus supporting trust in the system's IA features.

Attitude is considered an important aspect of trust in automation (Lee and See 2004), which is supported by the present study. Although being an individual trait, it can be supported socially: information occurring in a wider social environment, e.g., in indirect social relationships, can shape trusting attitudes and behavior. This can include, e.g., interaction with service providers and the use of complementary technologies (Bachmann 2001; Kosonen et al. 2008; Shapiro 1987; van der Werff et al. 2021.) In addition, the study participants expressed positive sentiments towards the developers and their willingness to listen to the users' and their suggestions, indicating developers' benevolence and ability to create helpful products. This underlines the importance of social trust and interpersonal trust (Mayer et al. 1995) and can also be acknowledged in social practices inside an organization, with more active and receptive users supporting their colleagues to use and trial the system. This relatively strong social trust between

the actors offered a fruitful starting point for the experts trusting even a new type of technology with a high level of automation.

6.2 Aspects of trust in human-AI interaction

Research on trust in technology underlines functionality, reliability, and helpfulness as the main factors in trust in technology (Lankton and McKnight 2011; Muir and Moray 1996). These characteristics are considered to enable users to trust the system, which is understandable—without them, the technology would be useless. This can be considered obvious regarding novel technologies, including AI-based applications. For instance, Bingley et al. (2023) explored user beliefs about AI, and functionality was the most mentioned theme for laypeople. Functionality and reliability were also present in our study findings and are clearly essential to accountants' trust in IA.

What is interesting about our study findings is the emphasis on domain experts' personal experience with reliability. The accountants underlined a need to evaluate, monitor, and double-check the system operations to have first-hand experience of its reliability. This helps the users assess the perceived ability of the system to conduct a given task, which is considered to influence trust in AI (Saßmannshausen et al. 2021). Such examination can also serve the users' need for AI transparency and explainability (Glikson and Woolley 2020; Bedué and Fritzsche 2022; Lockey et al. 2021), helping them to understand the situations in which ML systems can be considered trustworthy. Our study findings underline the relevance of assessing the appropriate use of IA for a given situation. This is especially important regarding the autonomous and adaptive nature of novel AI-based applications: if rule-based systems can be expected to operate in a predictable manner, this might not apply to continuously updating machine-learning systems.

According to Hoff and Bashir (2015), trust in automation is situational and context-dependent, and the decision to trust becomes visible per the situation or task at hand. In our study, this situatedness is constituted in the varied clientele and the accountant's knowledge of their client's operational environment. It is noteworthy that trust in IA comes with hesitation: trust in the system should be calibrated appropriately (Parasuraman and Riley 1997; Lee and See 2004), and the study participants do not seem comfortable with relying on the system too much. Appropriate trust refers to a situation in which the user's trust and the technology's capabilities are aligned, and the user knows when to trust the technology, avoiding over-trusting and distrusting. To gain appropriate trust in AI, users should be able to utilize their knowledge to improve the outcomes in situations where AI models may have limitations (Zhang et al. 2020a, b). This was also visible in our findings: expert knowledge guided the trust calibration, helping the users to decide when to (or not) trust the systems' suggestion.

6.3 Trustor characteristics: domain experts as AI users

In addition to technological characteristics in trust, our findings reveal interesting trustor (user) characteristics. For instance, certain individual traits, such as users' propensity or tendency to trust (e.g., Mayer et al. 1995; van der Werff et al. 2021; Saßmannshausen et al. 2021), influence trust in technology, including AI. Our findings support the idea that trust in novel technology is a cumulative process built on previous experiences. However, while *digital affinity* (i.e., interest and pleasure in learning and using digital technologies) has been found to influence domain experts' trust in AI (Saßmannshausen et al. 2021), this is probably a common characteristic for all techno-savvy users who are more prone to use (and trust) novel technologies than those with a more hesitant attitude. To extend these prior findings, our study reveals a few additional aspects of trust in IA.

First, the importance of expert knowledge is emphasized in our study. The use of IA intertwines with general accounting knowledge and an understanding of their client's specific needs, intricacies, and variations. To be able to appropriately monitor the system and permit higher levels of automation, they need to have first-hand knowledge of their clients' operations. This is very case-specific and refers to tailored use of novel AI applications: within the same organization, the use of IA might vary a lot. This also explains a need for personal reassurance regarding the system functionality and becomes visible also in participants' comparison between IA and a colleague: whereas a colleague might have more knowledge of accounting, they lack knowledge of a specific clientele. This is an interesting observation, showing that the notion of expertise is not related only to explicit knowledge but also tacit knowledge (Rinta-Kahila et al. 2018). Whereas previous research suggests that expert status (e.g., competence, skills, and experience) is not significant in trust in AI (Saßmannshausen et al. 2021), our study clearly underlines its relevance: the use of IA is based on domain expertise. This can probably be explained by the differences in technical characteristics of these studies: we chose to study trust in relation to a certain AI-based technology in a specific professional context instead of people's "general" perceptions of AI. Second, domain experts' role in their organization and in relation to their clients influences their trust in technology. Because the use of IA is client-specific, they are accountable for the system use and its possible consequences. This observation is aligned with previous research, in which the domain experts need to have the ability to understand, explain, and justify AI decisions to other stakeholders (Lockey et al. 2021). Work systems are considered to consist of people, technology, organizations, and the primary tasks (e.g., Bullinger-Hoffmann et al. 2021), underlining a holistic perspective when developing AI systems for knowledge work. Also, the perception of AI systems' trustworthiness covers work systems comprising both AI and humans and their interaction and collaboration in certain contexts and tasks. Therefore, the question of trust in AI might not focus on *whether* AI is trustworthy but on *how* AI could be used in a trustworthy manner,

which refers to both technical and trustor characteristics and the design and development of such systems.

6.4 Limitations and future research

While this study provides insights into domain experts' trust in intelligent automation, it is important to acknowledge certain limitations. The sample size used in this research is relatively small, which means that additional aspects of trust and subjective aspects may be found in further studies. However, this sample size is quite usual for qualitative interview studies. Also, the participation in the research interviews was voluntary, and hence, the interviewees represent people who have an initial interest in technology, therefore possibly having a propensity to trust novel technology solutions. Lastly, the study did not delve into the long-term effects or potential consequences of trust or mistrust in these systems. By addressing these limitations in future research, we can further advance our understanding of trust in intelligent automation and contribute to the development of trustworthy AI systems in various domains.

It is important to recognize that focusing on a particular domain, such as accountancy, presents unique challenges regarding the generalizability and transferability of findings to other domains. For example, our research indicates that accuracy and double-checking hold significant importance in the work of accountants, values they are reluctant to compromise with the introduction of novel AI technologies. However, such preferences may not be applicable across all knowledge-intensive fields in which AI is embedded. The main focus should be on understanding the domain-specific nuances and characteristics: a genuine interest in expert work in certain professional contexts and the end-users'preferences regarding their work practices (i.e., what they find valuable and what tasks they are willing to transfer to intelligent systems). Whereas we cannot expect the system developers to tailor each system in accordance with domain-specific nuances, it is essential to respect these nuances, including their respective goals, specific aims, and domain-specific values, when designing AI systems. Domain-specific know-how will always be part of expertise. It is noteworthy that these characteristics might also be invisible (Suchman 1995) and that possible organizational values or goal conflicts should be acknowledged during the design and development phase. This is also important considering the increasing agency of AI end-users in, e.g., system training: as domain experts are expected to transfer their expertise into AI systems, system developers might become more dependent on them. Thus, approaching this specific user group with respect is beneficial regarding the system use and acceptance, as our findings suggest. Furthermore, the concept of trust and its formation between domain experts and their clients may vary fundamentally across different professions. For instance, doctors may inherently enjoy a position of trust built upon their extensive education, whereas professionals in other domains, such as real estate, may need to

cultivate trust among their clients actively (Ala-Luopa et al. forthcoming). Novel AI applications can either be designed to support or diminish expertise, and these decisions (and embedded values) have both societal and domain-specific consequences that designers and other stakeholders should be aware of.

In addition to awareness and respect related to domain-specific nuances and design guidelines for AI in expert work, this study also offers other intriguing avenues for future research. We call for more research on domain experts' trustor characteristics, for instance, those with rejective attitudes on novel technologies. In addition, the perspectives of other stakeholders regarding domain experts' use of novel AI applications require also further exploration. This can refer to indirect trust relations, for instance, clients' attitudes towards AI-supported work practices that concern their businesses. In general, more research is needed to explore domain experts' interactions with specific AI applications, underlining the importance of research situated in real-life contexts, which oftentimes is more nuanced and complicated than we might anticipate.

7 Conclusion

The study underlines the context-specific, situated nature of trust in intelligent automation (IA) and reveals the importance of personal oversight and social aspects, such as collaborative design practices. The system use is based on expert knowledge, including tacit knowledge, which makes both the system use and trust in IA highly individual. Trust in IA in our study manifests at different levels: in human-AI interaction, in interpersonal communication and collaboration with the technology provider, and in acknowledging the domain experts' trust characteristics. The findings expand previous research on trust in technology by revealing new observations on trust in knowledge work, underlining variance in the sociotechnical phenomenon of trust.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s10606-024-09499-6.

Acknowledgments
We would like to express our gratitude to Dr *Lisa van der Werff*, Associate Professor at the Dublin City University Business School, Dr *Steve Lockey,* Postdoctoral Research Fellow at the University of Queensland Business School, and the FINT community (First International Network on Trust) for their valuable comments on earlier versions of this research article.

Author contributions All of the authors whose names appear on the submission contributed to the work: The interview questions were designed by the main author (S.A-L.) and discussed and modified together with all the authors. Six of the interviews were conducted by S.A-L., 2 interviews by M.H., and 1 interview by J.M.

The interviews were transcribed and analyzed by S.A-L., and the findings were discussed, challenged, and refined with all the authors. S.A-L. wrote the manuscript, which was then commented on and revised critically by T.O., and K.V. All the authors read the manuscript and approved the version to be published.

Funding Open access funding provided by Tampere University (including Tampere University Hospital). This research was conducted in the KITE project, funded by the European Regional Development Fund, Business Tampere, and the University of Tampere (A75453).

Data availability No datasets were generated or analysed during the current study.

Declarations

Ethical approval Ethical guidelines by the Finnish advisory board on research integrity (TENK), appointed by the Ministry of Education and Culture in Finland, were followed throughout the research. These guidelines cover integrity, meticulousness, and accuracy in conducting research and in recording, presenting, and evaluating the research results.

Competing interests The authors declare no competing interests.

References

Ackerman, Mark S. 2000. The intellectual challenge of CSCW: The gap between social requirements and technical feasibility. *Human-Computer Interaction* 15: 179–203.

Asatiani, Aleksandre, Esko Penttinen, Joona Ruissalo, and Antti Salovaara. 2020. Knowledge workers' reactions to a planned introduction of robotic process automation—empirical evidence from an accounting firm. In *Information systems outsourcing. Progress in IS*, eds. R. Hirschheim, A. Heinzl, and J. Dibbern. Cham: Springer. https://doi.org/10.1007/978-3-030-45819-5_17.

Ashoori, Maryam, and Justin Weisz. 2019. In AI We Trust? *Factors That Influence Trustworthiness of AI-infused Decision-Making Processes*. arXiv:1912.02675. https://doi.org/10.48550/arXiv.1912.02675.

Auernhammer, Jan. 2020. Human-centered AI: The role of Human-centered Design Research in the development of AI. In *Synergy - DRS International Conference*, 11–14 August 2020 eds. S. Boess, M. Cheung, and R. Cain, 11–14. https://doi.org/10.21606/drs.2020.282.

Bachmann, Reinhard. 2001. Trust, power and control in trans-organizational relations. *Organization Studies* 22 (2): 337–365. https://doi.org/10.1177/0170840601222007.

Bedué, Patrick, and Abrecht Fritzsche. 2022. Can we trust AI? An empirical investigation of trust requirements and guide to successful AI adoption. *Journal of Enterprise Information Management* 35 (2): 530–549. https://doi.org/10.1108/JEIM-06-2020-0233.

Benbasat, Izak, and Weiquan Wang. 2005. Trust in and adoption of online recommendation agents. *Journal of the Association for Information Systems* 6(3). https://doi.org/10.17705/1jais.00065.

Bingley, William J., Caitlin Curtis, Steven Lockey, Alina Bialkowski, S. Nicole Gillespie, Haslam Alexander, Ryan K. L. Ko, Niklas Steffens, Janet Wiles, and Peter Worthy. 2023. Where is the human in human-centered AI? Insights from developer priorities and user experiences. *Computers in Human Behavior* 141: 107617. https://doi.org/10.1016/j.chb.2022.107617.

Blandford, Ann, Dominic Furniss, and Stephann Makri. 2016. Qualitative HCI research: Going behind the scenes. *Synthesis Lectures on Human-Centered Informatics* 9 (1): 1–115.

Brynjolfsson, Erik, and Andrew McAfee. 2011. *Race against the machine: How the digital revolution is accelerating innovation, driving productivity, and irreversibly transforming employment and the economy*. Lexington: Digital Frontier Press.

Brynjolfsson, Erik, and Andrew McAfee. 2017. The Business of Artificial Intelligence: what it can and cannot do for your organization. Harvard Business Review Digital Articles. https://hbr.org/2017/07/the-business-of-artificial-intelligence. Accessed 14 November 2023.

Bullinger-Hoffmann, Angelika, Michael Koch, Kathrin Möslein, and Alexander Richter. 2021. Computer-Supported Cooperative Work – Revisited. *i-com* 20 (3): 215–228. https://doi.org/10.1515/icom-2021-0028.

Cairns, Paul, and Anna L. Cox. 2008. *Research methods for human-computer interaction*. Cambridge University Press.

Corritore, Cynthia. L., Susan Wiedenbeck, and Beverly Kracher. 2001. The elements of online trust. In *CHI '01 Extended Abstracts on Human Factors in Computing Systems (CHI EA '01)*, 504–505. New York: Association for Computing Machinery.https://doi.org/10.1145/634067.634355.

Corritore, Cynthia, L., Beverly Kracher, Susan Wiedenbeck. 2003. On-line trust: Concepts, evolving themes, a model. *International Journal of Human-Computer Studies*, 58 (6): 737–758. https://doi.org/10.1016/S1071-5819(03)00041-7.

de Visser, Ewart, Richard J. Pak, and Tyler H. Shaw. 2018. From 'automation' to 'autonomy': The importance of trust repair in human–machine interaction. *Ergonomics* 61 (10): 1409–1427. https://doi.org/10.1080/00140139.2018.1457725.

Dourish, Paul. 2003. The appropriation of interactive technologies: Some lessons from placeless documents. *Computer Supported Cooperative Work* 12 (4): 465–490. https://doi.org/10.1023/A:1026149119426.

Dourish, Paul, and Victoria Bellotti. 1992. Awareness and coordination in shared workspaces. In *Proceedings of the 4th ACM Conference on Computer-Supported Cooperative Work (CSCW'92)*, eds. J. Turner and R. Kraut, 107–114. ACM Press.

Ericsson, Anders K. 2014. *Expertise. Current Biology*. Cell Press. https://doi.org/10.1016/j.cub.2014.04.013.

Frey, Carl Benedikt, and Michael A. Osborne. 2017. The future of employment: How susceptible are jobs to computerisation? *Technological Forecasting and Social Change* 114: 254–280. https://doi.org/10.1016/j.techfore.2016.08.019.

Gefen, David, Elena Karahanna, and W. Detmar Straub. 2003. Trust and TAM in online shopping: AN integrated model. *MIS Quarterly: Management Information Systems* 27 (1): 51–90. https://doi.org/10.2307/30036519.

Glikson, Ella, and Anita Williams Woolley. 2020. Human trust in artificial intelligence: Review of empirical research. *Academy of Management Annals* 14 (2): 627–660. https://doi.org/10.5465/annals.2018.0057.

Hartikainen, Maria, Kaisa Väänänen, Anu Lehtiö, Saara Ala-Luopa, and Thomas Olsson. 2022. Human-centered AI design in reality: A study of developer companies' practices. *Nordic Human-Computer Interaction Conference (NordiCHI '22)*, 08–12 October 2022, Aarhus, Denmark, 1-11. New York: ACM. https://doi.org/10.1145/3546155.3546677.

HLEG AI. 2019. Ethics guidelines for trustworthy AI. Retrieved from High-Level Expert Group on Artificial Intelligence. https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai. Accessed 14 November 2023.

Hoff, Kevin Anthony, and Masooda Bashir. 2015. Trust in Automation: Integrating Empirical Evidence on Factors That Influence Trust. *Human Factors* 57 (3): 407–434. https://doi.org/10.1177/0018720814547570.

Hoffman, Robert R. 1998. How can expertise be defined?: Implications of research from cognitive psychology. In *Exploring expertise*, eds. R. Williams, W. Faulkner, and J. Fleck, 81–100. New York: Macmillan.

Hwang, Sungsoo. 2008. Utilizing qualitative data analysis software: A review of atlas.ti. *Social Science Computer Review* 26 (4): 519–527. https://doi.org/10.1177/0894439307312485.

Jacovi, Alon, Ana Marasović, Tim Miller, and Yoav Goldberg. 2021. Formalizing trust in artificial intelligence: Prerequisites, causes, and goals of human trust in AI. *FAccT 2021 - Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency,* 624–635. Association for Computing Machinery, Inc. https://doi.org/10.1145/3442188.3445923.

Jarrahi, Mohammad Hossein. 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.

Jarvenpaa, Sirkka L., Kathleen Knoll, and Dorothy E. Leidner. 1998. Is anybody out there? Antecedents of trust in global virtual teams. *Journal of Management Information Systems* 14(4): 29–64. http://www.jstor.org/stable/40398291.

Komiak, Sherrie Y. X., and Izak Benbasat. 2006. The effects of personalization and familiarity on trust and adoption of recommendation agents. *MIS Quarterly* 30(4): 941–60. https://doi.org/10.2307/25148760.

Kosonen, Miia, Kirsimarja Blomqvist, and Riikka Ellonen. 2008. Trust and its impersonal nature. *Encyclopedia of Networked and Virtual Organizations,* 1683–1690. IGI Global. https://doi.org/10.4018/978-1-59904-885-7.ch222.

Kuzel, A. J. 1992. Sampling in qualitative inquiry. In *Doing qualitative research*, eds. B. Crabtree, and W. Miller, 31–44. Newbury Park: Sage.

Lankton, Nancy K., and Harrison D. McKnight. 2011. What does it mean to trust Facebook? Examining technology and interpersonal trust beliefs. *Data Base for Advances in Information Systems* 42 (2): 32–54. https://doi.org/10.1145/1989098.1989101.

Lee, John, and Neville Moray. 1992. Trust, control strategies, and allocation of function in human-machine systems. *Ergonomics* 35 (10): 1243–1270. https://doi.org/10.1080/00140139208967392.

Lee, John D., and Katrina A. See. 2004. Trust in automation: Designing for appropriate reliance. *Human Factors* 46 (1): 50–80. https://doi.org/10.1518/hfes.46.1.50_30392.

Leitner-Hanetseder, Susanne, Othmar Lehner, Cristoph M. Eisl, and Carina Forstenlechner. 2021. A profession in transition: actors, tasks, and roles in AI-based accounting. *Journal of Applied Accounting Research* 22 (3): 539–556. https://doi.org/10.1108/JAAR-10-2020-0201.

Lockey, Steven, Nicole Gillespie, Daniel Holm, and Ida Asadi Someh. 2021. A review of trust in artificial intelligence: Challenges, vulnerabilities, and future directions. *Proceedings of*

the Annual Hawaii International Conference on System Sciences. January 2020, 5463–5472. IEEE Computer Society. https://doi.org/10.24251/hicss.2021.664

Madhavan, P., and D. A. Wiegmann. 2007. Similarities and differences between human–human and human–automation trust: An integrative review. *Theoretical Issues in Ergonomics Science* 8 (4): 277–301. https://doi.org/10.1080/14639220500337708.

Mayer, Roger C., James H. Davis, and David F. Schoorman. 1995. An integrative model of organizational trust. *Academy of Management Review* 20: 709–734. https://doi.org/10.2307/258792.

McDonald, Nora, Sarita Schoenebeck, and Andrea Forte. 2019. Reliability and inter-rater reliability in qualitative research: norms and guidelines for CSCW and HCI Practice. *Proceedings of the ACM on Human-Computer Interaction*, 3(CSCW), Article 72. https://doi.org/10.1145/3359174.

McKnight, Harrison D., Larry L. Cummings, and Norman L. Chervany. 1998. Initial trust formation in new organizational relationships. *Academy of Management Review* 23 (3): 473–490. https://doi.org/10.2307/259290.

McKnight, Harrison D., Vivek Choudhury, and Charles ("Chuck") Kacmar. 2002. Developing and validating trust measures for e-commerce: An integrative typology. *Information Systems Research* 13:334–359. https://doi.org/10.1287/isre.133.3.34.81.

McKnight, Harrison D., Michelle Carter, Jason Benne Thatcher, and Paul F. Clay. 2011. Trust in a specific technology: An investigation of its components and measures. *ACM Transactions on Management Information Systems* 2 (2): 1–25. https://doi.org/10.1145/1985347.1985353.

Muir, Bonnie M., and Neville Moray. 1996. Trust in automation: II. Experimental studies of trust and human intervention in a process control simulation. *Ergonomics* 39 (3): 429–460. https://doi.org/10.1080/00140139608964474.

Mumford, Enid. 2000. A Socio-Technical Approach to Systems Design. *Requirements Engineering* 5: 125–133. https://doi.org/10.1007/PL00010345.

Orlikowski, Wanda J. 1992. Learning from Notes: Organizational Issues in Groupware Implementation. *Proc. ACM Conf. Computer-Supported Cooperative Work CSCW'92,* Toronto, Ontario. New York: ACM.

Orlikowski, Wanda J. 1995. Evolving with notes: organizational change around groupware technology. Working Paper 186, *Center for Coordination Science.* Cambridge: MIT.

Pakarinen, Pauli, and Ruthanne Huising. 2023. Relational expertise: what machines can't know. *Journal of Management Studies*. https://doi.org/10.1111/joms.12915.

Parasuraman, Raja, and Victor Riley. 1997. Humans and automation: use, misuse, disuse abuse. *Human Factors* 39 (2): 230–253. https://doi.org/10.1518/001872097778543886.

Pavlou, Paul A. 2003. Consumer acceptance of electronic commerce: Integrating trust and risk with the technology acceptance model. *International Journal of Electronic* 7 (3): 101–134.https://doi.org/10.1080/10864415.2003.11044275.

Pavlou, Paul A., and David Gefen. 2004. Building effective online marketplaces with institution-based trust. *Information Systems Research* 15 (1): 37–59. https://doi.org/10.1287/isre.1040.0015.

Pennanen, Kyösti., Tarja Tiainen, and Harri T. Luomala. 2007. A qualitative exploration of a consumer's value-based e-trust building process: A framework development. *Qualitative Market Research* 10 (1): 28–47. https://doi.org/10.1108/13522750710720387.

Rinta-Kahila, Tapani, Esko Penttinen, Antti Salovaara, and Wael Soliman. 2018. Consequences of discontinuing knowledge work automation: surfacing of deskilling effects and methods of recovery. *Proceedings of the 51st Hawaii International Conference on System Sciences (HICSS 2018),* 5244–5253. University of Hawai'i at Manoa. https://doi.org/10.24251/hicss.201.

Saßmannshausen, Till, Peter Burggräf, Johannes Wagner, Marc Hassenzahl, Thomas Heupel, and Fabian Steinberg. 2021. Trust in artificial intelligence within production management – An exploration of antecedents. *Ergonomics* 64 (10): 1333–1350. https://doi.org/10.1080/00140139.2021.1909755.

Saunders, Mark N. K. 2012. Choosing research participants. In *Qualitative organizational research: core methods and current challenges*, eds. G. Symon and C. Cassell, 35–52. London: Sage.

Schmidt, Kjeld, and Liam Bannon. 1992. Taking CSCW seriously. *Computer Supported Cooperative Work* 1: 7–40. https://doi.org/10.1007/BF00752449.

Schoorman, David F., Roger C. Mayer, and James H. Davis. 2007. An integrative model of organizational trust: Past, present and future. *Academy of Management Review* 32 (2): 344–354.

Shapiro, Susan P. 1987. The social control of impersonal trust. *American Journal of Sociology* 93: 623–658. https://doi.org/10.1086/228791.

Siau, Ken, and Weiyu Wang. 2018. Building trust in artificial intelligence, machine learning, and robotics. *Cutter Business Technology Journal*, 31(2), 47–53. https://www.cutter.com/article/building-trust-artificial-intelligence-machine-learning-and-robotics-498981. Accessed 14 November 2023.

Suchman, Lucy. 1995. Making work visible. *Communications of the ACM* 38 (9): 56–64. https://doi.org/10.1145/223248.223263.

Vereschak, Oleksandra, Gilles Bailly, and Baptiste Caramiaux. 2021. How to evaluate trust in ai-assisted decision making? A survey of empirical methodologies. *Proc. ACM Hum.-Comput. Interact*., 5, CSCW2, 327. https://doi.org/10.1145/3476068

Van der Werff, Lisa, Kirsimarja Blomqvist, and Sirpa Koskinen. 2021. Trust cues in artificial intelligence: A multilevel case study in a service organization. *Understanding Trust in Organizations: A Multilevel Perspective*. Routledge. https://doi.org/10.4324/9780429449185-13

Wang, Weiquan, and Izak Benbasat. 2008. Attributions of trust in decision support technologies: A study of recommendation agents for e-commerce. *Journal of Management Information Systems* 24 (4): 249–273.https://doi.org/10.2753/MIS0742-1222240410.

Yang, Qian, Aaron Steinfeld, Carolyn Rosé, and John Zimmerman. 2020. Re-examining whether, why, and how human-ai interaction is uniquely difficult to design. *Conference on Human Factors in Computing Systems - Proceedings*, 1–13. https://doi.org/10.1145/3313831.3376301

Zhang, Yunfeng, Liao Q. Vera, and Rachel K. E. Bellamy. 2020a. Effect of confidence and explanation on accuracy and trust calibration in AI-assisted decision making. *FAT 2020 - Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency,* 295–305. Association for Computing Machinery, Inc. https://doi.org/10.1145/3351095.3372852.

Zhang, Yingying, Feng Xiong, Yi Xie, Xuan Fan, and Haifeng Gu. 2020b. The impact of artificial intelligence and blockchain on the accounting profession. *IEEE Access* 8: 110461–110477. https://doi.org/10.1109/ACCESS.2020.3000505.

Zicari, Roberto V., John Broderser, James Brusseau, Boris Düdder, Timo Eichhorn, Todor Ivanov, Georgios Kararigas, Pedro Kringen, Melissa McCullough, Florian Möslein, Naveed Mushtaq, Gemma Roig, Norman Stürtz, Karsten Tolle, Jesmin Jahan Tithi, Irmhild van Halem, and Magnus Westerlund. 2021. Z-Inspection®: A process to assess trustworthy AI. *IEEE Transactions on Technology and Society* 2 (2): 83–97. https://doi.org/10.1109/TTS.2021.3066209.