AI Literacy - Towards Measuring Human Competency in Artificial Intelligence

Marc Pinski
Technical University of Darmstadt
pinski@ise.tu-darmstadt.de

Alexander Benlian
Technical University of Darmstadt
benlian@ise.tu-darmstadt.de

Abstract

Artificial intelligence (AI) has gained significant traction in information systems (IS) research in recent years. While past studies have identified many effects of AI technology on human-AI collaborations, there is a paucity in IS literature on the competencies of humans that affect this relationship. In this study, we set out to develop a measurement instrument (scale) for general AI literacy, that is humans' socio-technical competencies regarding AI. We conducted a systematic literature review followed by five expert interviews to define and conceptualize the construct of general AI literacy and to generate an initial set of items. Furthermore, we performed two rounds of card sorting with six and five judges and a pre-test study with 50 participants to evaluate the developed scale. The validated measurement instrument contains five dimensions and 13 items. We provide empirical support for the measurement model and conclude with future research directions.

Keywords: Artificial intelligence, AI competencies, AI skills, human-AI interaction, future of work

1. Introduction

Artificial intelligence (AI) and its impact on individual humans, their organizations, and their work have gained enormous traction within information systems (IS) research in recent years (Benbya et al., 2021; Berente et al., 2021; Jain et al., 2021). AI competence has risen to a key skill for humans with increasing importance for future work, whereas research increasingly calls for ways to improve AI competencies (Tarafdar et al., 2019). While IS research has produced an abundance of studies and frameworks of technical AI features to improve human-AI collaboration (Fügener et al., 2021b), hitherto, there are no mature conceptualizations and instruments available to measure general AI competence. While the impact of technical features on the success of human-AI collaborations is apparent, human socio-technical competence in AI drives it just as much (Cai et al., 2022). Academics did conceptualize and set up non-AI-specific measurement instruments for human IS competencies before AI's ascent (Bassellier et al., 2003). However, these lose their applicability when AI is involved since core assumptions of these IS frameworks are invalidated by AI (Schuetz & Venkatesh, 2020). While there are first conceptualizations to measure AI literacy in IS-adjunct fields (e.g., computer education (Wang et al., 2022)), these are focused on specific settings, such as humans in the role of users of AI applications. While this approach enables the measurement of highly specific aspects of AI literacy, it cannot be applied across different roles, for example, to compare different departments of a firm, such as R&D (e.g., developers of an AI tool) and sales (e.g., users of an AI tool). Assessing how the AI literacy of different roles compares to each other can yield valuable insights into the impact of AI literacy. A general scale independent of the human role is, to the best of our knowledge, still missing in core IS literature. Researchers have called for such a *general* AI competence construct in this still underexplored field to identify and measure key competencies for the future work with AI (Tarafdar et al., 2022).

We aim to fill this gap by defining and conceptualizing general AI literacy and by developing an instrument to measure the level of humans' general AI literacy. Therefore, we specify AI literacy as humans' socio-technical competence consisting of knowledge and experience, which are both distinct competency types that collectively constitute AI literacy. We draw on recent AI theorizing (Baird & Maruping, 2021; Berente et al., 2021; Schuetz & Venkatesh, 2020) and apply it to established design principles of human IS competence conceptualizations (Bassellier et al., 2003; Bassellier et al., 2015) to develop our measurement instrument.

We contribute to research and practice in four ways: First, we extend AI and human IS competence literature by specifying the competencies necessary for human-AI collaboration and establishing a connection between the IS research fields. Second, we contribute a measurement instrument that can be leveraged academically to investigate further relationships in the future of human work with AI or to enhance our understanding of AI acceptance. Third, the developed scale can be leveraged by practitioners, such as firms that can use the dimensions to analyze AI competencies of existing job roles or to determine AI skill requirements for new job roles. Humans working with AI can use the structure to better understand their future roles, assess AI task appropriateness, or develop AI-related ethical awareness. Educational institutions can leverage it to assess their AI curricula. Forth, we leverage the AI literacy construct to structure future research.

In the following, chapter 2 introduces the conceptual foundations of human IS competencies and AI. While chapter 3 describes our research design and results, chapter 4 discusses contributions, and chapter 5 examines limitations and future research directions.

2. Conceptual foundations

2.1 Human IS competencies

Human competencies have been established as a core IS research field with a major influence on IS-related interactions of individuals, organizations, and society (Wiesche et al., 2020). Since the IS literature is not fully consistent in its usage of the term 'competencies' (Chakravarty et al., 2013), we define competencies as human knowledge and experience, as opposed to concepts that include further organizational resources, such as information technology (IT) hardware assets (Sambamurthy et al., 2003). Within the field of human IS competencies, different studies have structured the relevant competencies and investigated their impact.

Looking at the conceptualized structures of human IS competencies, the literature acknowledged that relevant human IS competencies include not only technical competencies but also social competencies, such as business (Bassellier & Benbasat, 2004) or management competencies (Roepke et al., 2000). Thereby, competencies have been structured consistently in line with the socio-technical perspective which emphasizes the importance of interaction between the competence sets (Sarker et al., 2019). Beyond the content structuring of human IS competencies, research further agrees that human IS competencies can be divided qualitatively into 'explicit knowledge', which can be taught, read, and explained, and 'tacit knowledge', which is acquired by experience (Bassellier et al., 2015). Both knowledge forms have been shown to affect IS outcomes, such as performance, and need to be considered in interaction (Bassellier & Benbasat, 2004; Bassellier et al., 2003)

Furthermore, research has identified manifold effects of human IS competencies that exemplify the impact of the human component in IS (Chakravarty et al., 2013). To date, these studies predominantly focus on instrumental outcomes for the organization, such as performance (Croteau & Raymond, 2004) or innovativeness (Tarafdar & Gordon, 2007). Contrary, on the side of humanistic outcomes, such as the well-being of IT professionals, research did, to the best of our knowledge, not identify key relationships to IS competencies yet. Especially, when moving from general IT to the more immersive AI technology a thorough understanding also of the effects on humanistic outcomes becomes more important.

2.2 Emerging theory of AI and AI literacy

After a thorough evaluation of the literature on IS competencies, one might argue that it seems to be a fairly explored field of IS research. However, when assessing literature on emergent theorizing of AI, it becomes imperative to revise also our theories and knowledge on competencies regarding technology subsumed under the term, which differ qualitatively from prior non-AI technology (Berente et al., 2021).

While academics have defined and conceptualized AI from many angles, Berente et al. (2021) provide a concise view that distinguishes AI from non-AI technology by conceptualizing three unique facets of AI: *autonomy, learning,* and *inscrutability*. These three facets invalidate core assumptions that IS theory was built on for decades (Baird & Maruping, 2021; Schuetz & Venkatesh, 2020; Tarafdar et al., 2022). Schuetz and Venkatesh (2020) identified five broken assumptions and evaluated how they would need to be revised to reflect the changes triggered by AI (Table 1). These identified AI facets and revised IS assumptions are the basis to evaluate how AI impacts our understanding of human IS competence.

#	Broken IS assumption	Revised IS assumption
1	Humans are ***users***	Bilateral human-AI relationships
2	The developer defines the ***inputs***	AI is aware of the environment
3	IT artifact use leads to consistent ***outcomes***	AI can be functionally inconsistent
4	The ***way the tool derives*** its outcomes is comprehensible and can be verified	AI can be functionally not transparent
5	There is an artificial ***interface***	Humans can be unaware of their AI use

Table 1. Broken and revised IS assumptions by Schuetz and Venkatesh (2020)

Humans have always used technology, while the IT artifact had the passive role of a tool. With AI, artifacts are more *autonomous* and can assume an agentic role with their own goals (Baird & Maruping, 2021). This capacity enables AI artifacts to delegate tasks to humans which makes their relationship bilateral (#1, Table 1) (Fügener et al., 2021a). Furthermore, AI artifacts are more autonomous, because they are aware of their environment and process new types of input. Voice assistants, such as Alexa, listen continuously and process unstructured data like speech which have not been specified by a developer in advance (#2, Table 1). Contrary to non-AI artifacts, which produce consistent and deterministic outcomes, AI artifacts *learn* which implies functional inconsistency. The artifacts can incorporate feedback from their produced output and adjust their inner workings accordingly (#3, Table 1). Additionally, AI is often not transparent to its users and even developers. Neural networks are *inscrutable* because their enormous complexity makes it impossible for humans to understand how they derive their outcomes (#4, Table 1). Finally, AI artifacts differ from non-AI artifacts, because they do not always have an artificial interface that reveals to the user that they interact with technology. For example, human voice assistants are so close to the actual human voice, that they can interact with humans without them noticing (#5, Table 1) (Wang et al., 2017).

How these revised assumptions impact IS theory has been explored, for example, with regards to the ways humans and AI collaborate (Jain et al., 2021) or organizations function (Benbya et al., 2021). Both underline the human factor and hence also the role human IS competencies will play in AI theorizing. However, we have to assert that hitherto the AI study coverage in the IS literature is highly skewed towards the technical end of the socio-technical continuum, which holds especially for human competencies in AI (Sarker et al., 2019). Nevertheless, there are initial conceptualizations and definitions of AI literacy. Long and Magerko (2020) collocate 17 human competencies and 15 design considerations structured with five key questions: 'What is AI?', 'What can AI do?', 'How does AI work?', 'How should AI be used?', and 'How do people perceive AI?'. Heyder and Posegga (2021) draw on this work and structure the competencies into three conceptual blocks: Functional AI literacy, critical AI literacy, and sociocultural AI literacy. While this structure segments competencies by their content, Ng et al. (2021) structure human competencies by their skill type into three categories inspired by Bloom's taxonomy for competencies (know & understand, use & apply, evaluate & create) with the addition of AI ethics (Krathwohl, 2010). They synthesized a definition for each category based on a literature review (Table 2).

AI skill type	Definition
Know and understand AI	Know the basic functions of AI and how to use AI applications
Use and apply AI	Applying AI knowledge, concepts, and applications in different scenarios
Evaluate and create AI	Higher-order thinking skills (e.g., evaluate, appraise, predict, design) with AI applications
AI ethics	Human-centered considerations (e.g., fairness, accountability, transparency, ethics, safety)

Table 2. AI literacy skill type definitions by Ng et al. (2021)

3. Research design: Towards a scale for measuring general AI literacy

The key objective and contribution of this study is the development and evaluation of a scale to measure the level of general AI literacy. IS research has established systematic and rigorous approaches to develop such a scale (MacKenzie et al., 2011). Adhering to these guidelines we, set up a four-step research design to develop a measurement instrument (Figure 1).

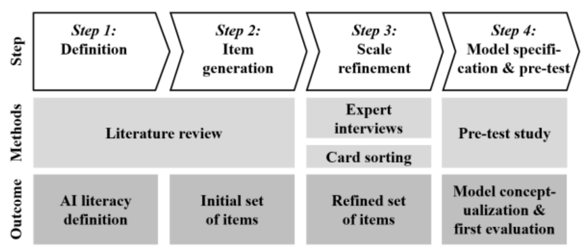

Figure 1. Research design

Initially, we defined general AI literacy to set up the foundation for our focal measurement construct by conducting a systematic literature review on key themes of AI and human IS competence (step 1). Then, we generated items based on our review for each of the conceptualized dimensions of general AI literacy (step 2). Thereafter, we refined the initial scale by interviewing experts and conducting a card sorting exercise to assess content validity (step 3). Finally, we specified the formal measurement model and conducted a pre-test study as a first evaluation (step 4).

The four outlined steps (chapters 3.1-3.4) are in line with MacKenzie et al.'s (2011) steps 1-5 of scale development. In the following, we elaborate on each step of the scale development process.

3.1 Definition of the focal measurement construct (step 1)

We initiated the scale development process by systematically reviewing the literature. To ensure a diligent literature review, we followed established guidelines practiced in IS research (Webster & Watson, 2002). We combined a direct search specifically on '*AI literacy*' with two supplementary searches on the more general fields of '*AI*' and '*IS competencies'* (Appendix B). For the direct search, we used a broad set of 18 IS journals and conferences to cover emergent research, while we focused on the Senior Scholars' Basket of IS Journals for the supplementary searches to only include theorizing with a certain maturity. The search resulted in 172 articles which were screened by reading titles and abstracts. After pre-selection followed by deep reading 21 studies remained relevant for the development of the focal measurement construct. The search was supplemented with relevant papers from adjacent fields identified via forward and backward reference search in the identified papers.

Chapter 2 gave an overview of the conceptual foundations from the underlying theories within human IS competencies and AI identified with the process described above. Furthermore, we assessed existing AI literacy definitions (Table 2, Ng et al. (2021)) and conceptualizations (Heyder & Posegga, 2021; Long & Magerko, 2020; Wang et al., 2022). In the following, we highlight how the presented research is synthesized into a definition to guide the scale development process of the general AI literacy.

Drawing on the literature, we first define the goal of our scale. We aim to establish a scale enabling us to measure AI competence in a general and inclusive way – an approach also followed by other IS constructs (Malhotra et al., 2004). In this context, we refer to AI in the sense of cognitive computing systems (Schuetz & Venkatesh, 2020). To avoid losing practical value or usability, the scale should neither be focused on a specific instance or design of AI, nor a specific job role. This will allow for broad practical applicability, for example, among firms when assessing their organization. The primary target audience shall be all employees in AI-related positions (direct & indirect). Next, we discuss how three synthesized themes (I-III) from the literature inform our adopted definition which concludes the first step.

(I) Core theme of different competency conceptualizations is the *socio-technical perspective* (Bassellier et al., 2015; Sarker et al., 2019). When defining general AI literacy for a scale, we follow this perspective, which implies that the dimensions and items should reflect both, competencies in AI technology as well as competencies in human factors involved in AI. Many AI competencies rely on a high interaction of social and technical aspects. Therefore, we decided to incorporate the theme in the definition by referring to competencies jointly, rather than splitting social and technical competencies on the first level.

(II) The *segmentation of competencies into explicit knowledge and tacit knowledge* is a common split in competence research and has already been applied in non-AI competence conceptualizations (Bassellier et al., 2015). Hence, we incorporate it in our AI literacy definition to guide the scale development accordingly. For further clarity, we distinguish our terminology into *knowledge* (explicit literacy) and *experience* (tacit literacy).

(III) So far, the themes were in line with IS competence theorizing. However, the theme of *broken IS assumptions* demands to accommodate AI's particularities. The first and the fifth revised assumptions state that there is a *bilateral relationship between humans and AI* and that for their interaction there is *no artificial interface* necessary anymore, which underlines the socio-technical perspective (Table 1). Therefore, general AI literacy needs to comprise competencies regarding technology subsumed under the term AI (agentic AI artifacts/actors), such as how AI is distinct from non-AI or where AI can be used. But it also needs to include competencies regarding the human actors involved in the human-AI collaboration, such as tasks where humans are superior to AI or which humans are involved in human-AI collaboration. Furthermore, humans need new competencies on how to recognize AI and what implications it has that humans can now be unaware of their AI interaction (Long & Magerko, 2020). The second, third, and fourth revised assumptions translate into the *steps of how AI handles input, processes the received information, and produces output* (Table 1). For each step, humans need competencies on how to handle what has fundamentally changed compared to non-AI. For example, humans need to know that AI perceives input differently and that input has different effects on an AI artifact compared to a non-AI artifact. Furthermore, humans must develop competencies to judge what it means for an AI application in a certain field (e.g., medicine or business) to not be functionally transparent (e.g., legal and ethical implications or effects on humans interacting with AI). When an AI artifact has derived an outcome, humans now need new competencies on how to handle and interpret it.

Considering the introduced definitions (chapter 2) as well as the identified themes (I – III) of the AI and human IS competence literature, we define for the purpose of developing a measurement instrument:

> *General AI literacy is humans' socio-technical competence consisting of knowledge regarding human and AI actors in human-AI interaction, knowledge of the AI process steps, that is input, processing, and output, and experience in AI interaction.*

3.2 Item generation (step 2)

Based on the literature review described in step 1 and the subsequently adopted definition of general AI literacy, we generated an initial set of items. We considered the item style of previous competence conceptualizations when setting up the items (Bassellier et al., 2003).

In our scale development for general AI literacy, our focus was to measure the human perception of competencies. AI research has shown that metaknowledge, that is one's knowledge about one's knowledge, is a key determinant for the success of human-AI collaboration (Fügener et al., 2021a). While being aware of the drawbacks of a subjective scale, we considered perception which measures the assessment of the own knowledge as a first step towards the measurement of general AI literacy most relevant. Research has further specifically called for better assessment of metaknowledge which our scale contributes to (Fügener et al., 2021a). Additionally, a self-assessment serves the purpose of an inclusive scale that is neither focused on users nor developers as both assess the perception of their literacy for their role in the general construct dimensions.

We aimed to set up the items at the intersections of the three introduced themes 'socio-technical', 'explicit/tacit', and 'revised IS assumptions'. For example, each revised assumption (Table 1) was targeted to be itemized regarding its social and technical implications. Overall, we aimed to start the process with a list balanced around the themes. The initial list comprised 46 items structured along six dimensions (AI & human actors, AI interface, AI input, AI processing, AI output, AI experience) which entered the refinement process.

3.3 Scale refinement (step 3)

In step 3, we used two refinement methods to assess content validity and scale design: First, a round of expert interviews was conducted to incorporate different viewpoints on AI. Second, two rounds of card sorting were performed to assess whether the items are correctly associated with the dimensions.

The combination of systematic literature review and *expert interviews* is recommended by the literature and assumed to identify a set of potential items with high validity (Moore & Benbasat, 1991). Therefore, we conducted partially open-ended expert interviews to identify further dimensions and aspects of general AI literacy as well as gather feedback on the initial set of items. Given that we chose a general approach to AI literacy, we aimed for AI and IS experts with different backgrounds and expertise. In total, we interviewed five experts. Two experts had an academic background and three were practitioners (Table 3).

Both interviewed academics had an AI/IS background with publications in highly ranked journals. The practitioners were selected from a consultancy, an established enterprise, and an AI startup to obtain a holistic view of AI in practice. Four interviews were conducted online and one face-to-face. Initially, experts were asked in an open-ended manner to describe their understanding of AI literacy and how they would conceptualize the construct. After we elicited the expert's views on AI literacy without prior cue through open-ended questions, we showed the expert our conceptualization and items and applied think-aloud techniques for further input. Leveraging open-ended and think-aloud techniques together gave us perspectives we had not been able to see before.

#	Field	Expert (Order of interview execution)
1	AI practitioner	Senior director at international strategy consultancy
2	AI/IS academic	Senior lecturer & researcher
3	AI/IS academic	Post-doctoral researcher
4	IS practitioner	Head of IT department in an established enterprise
5	AI practitioner	Founder of AI start-up

Table 3. List of interviewed experts

A key result of the expert interviews was that all experts intuitively confirmed the importance of the socio-technical perspective, as well as the explicit and tacit knowledge components. Also, from the six originally entered dimensions that were derived from the literature, the three dimensions aimed at an understanding of the AI steps (input, processing, output) could be validated as meaningful. However, the dimensions 'AI & human actors' and 'AI interface' which were also derived from the revised IS assumptions by Schuetz and Venkatesh (2020) did not intuitively resonate with a majority of the experts. Following suggestions from the experts for more clarity, we restructured the two dimensions into 'AI technology' and 'Human actors in AI'. Furthermore, it was recommended to separate the experience dimension into usage and design which we adopted. Subsequently, we restructured the construct into seven dimensions (Table 4). The dimensions are grouped into three categories: AI actor knowledge (explicit literacy), AI steps knowledge (explicit literacy), and AI experience (tacit literacy). Finally, we reworded the item set based on the feedback elicited via the think-aloud technique from the experts. The refined items then entered the card sorting process.

Category	Construct dimension
AI actor knowledge (AK)	**AI technology knowledge (TK)** *Definition: Knowledge of what makes AI technology distinct and the role of AI in human-AI collaboration and interaction*
	Human actors in AI knowledge (HK) *Definition: Knowledge of the role of human actors in human-AI collaboration and interaction*
AI steps knowledge (SK)	**AI input knowledge (IK)** *Definition: Knowledge of what AI input is and how humans should use it*
	AI processing knowledge (PK) *Definition: Knowledge of how AI processes information and what effects it has on humans*
	AI output knowledge (OK) *Definition: Knowledge of what AI output is and how humans should use it*
AI experience (EX)	**AI usage experience (UE)** *Definition: Experience in interacting with AI*
	AI design experience (DE) *Definition: Experience in designing and setting up AI*

Table 4. General AI literacy construct dimensions

Next, we performed two rounds of *card sorting* to ensure further content validity of the items. The method is considered appropriate to validate that items are individually representative of their dimension and that items within a dimension are collectively representative of the entire content of that dimension (MacKenzie et al., 2011). We selected the item placement ratio ('hit-ratio') (Moore & Benbasat, 1991) and Cohen's kappa (Cohen, 2016) as two established measures for the inter-rater agreement to evaluate the card sorting results. Judges for the card sorting exercise have been acquired through the survey platform Prolific and pre-filtered for frequent technology use at work (>2 times a week). The judges were first instructed about the exercise and provided with the definitions of each dimension (Table 4). Thereafter, they were asked to allocate each item to exactly one of the seven dimensions, while additionally the option to choose 'n/a' was given. The items were shown in a randomized order and several attention checks were implemented to ensure that the judges exerted appropriate effort. After excluding judges that failed the attention checks, six judges remained in the first-round exercise and five judges in the second-round exercise.

In the first round, the judges were asked to assign each item from the initial set of 46 items which was the outcome of the expert interviews phase. The average hit-ratio of all dimensions was .52 with a Cohen's kappa of .27, both indicating a need for further refinement. As a result, items with the lowest hit ratios were dropped and the wording of the remaining items was adjusted. A set of 25 items was retained for the next iteration.

The second round of card sorting was conducted in the same setup but with a completely new set of judges. The average hit-ratio improved significantly to .86 which we deemed sufficient following prior research that considers .80 as the average hit-ratio threshold value (Moore & Benbasat, 1991). The hit-ratios in each dimension were also at individually appropriate levels, ranging from .70 to 1.00 (Table 5). Furthermore, Cohen's kappa improved to .74, which lies above the commonly used threshold of .70 (Boudreau et al., 2001). The range of all inter-rater kappa statistics was .62 to .89 which indicated strong inter-rater agreement (Landis & Koch, 1977). Based on the improved inter-rater agreement measures we considered the content validity of the refined item set appropriate.

			Theoretical dimensions							
			AK		SK			EX		-
			TK	HK	IK	PK	OK	UE	DE	N/A
Allocated	AK	TK	14	2	0	0	0	1	0	0
		HK	1	22	2	0	1	0	0	0
	SK	IK	0	0	18	0	0	0	0	0
		PK	0	0	0	17	0	0	0	0
		OK	1	0	0	1	19	0	0	0
	EX	UE	2	1	0	0	0	7	0	0
		DE	1	0	0	1	0	1	10	0
	-	N/A	1	0	0	1	0	1	0	0
Item placement			20	25	20	20	20	10	10	0
Hit-ratio			0.70	0.88	0.90	0.85	0.95	0.70	1.00	-

Table 5. Results of second-round card sorting

3.4 Model specification & pre-test (step 4)

In step 4, we first formally specified the measurement model and then conducted a pre-test study in line with established guidelines (MacKenzie et al., 2011). Subsequently, we assessed the measurement model of the general AI literacy construct by analyzing discriminant validity, convergent validity, internal consistency, multicollinearity, and item loadings (Fornell & Larcker, 1981). According to MacKenzie et al. (2011), the formal specification should capture the expected relationships between the items, construct dimensions, and focal construct. These relationships can be described either as formative ('defining characteristics of the construct') or reflective ('manifestations of the construct'). Based on the initial structure (Table 4), we define our construct as a multidimensional construct that is commonly found in IS literature (Croitor & Benlian, 2019). We specify the items as reflective of their dimensions, such as 'AI technology knowledge', because the dimensions exist at a deeper more embedded level than what the items describe. Furthermore, it seems likely that a change in one item would affect other items in the same dimensions. The dimensions themselves are specified as formative of general AI literacy, because it seems plausible that for example 'Human actors in AI' and 'AI technology' knowledge both increase AI literacy, but a change in the 'Human actors in AI' dimension does not necessarily cause a change in 'AI technology' dimension.

After our specification, we conducted a pre-test study with 50 participants who were asked to state their agreement with the dimension items and overall general construct items on a 7-point Likert scale (strongly disagree to strongly agree). Participants were acquired through the survey platform Prolific with a pre-screening for high technology usage (≥daily) and programming skills to ensure a sample with sufficiently discriminant validity. Several attention checks were applied to ensure that the participants carefully assessed each item. The participants were on average 32.8 years old; the gender distribution was 36% female, 62% male, and 2% other; and the highest educational achievement was a university degree for 66%, a high school diploma for 32%, and an apprenticeship for 2%. Taking into account that our construct is defined to measure general AI literacy, we considered the sample appropriate. While the sample size is at the lower recommended end, we considered it sufficient for an internal measurements pre-test.

Our *initial model* with seven dimensions showed an overall good fit ($R^2 = .81$), however, we discovered that the dimensions within AI steps knowledge (input, processing, output) suffered multicollinearity issues (Variance Inflation Factors (VIF) > 5.00) which we interpret as a too granular approach in a technology-driven dimension for a construct with a general approach.

Subsequently, we adjusted and simplified our measurement model by merging the 'AI steps knowledge' dimensions (IK, PK, OK) into one dimension (AI steps knowledge, SK) which has been conceptualized as a category in the previous chapter already. We excluded several items and retained one item from each of the step dimensions (IK, PK, OK) for the new unified step dimension (SK). Furthermore, we optimized the item selection for consistency in the other dimensions leaving a final set of 13 items for our *adjusted model*.

The adjusted model (Figure 2) also yielded an appropriate fit ($R^2 = .79$) but additionally satisfied all recommended model tests (Table 6) (Fornell & Larcker, 1981): The VIFs were below 5.00 for all dimensions ensuring no multicollinearity problems (Gefen et al., 2011). Furthermore, all dimensions satisfied the Fornell-Larcker criterion (square root of average variance extracted greater than all correlations to other latent variables) indicating sufficient discriminant validity. Cronbach's alpha (CA) was greater than .84 for all dimensions which lies above the commonly accepted threshold of .70 for internal consistency. Finally, item loadings were all above the recommended threshold of .70 (Figure 2) at a significance level of $p<0.001$ and the composite reliability (CR) is above .91 for all dimensions which exceeds the threshold of .80. Overall, the results indicate strong empirical support for the adjusted measurement model. Since the *initial model* satisfied all tests mentioned, except multicollinearity, indicating some empirical support, we include for future reference both in Appendix A (Initial model – 25 items; adjusted model – subset of 13 items)

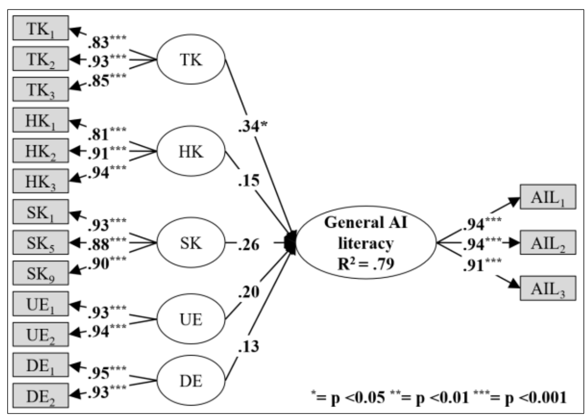

Figure 2. Path analysis of structural equation model (Adjusted model)

While the adjusted model explained 79% of the variance of the general AI literacy construct, only the dimension 'AI technology knowledge' ($\beta = .34$) had a strong and significant ($p<0.05$) effect on general AI literacy. Despite the low path coefficients of the other dimensions, we decided to retain them in the model, as practiced in other IS construct developments (Croitor & Benlian, 2019), because they add key content for the focal construct, the importance of the dimensions might differ in different contexts of AI literacy, and the dimensions do not have collinearity issues (Table 6).

		AK		SK	EX		AIL			
		TK	HK	SK	UE	DE	-	CR	CA	VIF
AK	TK	**.87**						.91	.84	4.47
	HK	.75	**.89**					.92	.87	2.26
SK	SK	.77	.58	**.90**				.93	.88	4.51
EX	UE	.57	.38	.45	**.94**			.94	.86	1.50
	DE	.47	.34	.74	.32	**.94**		.94	.87	2.34
AIL	-	.82	.67	.79	.61	.59	**.93**	.95	.92	-

Table 6. Correlation matrix (square root of average variance extracted in bold)**, Composite Reliability, Cronbach's Alpha, Variance Inflation Factors**

4. Contributions to research and practice

Our conceptualization and measurement instrument of general AI literacy contribute to research and practice in four ways. First, we provide an extension and specification of existing competency conceptualizations (Bassellier et al., 2003) with regards to AI (Schuetz & Venkatesh, 2020). Our developed construct picks up established aspects of IS competency literature and applies AI specificities to them, yielding an empirically tested instrument to measure the level of general AI literacy. By bridging the human IS competence and AI research streams, our instrument also extends the AI literature in IS through structuring human AI competencies. Second, we contribute an instrument to IS research that enables further exploration of the relationships of AI literacy to other effects of interest (including instrumental and humanistic outcomes), such as AI delegation intentions, trust in AI, or the intention to follow AI advice. Thereby, we provide an answer to AI research that called for further exploration of metaknowledge in AI (Fügener et al., 2021a). Furthermore, the instrument can yield insights into the AI-specific aspects of technology acceptance. Potential applications are the assessment of AI literacy within different corporate functions and how it impacts the work, such as setting a strategic AI agenda for managers or how AI features are implemented by product managers. Third, the instrument constitutes a useful and universal tool for practitioners. Without focusing on a specific instance of AI or human role, it can be leveraged as a general tool in the organizational context. For example, it enables companies to analyze and define AI literacy requirements of different roles (product manager, top manager, developer, etc.). Consequently, the respective organizations can identify AI literacy deficits and set up targeted training programs for their employees. Lastly, our instrument's conceptualization structures future AI research within IS. Our five construct dimensions (adjusted model) invite several future research questions which are discussed in the last chapter.

5. Limitations and future research

Our research has several limitations. First, while we deem the applied empirical tests for the first evaluation of the scale development appropriate, further steps (e.g. cross-validation) need to be applied to gain more validity (MacKenzie et al., 2011). Furthermore, our pre-test study was an online sample that pre-selects English-speaking subjects with access to a computer and hence a minimum level of general computer literacy which likely impacts AI literacy. To gain additional insights, the sample needs to be extended to also include other segments of society. Lastly, our sample size was at the lower end of the recommended size which invites future research to test the model with a larger number of observations.

Each of the identified construct dimensions poses an interesting future research direction. Further investigations within each dimension will not only enable further refinement of the instrument but also potentially uncover yet undescribed effects of AI. Potential research questions are summarized below (Table 7). The first question in each row exemplifies further AI content exploration, while the second question indicates potential paths to enhance the scale.

Dimension	Potential research questions
AI technology knowledge (TK)	- *Content*: How does AI technology knowledge in different organizational roles (e.g., developer and product manager) impact the effectiveness of their cooperation? - *Scale:* Knowledge on which AI features is especially decisive to measure AI technology knowledge?
Human actors in AI knowledge (HK)	- *Content:* How does the knowledge of specific human advantages and disadvantages over AI impact human-AI collaboration? - *Scale:* What are the key human actors that should be represented in the general AI literacy of a human?
AI steps knowledge (SK)	- *Content:* Is knowledge of AI input and AI output interpretation sufficient to enable humans to handle ethical dilemmas with AI? - *Scale:* Which AI step (input, processing, output) has the highest impact on general AI literacy?
AI usage experience (UE)	- *Content:* In which organizational roles is AI usage experience most needed and more important than explicit knowledge? - *Scale:* Which specific AI experiences can be itemized in a scale and describe usage experience most appropriately?
AI design experience (DE)	- *Content:* Is high-level AI design experience (e.g., simple visual modeling) beneficial for managers in their role (e.g., enhancing communication with technical employees?)

- *Scale:* Which specific AI experiences can be itemized in a scale and describe design experience most appropriately?

Table 7. Future research directions

6. References

Baird, A., & Maruping, L. M. (2021). The Next Generation of Research on IS Use: A Theoretical Framework of Delegation to and from Agentic IS Artifacts. *MIS Quarterly*, *45*(1), 315-341. https://doi.org/10.25300/misq/2021/15882

Bassellier, G., & Benbasat, I. (2004). Business Competence of Information Technology Professionals: Conceptual Development and Influence on IT-Business Partnerships. *MIS Quarterly*, *28*(4). https://doi.org/10.2307/25148659

Bassellier, G., Benbasat, I., & Reich, B. H. (2003). The Influence of Business Managers' IT Competence on Championing IT. *Information Systems Research*, *14*(4), 317-336. https://doi.org/10.1287/isre.14.4.317.24899

Bassellier, G., Reich, B. H., & Benbasat, I. (2015). Information Technology Competence of Business Managers: A Definition and Research Model. *Journal of Management Information Systems*, *17*(4), 159-182. https://doi.org/10.1080/07421222.2001.11045660

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

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

Boudreau, M.-C., Gefen, D., & Straub, D. W. (2001). Validation in Information Systems Research: A State-of-the-Art Assessment. *MIS Quarterly*, *25*(1). https://doi.org/10.2307/3250956

Cai, W., Jin, Y., & Chen, L. (2022). *Impacts of Personal Characteristics on User Trust in Conversational Recommender Systems* 2022 CHI Conference on Human Factors in Computing Systems,

Chakravarty, A., Grewal, R., & Sambamurthy, V. (2013). Information Technology Competencies, Organizational Agility, and Firm Performance: Enabling and Facilitating Roles. *Information Systems Research*, *24*(4), 976-997. https://doi.org/10.1287/isre.2013.0500

Cohen, J. (2016). A Coefficient of Agreement for Nominal Scales. *Educational and Psychological Measurement*, *20*(1), 37-46. https://doi.org/10.1177/001316446002000104

Croitor, E., & Benlian, A. (2019). Perceived input control on online platforms from the application developer perspective: conceptualisation and scale development. *Journal of Decision Systems*, *28*(1), 19-40. https://doi.org/10.1080/12460125.2019.1616977

Croteau, A.-M., & Raymond, L. (2004). Performance Outcomes of Strategic and IT Competencies Alignment. *Journal of Information Technology*, *19*(3), 178-190. https://doi.org/10.1057/palgrave.jit.2000020

Fornell, C., & Larcker, D. F. (1981). Evaluating Structural Equation Models with Unobservable Variables and Measurement Error. *Journal of Marketing Research*, *18*(1). https://doi.org/10.2307/3151312

Fügener, A., Grahl, J., Gupta, A., & Ketter, W. (2021a). Cognitive Challenges in Human–Artificial Intelligence Collaboration: Investigating the Path Toward Productive Delegation. *Information Systems Research*. https://doi.org/10.1287/isre.2021.1079

Fügener, A., Grahl, J., Gupta, A., & Ketter, W. (2021b). Will Humans-in-the-Loop Become Borgs? Merits and Pitfalls of Working with AI. *MIS Quarterly*, *45*(3), 1527-1556. https://doi.org/10.25300/misq/2021/16553

Gefen, Rigdon, & Straub. (2011). Editor's Comments: An Update and Extension to SEM Guidelines for Administrative and Social Science Research. *MIS Quarterly*, *35*(2). https://doi.org/10.2307/23044042

Heyder, T., & Posegga, O. (2021). Extending the foundations of AI literacy. International Conference on Information Systems, Austin, USA.

Jain, H., Padmanabhan, B., Pavlou, P. A., & Raghu, T. S. (2021). Editorial for the Special Section on Humans, Algorithms, and Augmented Intelligence: The Future of Work, Organizations, and Society. *Information Systems Research*, *32*(3), 675-687. https://doi.org/10.1287/isre.2021.1046

Krathwohl, D. R. (2010). A Revision of Bloom's Taxonomy: An Overview. *Theory Into Practice*, *41*(4), 212-218. https://doi.org/10.1207/s15430421tip4104_2

Landis, J. R., & Koch, G. G. (1977). The Measurement of Observer Agreement for Categorical Data. *Biometrics*, *33*(1). https://doi.org/10.2307/2529310

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

MacKenzie, S. B., Podsakoff, P. M., & Podsakoff, N. P. (2011). Construct Measurement and Validation Procedures in MIS and Behavioral Research: Integrating New and Existing Techniques. *MIS Quarterly*, *35*(2). https://doi.org/10.2307/23044045

Malhotra, N. K., Kim, S. S., & Agarwal, J. (2004). Internet Users' Information Privacy Concerns (IUIPC): The Construct, the Scale, and a Causal Model. *Information Systems Research*, *15*(4), 336-355. https://doi.org/10.1287/isre.1040.0032

Moore, G. C., & Benbasat, I. (1991). Development of an Instrument to Measure the Perceptions of Adopting an Information Technology Innovation. *Information Systems Research*, *2*(3), 192-222. https://doi.org/10.1287/isre.2.3.192

Ng, D. T. K., Leung, J. K. L., Chu, S. K. W., & Qiao, M. S. (2021). Conceptualizing AI literacy: An exploratory review. *Computers and Education: Artificial Intelligence*, *2*. https://doi.org/10.1016/j.caeai.2021.100041

Roepke, R., Agarwal, R., & Ferratt, T. W. (2000). Aligning the IT Human Resource with Business Vision: The Leadership Initiative at 3M. *MIS Quarterly*, *24*(2). https://doi.org/10.2307/3250941

Sambamurthy, V., Bharadwaj, A., & Grover, V. (2003). Shaping Agility through Digital Options:

Reconceptualizing the Role of Information Technology in Contemporary Firms. *MIS Quarterly*, *27*(2). https://doi.org/10.2307/30036530

Sarker, S., Chatterjee, S., Xiao, X., & Elbanna, A. (2019). The Sociotechnical Axis of Cohesion for the IS Discipline: Its Historical Legacy and its Continued Relevance. *MIS Quarterly*, *43*(3), 695-719. https://doi.org/10.25300/misq/2019/13747

Schuetz, S., & Venkatesh, V. (2020). Research Perspectives: The Rise of Human Machines: How Cognitive Computing Systems Challenge Assumptions of User-System Interaction. *Journal of the Association for Information Systems*, 460-482. https://doi.org/10.17705/1jais.00608

Tarafdar, M., Beath, C. M., & Ross, J. W. (2019). Using AI to Enhance Business Operations. *MIT Sloan Management Review*, *60*(4), 37-44. https://sloanreview.mit.edu/article/using-ai-to-enhance-business-operations/

Tarafdar, M., & Gordon, S. R. (2007). Understanding the influence of information systems competencies on process innovation: A resource-based view. *The Journal of Strategic Information Systems*, *16*(4), 353-392. https://doi.org/10.1016/j.jsis.2007.09.001

Tarafdar, M., Page, X., & Marabelli, M. (2022). Algorithms as co-workers: Human algorithm role interactions in algorithmic work. *Information Systems Journal*. https://doi.org/10.1111/isj.12389

Wang, B., Rau, P.-L. P., & Yuan, T. (2022). Measuring user competence in using artificial intelligence: validity and reliability of artificial intelligence literacy scale. *Behaviour & Information Technology*, 1-14. https://doi.org/10.1080/0144929x.2022.2072768

Wang, Y., Skerry-Ryan, R., Stanton, D., Wu, Y., Weiss, R. J., Jaitly, N., Yang, Z., Xiao, Y., Chen, Z., Bengio, S., Le, Q., Agiomyrgiannakis, Y., Clark, R., & Saurous, R. A. (2017). Tacotron: Towards End-to-End Speech Synthesis. Interspeech 2017, Stockholm, Sweden.

Webster, J., & Watson, R. T. (2002). Analyzing the past to prepare for the future: Writing a literature review. *MIS Quarterly*, *26*(2), 1195-1215. https://www.jstor.org/stable/4132319

Wiesche, M., Joseph, D., Thatcher, J., Gu, B., & Krcmar, H. (2020). *"IT Workforce" in MIS Quarterly Research Curations* (A. Bush & A. Rai, Eds.). http://misq.org/research-curations

Appendix A: AI literacy scale items

Dim.		ID	Item All shown items are included in the *initial model* Δ-marked items are included in the *adjusted model*
			I have knowledge of…
AI technology knowledge		TK1	…of the types of technology that AI is built on$^{\Delta}$
		TK2	…of how AI technology and non-AI technology are distinct$^{\Delta}$
		TK3	…of use cases for AI technology$^{\Delta}$
		TK4	…of the roles that AI technology can have in human-AI interaction
Human actors in AI knowledge		HK1	…of which human actors beyond programmers are involved to enable human-AI collaboration$^{\Delta}$
		HK2	…of the aspects human actors handle worse than AI$^{\Delta}$
		HK3	…of the aspects human actors handle better than AI$^{\Delta}$
		HK4	…of the human actors involved to set up and manage human-AI collaborations
		HK5	…of the tasks that human actors can assume in human-AI collaboration
AI steps knowledge	AI input	SK1	…of the input data requirements for AI$^{\Delta}$
		SK2	…of how input data is perceived by AI
		SK3	…of potential impacts that input data has on AI
		SK4	…of which input data types AI can use
	AI Processing	SK5	…of AI processing methods and models$^{\Delta}$
		SK6	…of how information is represented for AI processing
		SK7	…of the risks AI processing poses
		SK8	…of why AI processing can be described as a learning process
	AI Output	SK9	…of using AI output and interpreting it$^{\Delta}$
		SK10	…of AI output limitations
		SK11	…of how to handle AI output
		SK12	…of which AI outputs are obtainable with current methods
			I have experience in…
AI usage exp.		UE1	…in interaction with different types of AI, like chatbots, visual recognition agents, etc.$^{\Delta}$
		UE2	…in the usage of AI through frequent interactions in my everyday life$^{\Delta}$
AI des. exp.		DE1	…in designing AI models, for example, a neural network$^{\Delta}$
		DE2	…in development of AI products$^{\Delta}$
AI literacy (Overall items)		AIL1	In general, I know the unique facets of AI and humans and their potential roles in human-AI collaboration$^{\Delta}$
		AIL2	I am knowledgeable about the steps involved in AI decision-making$^{\Delta}$
		AIL3	Considering all my experience, I am relatively proficient in the field of AI$^{\Delta}$

Appendix B: Literature review sources

#	Search terms	Included journals & conferences
1	**AI**, Artificial intelligence	*Senior Scholars' Basket of IS Journals*: EJIS, ISJ, ISR, JAIS, JIT, JMIS, JSIS, MISQ
2	**IS / IT competenc***, IS / IT capabilit*,	
3	**AI literacy,** AI competenc*, AI capabilit*	*1. Senior Scholars' Basket of IS Journals:* see above *2. AIS Special Interest Group AI:* DSS, ES, ES with Applications, IEEE IS, ISA *3. Key IS conferences:* AMCIS, ECIS, HICSS, ICIS