



## OPEN ACCESS

EDITED BY  
Annette Kluge,  
Ruhr University Bochum, Germany

REVIEWED BY  
Benedikt Leichtmann,  
Johannes Kepler University of Linz, Austria  
Marcus Grum,  
University of Potsdam, Germany

\*CORRESPONDENCE  
Athanasios Mazarakis  
✉ ama@informatik.uni-kiel.de  
Christian Bernhard-Skala  
✉ bernhard-skala@die-bonn.de  
Martin Braun  
✉ martin.braun@aio.fraunhofer.de  
Isabella Peters  
✉ ipe@informatik.uni-kiel.de

RECEIVED 11 July 2023  
ACCEPTED 05 October 2023  
PUBLISHED 27 October 2023

CITATION  
Mazarakis A, Bernhard-Skala C, Braun M and  
Peters I (2023) What is critical for  
human-centered AI at work? – Toward an  
interdisciplinary theory.  
*Front. Artif. Intell.* 6:1257057.  
doi: 10.3389/frai.2023.1257057

COPYRIGHT  
© 2023 Mazarakis, Bernhard-Skala, Braun and  
Peters. This is an open-access article  
distributed under the terms of the [Creative  
Commons Attribution License \(CC BY\)](#). The use,  
distribution or reproduction in other forums is  
permitted, provided the original author(s) and  
the copyright owner(s) are credited and that  
the original publication in this journal is cited, in  
accordance with accepted academic practice.  
No use, distribution or reproduction is  
permitted which does not comply with these  
terms.

# What is critical for human-centered AI at work? – Toward an interdisciplinary theory

Athanasios Mazarakis<sup>1\*</sup>, Christian Bernhard-Skala<sup>2\*</sup>,  
Martin Braun<sup>3\*</sup> and Isabella Peters<sup>1\*</sup>

<sup>1</sup>ZBW – Leibniz Information Centre for Economics, Web Science, Kiel, Germany, <sup>2</sup>Department of Organisation and Program Planning, German Institute for Adult Education - Leibniz Centre for Lifelong Learning, Bonn, Germany, <sup>3</sup>Fraunhofer Institute for Industrial Engineering, User Experience, Stuttgart, Germany

Human-centered artificial intelligence (HCAI) has gained momentum in the scientific discourse but still lacks clarity. In particular, disciplinary differences regarding the scope of HCAI have become apparent and were criticized, calling for a systematic mapping of conceptualizations—especially with regard to the work context. This article compares how human factors and ergonomics (HFE), psychology, human-computer interaction (HCI), information science, and adult education view HCAI and discusses their normative, theoretical, and methodological approaches toward HCAI, as well as the implications for research and practice. It will be argued that an interdisciplinary approach is critical for developing, transferring, and implementing HCAI at work. Additionally, it will be shown that the presented disciplines are well-suited for conceptualizing HCAI and bringing it into practice since they are united in one aspect: they all place the human being in the center of their theory and research. Many critical aspects for successful HCAI, as well as minimum fields of action, were further identified, such as human capability and controllability (HFE perspective), autonomy and trust (psychology and HCI perspective), learning and teaching designs across target groups (adult education perspective), as much as information behavior and information literacy (information science perspective). As such, the article lays the ground for a theory of human-centered interdisciplinary AI, i.e., the Synergistic Human-AI Symbiosis Theory (SHAST), whose conceptual framework and founding pillars will be introduced.

## KEYWORDS

artificial intelligence, interdisciplinary, HCAI, human factors and ergonomics, information science, human-computer interaction (HCI), adult education, psychology

## 1. Introduction

The excitement around artificial intelligence (AI) is sparking a flurry of activity by researchers, developers, business leaders, and policy-makers worldwide. The promise of groundbreaking advances from machine learning and other algorithms drives discussions and attracts huge investments in, e.g., medical, manufacturing, and military innovations (Shneiderman, 2022). However, much of the debate in society is associated with aspects of whether or not AI will replace people in business activities (Del Giudice et al., 2023). In addition, trust in AI systems, transparency, and explaining such systems is not straightforward to end users (Laato et al., 2022).

In a survey in Germany in May 2023, 46% of 1,220 respondents considered AI technologies to be more of a risk for them personally, while only 39% saw opportunities in AI solutions. However, openness to these new technologies decreases with age and increases

with education: most younger people up to age 34 see AI as an opportunity, as do individuals with a university degree (Infratest dimap, 2023). The results of previous surveys conducted worldwide are confirmed, although the results differ greatly depending on the economic development in each country (Ipsos, 2022).

More particularly, in the workplace, there is a risk of creating a defeatist mentality among the employees when ignoring human aspects of AI implementation. Similar examples exist from the past, e.g., knowledge management faced the same challenges around individual, organizational, and technological barriers (Riege, 2005). Nevertheless, it is argued that high levels of human control and automation are likely to simultaneously empower people and not just emulate humans (Shneiderman, 2020). The idea of human centeredness in AI implementation binds these critical research results together.

Artificial intelligence is, by definition, a sequence of mathematical models created by humans, which are executed by computers. The OECD defines AI more precisely: “An AI system is a machine-based system that can, for a given set of human-defined objectives, make predictions, recommendations, or decisions influencing real or virtual environments. [...] AI systems are designed to operate with varying levels of autonomy” (OECD, 2019, p. 23–24). This is a refinement of the definition of McCarthy (2007, p. 2), in which AI is defined as “[...] the science and engineering of making intelligent machines, especially intelligent computer programs. It is related to the similar task of using computers to understand human intelligence, but AI does not have to confine itself to methods that are biologically observable.”

In general, Artificial Narrow Intelligence (ANI) and Artificial General Intelligence (AGI) are distinguished. An AGI system would be an autonomous agent that can learn unsupervised (McLean et al., 2021). ANI has achieved enormous success in determined situations with a low-dimensional phase space, such as strategy games (Lenzen, 2019). However, the methodology of ANI performance can only be applied to a limited range of tasks (Landgrebe and Smith, 2022).

The organization of working processes may profit from a design approach that integrates human and technical intelligence. Such an organization is achieved, among other things, when humans and machines can use their specific skills and when humans and machines mutually support each other in gaining capabilities (Braun, 2017). In this sense, human-centered AI (HCAI) already consists of a set of standards, concepts, and principles, like e.g., fairness, accountability, beneficence, justice, and explicability, to name a few (Huchler et al., 2020). However, these principles are not consistently implemented in practice because of competing goals of productivity and cost-cutting, which has been a traditional challenge of HFE (Spitzley, 1980).

The present article aims at setting the foundations for an interdisciplinary theory of human-centered AI. For this, the article reflects in Section 6.2 on learnings from an interdisciplinary research project, which is considered as a demonstrator. This demonstrator combines normative, theoretical, and methodological concepts from human factors and ergonomics (HFE), psychology, human-computer interaction (HCI), information science, and adult education to study AI at work.

The article introduces each discipline’s perspective on human-centricity and AI at work. It discusses the implications for developing, transferring, and implementing AI at the workplace, including disciplines not yet in the spotlight about HCAI at work. By explaining the fruitful interplay of these disciplines and how they contribute to human-centricity, we will ultimately argue that a theory of human-centered AI will immensely benefit from incorporating critical concepts, norms, and theories from the presented disciplines as well as from an interdisciplinary approach. This article, therefore, sheds light on different co-existing perspectives on and criteria of human-centered AI and will show how they can be meaningfully integrated to answer the research question of *what is critical for human-centered AI at work*. So far, this has only been conducted for, e.g., social sciences (Miller, 2019) and thus has not acknowledged or even included more diverse disciplines. However, many authors have identified an urgent need for collaboration across disciplines for human stakeholders, e.g., for explainable AI (Langer et al., 2021). This article contributes to this research gap by introducing views from heterogeneous disciplines with either a focus on individuals, such as psychology, or a focus on technology, such as HCI.

Moreover, the article leverages insights from disciplines that are either primarily concerned with a work context, such as HFE, or disciplines that study the work context as one research object amongst others, such as adult learning and information science. Furthermore, it systematizes results from the different disciplines by using the five perspectives on human-centricity by Wilkens et al. (2021). In addition, it identifies fields of action for human-centered AI implementation—such as supporting balanced workload, information literacy, providing tailored learning opportunities for low-skilled workers, enhancing technology acceptance, and building trust. By that, it finally sets the foundation for a synergistic human-AI symbiosis theory, which includes these aspects of AI implementation. Overall, this article is written for an interdisciplinary readership interested in human-centered AI. As a result, this article has more of an explorative, descriptive, and conceptual character.

## 2. Supporting balanced workload: HCAI in human factors and ergonomics

Human factors and ergonomics (HFE) has developed concepts and principles for work design, especially for human-technology interaction. These concepts can be applied to AI systems in terms of a division of functions between humans and AI. The principles of HFE inhibit a mutually reinforcing relationship between humans and AI based on capabilities and ethical design principles.

### 2.1. Defining ergonomics and illustrating the methodology

Ergonomics is the scientific discipline concerned with the understanding of interactions among humans and other elements of a system; ergonomics is also the profession that applies theory,

principles, data, and methods to design in order to optimize human wellbeing and overall system performance (IEA, 2000). The terms *ergonomics* and *human factors* are often used interchangeably or as a unit (HFE). Domains of HFE are physical ergonomics (i.e., human anthropometric, physiological, and biomechanical characteristics as they relate to work), cognitive ergonomics (i.e., concerning mental processes), and organizational ergonomics (i.e., optimization of socio-technical systems, including their structures and processes). HFE is a multidisciplinary, human-centered integrating science that is not domain-specific (ILO and IEA, 2021). HFE encompasses not only safety and health but also the cognitive and psycho-social aspects at work. Additionally, HFE can focus on micro ergonomic design aspects—including the design of the procedures, the context, and the equipment and tools used to perform tasks—as well as macro ergonomic design aspects—including the work organization, types of jobs, technology, work roles, and communication (Wilson, 2014). Through their work activity, human beings acquire experiences about physical and social, external as well as internal reality, and they change by self-reflection (Schön, 1983).

A socio-technical “work system” is part of a work process in which a task is accomplished. Through the interaction of working humans with work equipment, the function of the system is fulfilled within the workspace and work environment under specific working conditions (ISO 6385, 2016). “Work design” is a collective term for measures for the purposeful and systematic design of work objects, work processes, and working conditions. The aim of work design is the optimal fulfillment of work tasks, taking into account human development requirements and economic efficiency (Dyckhoff, 2006). In order to increase productivity, work design is based on the rational principle of labor division and the systematization of work processes. The division of labor supports specialization, which makes it possible to use machines and automate work processes. Work design determines which functions are automated and which remain with humans (Baxter and Sommerville, 2011). The systematization of work sequences aims at their method-based optimization (Schlick et al., 2018). Since the division of labor goes along with external supply and a loss of autonomy, the social dimension of human-centered work seems to be indispensable; it is expressed, among other things, in committed cooperation, fair work relationships, and learning opportunities (Ulich, 2011).

## 2.2. Human-centered work design and interactive human-technology systems

Human-centered work design refers to a problem-solving approach from a human perspective and the interactions of the working human with technical and organizational factors. One application focus concerns interactive human-technology systems (i.e., interaction design). For this purpose, knowledge and methods from human factors, ergonomics, usability, and user experience are applied. Human-centered design criteria are documented in legal regulations, standards, and rules (Karwowski et al., 2021).

Human-centered work is characterized by balanced workload situations in order to avoid over- and under-strain and thus

to promote performance, job satisfaction, and learning (Schlick et al., 2018). Work activities should not harm the health of the working person, should not-or at most temporarily-impair their wellbeing, meet their needs, enable individual and collective influence on working conditions and work systems, and contribute to the development of their personality in the sense of promoting their capability potentials and competencies (Ulich, 2011). Design dimensions include work content, time, process, conditions, workplace, or equipment.

Concerning the design of interactive human-technology systems, ISO 9241-2 (1992) Part 2 specifies the human-centered requirements of user orientation, variety, holism, meaningfulness, the scope for action, social support, and development opportunity. In ISO 9241-110 (2020) Part 110 (interaction principles), the human-centered design criteria of task appropriateness, self-describability, expectation conformity, learnability, controllability, error robustness, and user binding of technical systems are concretized. ISO 9241-210 (2019) Part 210 (Human-centered design for interactive systems) emphasizes the user experience, which describes the functional and emotional impressions of a user when interacting with a product or service. If users' requirements are met in a useful way, it is pleasant to use a product.

## 2.3. Human work and human factors

Technical rationalization measures aim to raise work productivity and reduce costs also by substituting the human workforce with machines. Such automation efforts can affect both physical and mental work. To automate sophisticated forms of work that were previously preserved by humans, the use of AI aims to imitate human intelligence (Landgrebe and Smith, 2022). In order to assess the extent to which AI can explain, predict, and influence human behavior requires HFE expertise.

Humans represent a physical, biological, social, and mental entity. They are generally aware that they have a perception, a mind, and a will. A basic requirement of human existence is to access the natural and social environment through interpretive understanding to gain meaning. In communicating with other humans and interacting with the environment, humans gain a deep understanding of the world (Bender and Kolle, 2020). On this basis, they can act purposefully even in the face of incomplete and contradictory information (Wilkins et al., 2014), transfer existing experiential knowledge to new contexts through the understanding of meaning, adopt different perspectives, and anticipate the consequences of their actions. However, the world can only be understood through knowledge of concepts and symbols that emanate from human intelligence (Whorf, 1963). Human intelligence is a complex neurological capacity including reasoning, memory, consciousness, emotions, will, intention, and moral judgment. Humans can take responsibility for acting solidaristic and morally (Böhle, 2009). Such subjective and implicit competencies are the basis of knowledge and innovation work:

- “Knowledge work” refers to the execution of work orders that are to be carried out according to available (un-) complete rules that the working person knows. Knowledge includes

explicit or implicit components that guide action but cannot be verbalized.

- “Innovation work” refers to problem-finding and problem-solving work assignments in which the goal and the path to achieving the goal are not predetermined. Innovation work includes unplannable or poorly plannable, unpredictable intellectual performances, prognostic intellectual performances that do not have precisely defined goals, and diagnostic intellectual performances for which no algorithms can exist because it is unclear what is being searched for in the first place (Hacker, 2018).

There is increased interest to substitute knowledge work by AI. However, as expected, significant aspects of innovation work will remain with humans in the future (Hacker, 2018).

## 2.4. HFE design criteria for HCAI

HCAI encompasses individual safety, the trustworthiness of AI operation, an appropriate division of functions between humans and machines, and conducive working conditions (Huchler et al., 2020). Insofar as AI applications contribute to the automation of knowledge work, the design criteria of conduciveness to learning and social compatibility take on increased importance.

HCAI ties in with the human capabilities and places the mutual reinforcement of humans and AI at the center of their interaction. Such reinforcement will be reached through a complementary division of functions between humans and AI that takes into account the differences between human capabilities and technical functionalities (Rammert, 2009). The core principle of the complementary division of functions strives for higher productivity and adaptivity through a lower overall degree of automation with increasing partial automation and systematic integration of the capabilities of the working human (Huchler, 2022). In order to cope with the uncertainties of work systems, human options of control are extended, e.g., by informal work actions and work-integrated learning processes. Appropriate competencies are to be maintained by designing the interaction between humans and AI in a learner-friendly way (Grote et al., 2000).

Regarding the social responsibility of AI applications, ethical aspects need to be clarified. Ethics focuses on specifically moral action, especially with regard to its justifiability and reflection (Bostrom and Yudkowsky, 2013). In AI use, ethical questions are concretized in the phenomena of uncertainty and social inequality (Brynjolfsson and McAfee, 2012):

- *Uncertainty*: Purposes of the use of technology are not or not exclusively achieved, i.e., the relationship between means and ends is not always comprehensible; this manifests in insufficient transparency, comprehensibility, and manageability, as well as the irreversibility of decision-making processes.
- *Social inequality*: The people who suffer harm are not the beneficiaries; inequality affects personal and informational autonomy in the use of data, the possibility of personal development, decision-making power, and the economic exploitation of work results.

When designing HCAI systems, ethical rules should ensure that human autonomy of decision and action is preserved with moral intent (Floridi et al., 2018). Currently, AI is not yet comprehensively capable of making moral decisions. Moral principles are instead specified by humans and implemented in the form of algorithms that can lead to morally grounded actions as a result. An essential principle of human-centered design is preserving and appropriately using these moral means of control and access (Bülchmann, 2020). This also relates to the human influence of exit points, if necessary.

Bostrom and Yudkowsky (2013) recommend four ethical design principles of HCAI: An HCAI functioning should be (1) comprehensible and (2) its actions predictable in principle; there should be sufficient time for users to react and veto control in the event of a potential malfunction. HCAI should (3) not be easily manipulated, and if a malfunction does occur, (4) responsibility should be clearly defined.

HFE traditionally incorporates a variety of perspectives on human work. In this chapter, many methods and definitions of HFE were presented, which are also relevant to the other disciplines. HFE has a pragmatic concept for the design of work systems, processes, and tools, including AI, some of which are documented in regulations and standards. Knowledge of these methods and definitions is cross-disciplinary relevant, and not limited to HFE, and additionally necessary to prepare the implementation of HCAI at work.

## 3. Information is the key: HCAI in information science

The topic “human-centered AI” has not gained much attention in information science; even “human-centricity” is not a much-discussed concept due to how information science’s central object of study — information — is conceptualized and defined. In the following, it will be argued that there is no need to be explicit about “human-centricity” because, for one, information does not exist outside of human beings, and second, information behavior is central to the development and evolution of humans. If human-centered AI systems are considered information systems, then several implications can be deduced for their design and handling from this argumentation. Overall, there are three ways how information science can conceptually approach human-centered AI:

- (1) The meta-level that discusses HCAI against the discipline’s pragmatic understanding of information.
- (2) The information behavior perspective, which is considered central to human life and that leads to the creation of user models.
- (3) The literacy aspect that reflects on the skills humans need to handle AI successfully.

### 3.1. The meta-level: what is information?

In the literature, *information* is often characterized by using the Semiotic Triangle by Charles Sanders Peirce since it has been argued that information is the basis for the communicative action (in the sense of information as a message, information

as a state) and a communicative act itself (in the sense of exchanging knowledge/being informed, as a process; [Henrichs, 2014](#)). Accordingly, information has a triadic structure consisting of the following:

- *Object*, meaning of the information or semantics,
- *Sign(s)/syntax*, signals for or physical carrier of knowledge and formal-syntactic representation of signs, and
- *Interpretant*, user/usage of information, or pragmatics.

This structure is dynamical since forming the relations between the parts always entails some sort of transmission. The structure is also relational, which results in the need to consider all three parts of the triangle simultaneously when referring to *information* since they are linked inseparably. Information science deals with all three aspects of information, leading to multi- and interdisciplinary studies with, e.g., computer science (that mostly focuses on the signal part, for example, when building digital libraries) or linguistics (focusing on the object part, for example when constructing ontologies or other knowledge organization systems). The most fruitful — and therefore most central — avenue for information science concerns, however, the pragmatics of information, which focuses on the human part of information processing and how humans make use of information. In a nutshell, information science seeks to understand how and for what reasons humans need, gather, and use information. It primarily asks from the interpretant point of view: What is information used for? Which actions are possible with that information, and what do humans need to act properly?

This understanding of information differs from the definition used in, e.g., computer science or telematics that focuses on the signal and disregards the meaning of the information ([Shannon, 2001](#)). Information science considers knowledge as the raw material for the creation of information — knowledge is possible information ([Rauch, 1988](#)), and information is a manifestation or representation of knowledge ([Kuhlen, 2004](#)). Since (formless) knowledge, which exists independently from signals, needs to be brought into a (physical) form to be transmitted ([Stock and Stock, 2015](#)), it can be argued that “information is a thing — knowledge is not” ([Jones, 2010](#)). If humans use that potential information for further action, information materializes. In general, information is used to decrease the amount of uncertainty a human experiences ([Wersig, 1974](#)).

Therefore, information depends on the context in which humans perceive it, and it can be different with different contexts and different humans. Humans construct information by decoding the signal — information does not just exist (whereas knowledge exists even without signals; [Stock and Stock, 2015](#)). This construction takes place in social environments and via means of communication. The recent popularity of ChatGPT and generative language models for AI, as well as how interaction (or communication) with those systems has been designed, is reflective of the relevance of communication in information processes and for information behavior.

### 3.2. Information behavior and user modeling: traces toward the human perspective

Reflections on the concept of *information* and studies on how humans engage with information have mutually influenced each other, developing a shared understanding of the subject. Information behavior research, as a sub-discipline of information science, is concerned with how and why humans interact with information in different informational contexts, including how they use, create, and seek information ([Bates, 2017](#)), actively or passively ([Wilson, 2000](#)), individually, collectively, or collaboratively ([Reddy and Jansen, 2008](#)). Information non-use, such as information avoidance ([Golmann et al., 2017](#)), is also part of the research agenda as well as information sharing, (personal) information management, information practices, information experiences, and information discovery ([Greifeneder and Schlebbe, 2023](#)).

It is remarkable that, similarly to the triadic structure that considers the human, the interpretant, an inseparable part of *information*, the concept of *Human Information Behavior* was never adopted by the research community ([Bates, 2017](#)). Information behavior research underwent several so-called conceptual and methodological turns that are also reflective of the increasing relevance and attention the human being has been attributed over the years: from understanding which information sources and systems humans use to gather information, learning about the information need that motivates humans to interact with information (cognitive turn) and their emotions involved (affective turn), to the role of socio-cultural contexts (socio-cognitive turn) and habitualized information practices (social-constructionist turn) ([Hartel, 2019](#)). Information behavior can only be exposed by humans — manifesting their (for outsiders’ implicit) relevance for information science again.

This also becomes apparent by one central activity of information behavior, i.e., the humans’ engagement in looking for information. [Case and Given \(2016\)](#) emphasize that from birth onwards, humans are prompted to seek information to meet their fundamental needs. Information needs are driven by those fundamental needs, often because the human recognizes a lack of information to meet the fundamental needs ([Stock and Stock, 2015](#)). Information seeking behavior is activated by concrete information needs and is, therefore, active and intentional ([Case and Given, 2016](#)). If a computer or IT system is used to look for information, then [Wilson \(1999\)](#) speaks of Information Searching.

Information behavior research is not a goal in itself — like other disciplines, it seeks to advance information systems to tools that can be easily and efficiently used, that automatically adapt to changing situations, and that are adaptable to the needs of their users ([Elbeshausen, 2023](#); [Lewandowski and Womser-Hacker, 2023](#)). As has been argued, understanding information and information behavior always requires knowledge about humans, e.g., users of an AI system. The complexity of information behavior often prevents the use of quantitative or statistical methods, so that qualitative methods are the main approach.

User modeling is an important activity in this regard, which aims to describe individual users to enable, for example, personalization of search results or groups of typical users that share certain characteristics (e.g., novices and experts). Latter is often realized via personas that represent typical users of a system with very concrete properties (an approach sometimes criticized for replicating stereotypes; Marsden and Haag, 2016). If the information system targets a broad user base, user modeling can be a tough challenge since there is not only a large, diverse group of (possible) users, but user behavior is also dynamic and can change while using the system for particular tasks or over time. Humans adjust informational practices, tactics, and behavior dynamically to match contexts and to maximize the amount of information they can get, e.g., by changing search terms (Pirolli and Card, 1999). In addition to the informational environment and contexts, the information behavior of a person is also connected to their personality (Lewandowski and Womser-Hacker, 2023). The principle of least effort (Zipf, 1949) is also applicable to information behavior: humans tend only to spend the minimum effort to accomplish tasks, often only resulting in only a satisfying (but not the best) result.

Interactive information systems that more dynamically react to users' information behavior and that serve a variety of tasks, need to even better understand humans, their needs, and context to be accepted and add value. Ingwersen and Järvelin (2005) argue — similar to the pragmatic definition of information — that information systems are never used in isolation but are always embedded in personal, organizational, cultural, and more contexts and therefore need to be designed and evaluated within those contexts. Users should be given the opportunity to use information systems purposefully to focus on the task to be fulfilled without being bothered by the challenges of handling the system (Elbeshhausen, 2023, p. 474). For economic reasons, in the corporate context, it is of paramount importance to know which information types and information services are meaningful for employees and which information needs arise (Stock and Stock, 2015). Gust von Loh (2008) distinguishes between objective information needs from workers and employees that arise from a certain job position (e.g., because of the company strategy) and that are independent of a particular staff member, and subjective information needs that are articulated by a specific job holder and that stem from user studies.

Here, strong connections to the human-computer interaction (HCI) field become apparent. Information science and HCI share their focus on humans interacting with information systems, their cognitive and contextual embeddedness while doing so, and the subjective and objective information needs a system has to satisfy (Jetter, 2023). Both disciplines acknowledge that the design of humane (or human-centered) information systems (empirically proven via usability- and user experience methods) benefits from the enrichment of information and contextualization.

Technical information systems may be unable to fit all the information behavioral aspects of a broad user base but may need to focus on a selection of tasks or user types. Furthermore, information behavior also takes place outside of technical or digital environments. Then, educating the users toward a certain behavior and increasing their knowledge about the information system and environment could be an additional approach.

### 3.3. Information literacy as a prerequisite to deal with AI

Although information is central to human development and life, dealing with information in a good and meaningful way is a skill that has to be acquired and cultivated — especially with regard to the ever-increasing complexity of today's digital information environments. To be able to efficiently and ethically deal with information in a particular context, to understand how information is produced, evaluated, and distributed, how it can be effectively searched for, and to assess the personal informational and thinking competencies critically are skills that are subsumed under the term “information literacy” (Griesbaum, 2023). The UNESCO (2013) considers media and information literacy as a core competency for democratic societies that enables citizens to successfully engage and participate in private, vocational, and societal activities. Information literacy is, however — and similarly as the concept at its core: information — a relational concept. Its characteristics change with the information environments and contexts in which human beings have to deal with information (Griesbaum, 2023). This also presumes that an information-literate person has a certain amount of knowledge about the topic or circumstances they are dealing with — but this is not always the case. Hence, the more the person lacks expertise and knowledge, the more trust the person needs to put into the information ecosystem. Information literacy then transfers from the topic or situation itself to the evaluation of other information sources or experts whose recommendations have to be trusted (Griesbaum, 2023).

Despite its stated relevance, often, information literacy is not an integral part of school education but rather embedded in higher education and services of university libraries (ACRL, 2016). In work-related contexts, information literacy issues become apparent in enterprises with structured knowledge management approaches (Travis, 2017). However, Lloyd (2013) has found that information literacy at the workplace is mainly reduced to socio-cultural practices for collaboration. Middleton et al. (2018) could show that information literacy is strongly connected to innovative work practices.

This hints toward an increasing need for information literacy in complex information contexts as induced by AI systems. Consequently, AI literacy is an emerging field in information science, borrowing most of the central aspects already embedded in information literacy but also highlighting further skills and normative claims (Touretzky et al., 2019). Ng et al. (2021a,b) performed a literature search on AI literacy to derive aspects this concept entails. They found that all selected articles consider knowing the basic functions of AI and how to use AI applications in everyday life ethically (know and understand AI), as well as applying AI knowledge, concepts, and applications in different scenarios (apply AI), the core competencies of AI-literate humans. Two-thirds of the analyzed articles also mention critical higher-order thinking skills (such as evaluating, appraising, predicting, and designing) as part of AI literacy (evaluate and create AI).

Furthermore, the literature states that AI literacy is central to the future workforce, simultaneously preparing humans to efficiently use and critically evaluate AI and sparking career interest in this field (Chai et al., 2020). Interestingly, the study revealed

that only 50% of the articles considered educating humans about socially responsible behavior when using or designing AI as part of AI literacy. Here the authors see room for improvement: “[...] conceptualizing AI literacy with human-centered considerations is crucial to building a future inclusive society” (Ng et al., 2021a, p. 507). The evaluation of AI literacy itself is conducted via knowledge tests, self-reporting, questionnaires, or observations when interacting with AI.

## 4. Designing learning opportunities for all: HCAI in adult education research and adult learning

The core assumption of HFE is that human beings are able and willing to learn and shape their working lives. Considering this first assumption, aspects of lifelong learning, adult and continuing education, and adult learning at the workplace touch the core of work design and human centeredness. At the same time, for many years adult education policies and research have dealt with how educational systems can effectively provide knowledge and skills for a technologically changing (working) society (Merriam and Bierema, 2013). The results of numerous research activities within the adult education scientific community contribute to shaping AI-affected workplaces in a human-centered way, which from an adult educational perspective means a learning-centered way (Harteis, 2022).

From the perspective of adult education research and policy, it is a consensus that educational systems target fostering the quality of educational processes and providing learning opportunities equally (UNESCO, 2019; BMAS and BMBF, 2021; Council of the European Union, 2021; Autor:innengruppe Bildungsberichterstattung, 2022; OECD, 2023).

Educational systems contribute to designing and establishing *learning opportunities*. That means they contribute to channeling, organizing, and monitoring informal, non-formal, and formal (adult) learning processes. In democratic states, educational systems aim to provide skills to individuals so that they can actively participate in public and working life. At the same time, quality learning processes within educational systems underlie the expectancy to provide a qualified and employable workforce. Scientific discourse treats these aspects using the two concepts of individual self-regulation, on the one hand, and human resources, on the other hand (Autor:innengruppe Bildungsberichterstattung, 2022).

Work has a major role to play in education. First, work and working life are significant fields of adult learning. Human subjects acquire skills for and within their employment to stay employable. Second, in Western-so-called working societies-work is a major part of active participation in society, as it impacts, e.g., social status and social and financial resources as much as professional and, therefore, social identity (Kraus, 2008; Gericke, 2017). Third, work in terms of work-based learning is a learning and teaching methodology (Bauer et al., 2004; Dehnbostel, 2022). Therefore, it is an important quality criterion of professionally designed adult and continuing education to take aspects of individuals’ (working) lives, such as possible ruptures in (working) biographies and career

development, as one starting point for developing and creating learning opportunities.

*Equality of opportunities* in adult education refers to equal access to learning opportunities as a major challenge for education systems (Käpplinger and Lichte, 2020; Council of the European Union, 2021). It targets especially vulnerable and marginalized groups such as migrants, low-qualified, unemployed, disabled, or illiterate persons, who have different learning needs regarding content but also need differently structured learning opportunities than high-skilled workers or middle-class citizens. Work has an important role to play in the equality of learning opportunities. According to the Adult Education Survey (BMBF, 2022), for years, more than 70 up to 75% of the adult learning activities of the German population aged 16–65 have taken place during daily working time or were financed by the employer. A much smaller and even decreasing part of adult learning activities was work-related but based on individually generated financial and time resources (13% in 2012 – 8% in 2020), while the share of individual non-work related learning activities is relatively stable at about 17–18% (BMBF, 2022, p. 22). Major differences exist in the participation rates of different social groups. The employed population shows a higher participation rate (46% in 2012 – 60% in 2020) than the unemployed population (13% in 2012 – 19% in 2020). The same applies when comparing the un- or low-skilled population (30% in 2012 – 46% in 2020) with high-qualified persons (about 70% in 2012 – 81% in 2020). Remarkably, in the German adult population, learning activities are on the rise in absolute numbers. At the same time, the differences in share between certain social groups have not remarkably diminished. These findings concerning the participation rates in adult learning vary across countries, still the gap between employed and unemployed, as much the high – and the low-skilled persons, shows to be a central challenge in more or less all OECD countries (European Union, 2021).

Against this background, an adult education research perspective on human-centered AI implementation will concentrate not only on how to provide quality learning opportunities but it will focus as well on how to tailor these quality opportunities for each social group. So, an adult education research perspective contributes to human-centered AI, first of all, by analyzing if an educational system, an employer, or a single workplace offers learning opportunities for AI-based workplaces, whom these learning opportunities are made for, and what kind of learning opportunities are proposed.

### 4.1. Designing learning: what skills should we qualify for?

Regarding contents and needs for skills, there is consensus that in a digitalized and AI-based world, life and work tasks will become more complex. There are catalogs trying to capture and describe important future skills. In the context of education and lifelong learning, the European Commission’s Framework DigComp has had quite an impact in the field in Germany (Joint Research Center (JRC), 2022). Moreover, in higher education, the so-called twenty first-century skills play an important role (Anandiadou and Claro,

2009; Schnabel, 2017). These two catalogs represent important examples for a whole discussion that brings the importance of future skills to the fore. They concentrate on skills in

- Working with media, technology, information, and data
- Virtual and face-to-face communication and collaboration in diverse (e.g., interdisciplinary, intercultural, intergenerational) contexts
- Creative problem solving, innovation, analytical and critical thought
- Flexibility, coping with ambiguity, self-motivation, and working independently (Schnabel, 2017)

When thinking human centeredness from an adult learning perspective, it is important to note that these skills will not replace professional skills but will additionally be on top of professional skills. Even more, they will be interlinked with professional skills. So when preparing a workforce for an AI working world, degrees will have to encompass professional skills as a basis plus these future skills.

## 4.2. Designing learning: who should we target and how?

Quality is not determined by knowing the skill needs and contents of learning but also by methodologies that help to teach these skills professionally, effectively, and efficiently to a whole range of target groups. When it comes to skill delivery in companies, there is a vivid research landscape on how to deliver sustainable learning success in digital transformation. The learning and teaching methodologies in focus range from informal learning in the workplace, learning nuggets, non-formal workshop settings, or formal learning arrangements within chambers and universities (e.g., Rohs and Ganz, 2015; Anderson and Rivera-Vargas, 2020). From the company's viewpoint, where there are financial and economic restrictions, these discussions are critical. With a company's decision to invest in one or another kind of learning opportunity, it shapes structures and methodologies of learning and, finally, participation rates in adult learning opportunities to a high degree.

Discussions of teaching and learning methodology differentiate along the question of which knowledge or skill can be efficiently and professionally taught in which setting to which target group. Taking marginalized groups as an example, it is a common educational argument based on Bourdieu (Watkins and Tisdell, 2006) or biographical research (Alheit, 2021) that low-skilled people or functional illiterates have rather negatively experienced learning throughout their lives, sometimes they have gone through biographies of failing in an educational system. Therefore, it is highly challenging for professionally organized adult education to get access to these groups and to teach them effectively—much more challenging than teaching high-skilled people or managers who have had successful learning careers.

Therefore, when implementing AI in a human-centered way, quality learning opportunities need to ensure that all target groups who are affected by AI in the workplace get the

opportunity to qualify for these changes. At the same time, different target groups will need different skills in the workplace and different learning methodologies for acquiring these skills. In addition, it is a professional adult education task to create good learning opportunities with a well-fitted methodology that facilitates between the affordances of a company within the digital transformation, on the one hand, and the needs of the target group and their learning habits, on the other hand.

In terms of learning methodologies, recent projects have shown

- (a) That professionally implemented learning projects in the workplace can effectively qualify low-skilled workers on the job within digital transformation processes (Goppold and Frenz, 2020).
- (b) That worker's councils have an important role to play as facilitators of bringing together unskilled and low-skilled workers or functionally illiterate employees with continuing education activities. Still, the members of worker's councils need to be qualified to fulfill their role (Lammers et al., 2022; Arbeiter, 2023).

In recent years, networked structures in adult education have been brought to the fore (e.g., UNESCO, 2015). In the case of implementing human-centered AI networks between adult education providers, companies and worker's councils will probably be in favor of channeling professionally tailored learning opportunities into companies. Taking Germany as an example, vocational training providers create those networks in order to target marginalized groups; in the case of high-skilled individuals, universities of applied sciences have a mandate of developing continuing education to create opportunities in cooperation with companies (Dollhausen and Latke, 2020). It is an issue if these networks allow scaling up learning opportunities for a whole population or multiple companies and not to tailor adult learning for one single company.

## 5. Technology, autonomy, and trust: HCAI in psychology and human-computer interaction

Typically, the AI research community focuses on algorithmic advances, deeming a human-centered approach unnecessary, but at the same time, human-centered thinking is gaining popularity, and the AI community is diverse. However, this new thinking challenges established practices (Shneiderman, 2022, p. 40). This new thinking also influences the perception of psychology and HCI, which are closely related, although they are separate disciplines, because psychology plays a significant role in HCI. For this reason, both disciplines are discussed together in this section (Clemmensen, 2006). From a psychological and human-computer interaction (HCI) point of view, technology acceptance and adoption are also becoming essential aspects of human-centered AI (Del Giudice et al., 2023), especially in human autonomy (Bennett et al., 2023) and the development of guidelines for human-AI interaction (Amershi et al., 2019). Such guidelines need to consider issues with information overflow and should assist in using complex

systems (Höök, 2000). This consideration can be achieved by putting in place verification measures or regulating levels of human-controlled autonomy to prevent unintended adaptations or activities by intelligent systems (Amershi et al., 2019; Xu et al., 2023). In addition, AI-driven influence techniques like psychological targeting or digital nudging have raised ethical worries about undermining autonomy (Bermúdez et al., 2023). Moreover, a series of recent studies found that employees who work with AI systems are more likely to suffer loneliness, which can lead to sleeplessness and increased drinking after work (Tang et al., 2023).

Nevertheless, in HCI, the understanding of human autonomy remains ambiguous (Bennett et al., 2023). This ambiguity might be attributed to an old controversy if people and computers being in the same category or if, as many HCAI sympathizers believe, vast differences exist (Shneiderman, 2022, p. 25), with Shneiderman supporting the latter (Shneiderman, 2022, p. 31). However, AI and its impact on the workplace are said to be disruptive, including chatbot-based communication systems that can demonstrate empathy through an understanding of human behavior and psychology, allowing the chatbot to connect with customers emotionally to ensure their satisfaction and thus support the adoption of AI systems (Krishnan et al., 2022). AI differs from HCAI by two key human-centered aspects in terms of performance and the product. The human-centered process is based on user experience design methods and continuous human performance evaluation. Furthermore, the human-centered product is emphasized by human control to enhance human performance by designing super tools with a high level of automation (Shneiderman, 2022, p. 9).

Still, HCI acknowledges the importance of AI by highlighting it in almost all of the current HCI grand challenges, like human-technology symbiosis and human-environment interactions, to name a few (Stephanidis et al., 2019). This is accompanied by six grand challenges of human-centered AI: human wellbeing, responsible design of AI, privacy aspects, AI-related design and evaluation frameworks, the role of government and independent oversight, and finally, HCAI interaction in general (Garibay et al., 2023). HCAI interaction especially plays a vital role at work, as economic challenges meet with ethical and organizational considerations (Garibay et al., 2023). This collection of grand challenges reflects the almost symbiotic relationship between HCI and AI.

Finally, the transition to human interaction with AI systems by moving on from siloed machine intelligence to human-controlled hybrid intelligence can be considered a new opportunity for HCI professionals to enable HCAI (Xu et al., 2023). A potential goal of human-centered AI design is to create human-controlled AI using human-machine hybrid intelligence, which emphasizes the integration of humans and machines as a system, aided by the introduction of human functions and roles that ensure human control of the system (Xu et al., 2023, p. 503). However, such integration of humans and machines is not without obstacles and unrealistic user expectations, and negative emotional responses are often a source of concern.

## 5.1. Unrealistic user expectations

Exaggerated and unrealistic user expectations about AI-based applications and absent design solutions to support human-centered work can lead to frustration and questioning the “intelligence” of such applications (Luger and Sellen, 2016). For example, high efficiency of search functions should be combined with curated content and meaningful recommendations even without the necessary meta-information. These demands raise hopes that may neglect the actual software and hardware capabilities of research projects, which may only be feasible for very large companies. In addition, it is requested to combine, match, and recommend different kinds of heterogeneous data, even on the internet, without considering resources. Although AI makes significant improvements daily, these are still quite unrealistic user expectations today.

One possible way to overcome these challenges of unrealistic user expectations is to divide the AI-based processes into different phases. For example, Amershi et al. (2019) offer 18 AI design guidelines separated into four phases: the initial phase when beginning to work with an AI-based application, during general interaction, when things go wrong, and aspects considering long-term experiences. A vital aspect of these guidelines is providing support and managing expectations. Such aspects are not unknown to technology acceptance models.

## 5.2. Building on technology acceptance and trust

Technology acceptance models, e.g., TAM (Davis, 1989), UTAUT (Venkatesh et al., 2003), and their extensions in various fields (Kao and Huang, 2023), offer a promising domain for an evaluation concerning human-centered AI. TAM is a conceptual model used to account for technology usage behavior, which has been confirmed to be valid in various technologies among different groups of people (Venkatesh et al., 2003; Choung et al., 2023). The original TAM model postulates that the intention to use technology in the future is determined by two key factors: perceived usefulness and perceived ease of use (Davis, 1989).

Choung et al. (2023) integrated trust as an additional variable in their extended TAM model. Their two studies confirm that trust is vital for accepting technology. Therefore, AI technologies should be designed and implemented in a human-centered way; consequently, their implementation should be easy to use, useful, and trusted. In general, empirical findings support the assumption that technology acceptance models help to explain the acceptance of AI technologies (Sohn and Kwon, 2020), including the aspect of trust (Choung et al., 2023). However, nevertheless, there are limitations to their usage, which are discussed by Bagozzi (2007).

Users' low level of trust in how their data is handled and processed must also be adequately considered psychologically. AI, in general, can predict user behavior in a wide range of applications by following digital traces of usage. Besides legal and ethical challenges, psychologists call this approach digital phenotyping when using elaborated smart sensing techniques and when it is successfully assisted and analyzed by data mining and machine

learning tools (Baumeister et al., 2023). This is not an entirely new topic, as, e.g., user behavior in an online environment relates to their personality and can be used to tailor content, improve search results, and increase the effectiveness of online advertising (Kosinski et al., 2014), which is backed by many empirical studies and summarized by Baumeister et al. (2023). At the same time, ethical challenges are addressed by the human-in-the-loop design, where individuals are asked to make a final decision or action (Shneiderman, 2022; Garibay et al., 2023; Xu et al., 2023), which can also help to improve the trust to HCAI.

This is in line with results from an experiment by Westphal et al. (2023), in which they empower users to adjust the recommendations of human-AI collaboration systems and offer explanations for the reasoning of the systems. The idea behind this approach is to counter low trust and limited understanding of users dealing with recommendations of an AI system, and at the same time, to keep in mind to achieve an adequate or calibrated trust, meaning that, e.g., not to over trust the AI system (Leichtmann et al., 2023b). However, interestingly, explanations could backfire because they can increase or signal task complexity, whereas enhanced decision control leads to higher user compliance with system recommendations (Westphal et al., 2023). These results affirm that well-explained support can be essential to accept and facilitate HCAI at work, leading to HCAI systems that explain themselves, so-called human-centered explainable AI, which can be accompanied by educational offers and measures for providing human-centered explainable AI.

### 5.3. A glimpse into the near future: explainable, understandable, and gamified AI

Such explainable AI, especially if it is human-centered, can be considered crucial. However, from a socio-technical standpoint, AI should also be understandable to stakeholders beyond explainability (Habayeb, 2022). This can be achieved when implementing user-participated experimental evaluation because it is necessary to overcome the relatively simple unilateral evaluation methods that only evaluate AI systems' performance.

One way to implement this is to use, e.g., a gamified crowdsourcing framework for explainability (Tocchetti et al., 2022), which uses game design elements in a non-game context (Deterding et al., 2011). However, current research focuses primarily on strategic and system issues related to AI system performance (Raftopoulos and Hamari, 2023), which limits the view of AI. Furthermore, HCI should promote the evaluation of AI systems as human-machine systems by including the end-user perspective (Xu et al., 2023, p. 505). Nevertheless, this makes it necessary for HCI to enhance its current methods. Constraints like focusing on single user-computing artifacts with a limited context of use, lab-based studies, or static human-machine functions are prevalent. Instead, "in-the-wild studies," the application of distributed contexts of use, and longitudinal study designs are encouraged to address the identified unique issues of AI systems to influence the development of AI systems in a human-centered way (Xu et al., 2023, p. 509–512).

Additionally, incentives for using artificial intelligence, e.g., gamification, are also emerging topics of interest for human-centered AI (Mazarakis, 2021). Gamification tries to bring the motivating effect associated with games to non-game situations with the help of elements like badges, leaderboards, and points (Mazarakis, 2021, p. 279, 283). First studies conducted with intelligent user interfaces and voice user interfaces like Amazon Alexa, which are also considered social robots and active appliances in artificial intelligence (Shneiderman, 2022), show the potential to focus on empirical research for the acceptance of these interfaces (Bräuer and Mazarakis, 2022a,b; Haghghat et al., 2023). Nevertheless, explainable AI is also, in this use-case, key to counteracting suspicion regarding the trust of social robots and active appliances in artificial intelligence. For example, to achieve transparent and accountable conversational AI and to include such a system in a gamified environment, interpretability, inherent capability to explain, independent data, interactive learning, and inquisitiveness are necessary (Wahde and Virgolin, 2023, p. 1856). Inquisitiveness is meant to be by the AI to show curiosity and not to annoy the user to achieve human centeredness. Curiosity means just displaying inquisitiveness in specific contexts, such as during learning, so as not to disturb the user (Wahde and Virgolin, 2023, p. 1865).

A further step is taken by Tocchetti et al. (2022), which propose a gamified crowdsourcing framework for explainability. Their crowdsourcing framework engages users on different levels than other platforms, primarily relying on extrinsic rewards. The provided user education, in particular, would raise users' understanding of the types of information that an AI system requires, learns, and produces, improving users' efficiency and developing the users' mindsets (Tocchetti et al., 2022, p. 7). Furthermore, their work shows that a symbiosis of HCAI, gamification, and explainable AI is also possible with greater effort and exertion. Consequently, this also results in an increased human centeredness. A first effort of studies in game-based environments yields promising results for explainable AI (Leichtmann et al., 2023a).

Different scenarios for using gamified AI and gamification in the context of AI, in general, are possible. For example, Tan and Cheah (2021) describe a work in progress and prototype for developing an AI-enabled online learning application for lecturing at a university physics. However, this scenario can be switched to a work-related setting without much effort. As education is one of the main application areas of gamification (Mazarakis, 2021), a combination with AI is obvious and already taking place (Kurni et al., 2023). In this case, first data is collected for AI processes, e.g., through step-by-step scaffolding instructions and feedback to students by studying students' progress in answering quiz questions. Then, it is possible to implement adaptive assessments to more accurately identify the student's level of mastery, adjust the difficulty level and the number of questions at each level of difficulty, and finally, step between each level of progression based on the student's answer. Thereby, individual feedback, which can then be used for learning analytics to improve, optimize, or redesign the curriculum to meet the needs of specific student cohorts, can be provided (Mazarakis, 2013; Tan and Cheah, 2021), e.g., for employer-provided training in different work scenarios.

## 6. Discussion

This chapter presents conclusions from the previous chapters' theory and practice and shows relations between them in order to inform a synergistic human-centered AI theory. First interdisciplinary human-centered AI perspectives are shown, according to Wilkens et al. (2021), and how they relate to the five disciplines. Then, interdisciplinary views of human-centricity and their interrelations are matched with observations from a demonstrator. These views have the goal of setting foundations for an interdisciplinary synergistic theory of human-centered AI, which is presented in Section 6.3.

### 6.1. Interdisciplinary perspectives on human-centered AI

Wilkens et al. (2021) found five co-existing views in a comprehensive literature review analyzing the significance of HCAI: a deficit-oriented, a data reliability-oriented, a protection-oriented, a potential-oriented, and a political-oriented understanding of how to achieve human-centricity while deploying AI in the workplace. These five perspectives reflect many aspects of AI's human-centricity, with varying levels of maturity along each dimension. In order to put the results of this article into context, the disciplines of information science, human-computer interaction, psychology, and adult learning are related in Table 1 to the five perspectives of Wilkens et al. (2021). HFE is inherent in Wilkens et al. (2021) and would cover all five perspectives, so it is not shown in Table 1.

The perspectives of Wilkens et al. (2021) largely coincide with the results of the demonstrator. Although all the perspectives are covered, HFE is particularly concerned about the data reliability-oriented understanding of HCAI and the potential-oriented understanding.

Potential deficits of data reliability are mainly considered from an ethical perspective since technical artifacts are not ascribed to any moral competence; at best, they can imitate human moral behavior. Taking into account the instrumental character of AI, it is rather necessary to consider fundamental conflicts of interest that might favor immoral behavior. This, however, leaves the field of AI design.

The potential-oriented understanding promotes a hybrid design approach corresponding to the "complementary division of functions between humans and AI." The mutual reinforcement of humans and HCAI appears to be suitable for satisfactorily coping with future, currently potentially unknown working requirements. It emphasizes the evolutionary principle of humans, whose behavior is imitated by intelligent machines, and that will ultimately be reflected in the precision and reliability of machine procedures.

The information science perspective on human-centered AI can be summarized mainly as protection-oriented, potential-oriented, and political-oriented. It is protection-oriented when it studies how information systems should be designed so that humans can easily and safely use them, and it is potential-oriented since it considers information systems sociotechnical environments in which humans co-construct information with the help of

technology. In social settings, and especially at the workplace, the political-oriented perspective is also part of information science's agenda, e.g., in terms of information literacy.

An adult education perspective, which in its tradition always includes an advocacy perspective, can be contextualized as a protection-oriented, a potential-oriented, and a policy-oriented approach. It is protection-oriented when talking about qualifying workers for correct decision-making in cooperation with AI systems. It is potential-oriented when thinking about how to use AI systems for quality learning opportunities, e.g., in learning analytics. Finally, when reflecting on the advocacy tradition of empowering social subjects, an adult education perspective is politically oriented.

It is not surprising that from a HCI and psychology point of view, most perspectives by Wilkens et al. (2021) are relevant, as they touch technology and individual aspects at the same time. Interestingly, the deficit-oriented understanding and data reliability-oriented understanding perspectives are more related to HCI. So assisted tools that work through elaborated sensor technology are common to, e.g., gamified and explainable AI, which are fields of HCI.

In contrast, the protection-oriented understanding and potential-oriented understanding perspectives are closely related to psychology. Unrealistic user expectations questioning the "intelligence" of AI, in connection with loss of autonomy and trust, are prevailing psychological aspects of human-centricity for these two perspectives. Nevertheless, these areas can also be found in HCI and, depending on the degree of technical implementation, are likely to be assigned to HCI.

It is clear from Table 1 that the disciplines do not cover all of Wilkens et al. (2021) perspectives, but there are different emphases, with an imbalance existing for the first two perspectives. The article at hand shows that all disciplines are important for HCAI, with HFE functioning as an umbrella discipline for HCAI at work. This makes the call for an interdisciplinary view obvious. The following section will detail these views that are enriched by findings from a demonstrator.

### 6.2. Interdisciplinary views of human-centered AI

This section presents the different fields of action, the relevant factors of human-centered AI from different disciplines, and the interdisciplinary views of HCAI, including possible areas of collaboration. Adding insights from a demonstrator, the foundations for a synergistic human-AI symbiosis theory are revealed.

For HFE, human-centricity means involving humans in the design of work systems, e.g., processes or tools. The design is based on a comprehensive understanding of users, tasks, and work environments. Human-centered design aims at balanced skill and performance development as the basis of work productivity and health. HFE combines the human potential with technology. From a HFE perspective, human-centered AI emphasizes that the decision-making competence of humans and their intervention in technical systems in doubt are weighted higher than that of

TABLE 1 Analysis of human-centered AI perspectives for information science, adult education, human-computer interaction, and psychology according to Wilkens et al. (2021).

Discipline	Deficit-oriented (Perspective 1)	Data reliability-oriented (Perspective 2)	Protection-oriented (Perspective 3)	Potential-oriented (Perspective 4)	Political-oriented (Perspective 5)
Information science			x	x	x
Adult education			x	x	x
HCI	x	x			
Psychology			x	x	

AI machines. Information science and HCI provide details and key features of how to calibrate this interaction of humans and computers. A key factor lies in aspects of designing AI systems with regard to personality issues, search behavior, and information literacy of users.

From an *adult education* perspective, the success of human-centered AI lies in providing learning opportunities across all target groups affected by AI. These learning opportunities need to adapt to learners and their individual learning habits and learning needs in terms of content and methodology instead of one-fits-all solutions. Especially marginalized learners will have to be a focus. Information science can add to this perspective with its differentiated insights on information literacy, a concept that might be helpful when implementing HCAI solutions and, up-to-date, is far away from being profoundly treated in the field of adult education with, e.g., low-skilled adults. Furthermore, HFE perspectives can help in elaborating adult learning methodologies, as HFE provides clear perspectives on workplace learning and how it adds to effective learning.

Additionally, mutual reinforcement of humans and AI, which considers ethical principles during design, is essential and is actively considered in *information science*, e.g., in aspects of usability and interaction design. Especially the more profound understanding of sociotechnical information systems and how humans interact with digital environments are essential to information science.

Finally, for *HCI and psychology*, human-centered processes and products are the foundation of human-centricity. As a result, user experience design approaches and ongoing human performance evaluation are required. Additionally, human-centered AI products are emphasized by human control (Shneiderman, 2022, p. 9). These elements should be utilized to help tackle the six grand challenges of human-centered AI: human wellbeing, responsible design of AI, privacy aspects, AI-related design and evaluation frameworks, the role of government and independent oversight, and finally, HCAI interaction in general (Garibay et al., 2023). The overlap between HFE and HCI is visible where interaction design is considered from a perspective when knowledge and experience of usability and user experience are applied. This helps, among other things, to implement technology acceptance models and, thus, to build trust and support when the overlap between HFE and HCI is consistently implemented. In addition, HFE and psychology (and partly also HCI) meet in the field of human autonomy. Furthermore, the relationship between information science and HCI is exemplified in the interaction of any kind of information and the focus on human-centered systems, again considering usability and user experience.

Recognition of the importance of information literacy should be considered an important link for HCAI in this regard.

In order to validate the findings compiled here, a demonstrator is used. The project “Connect & Collect: AI-based Cloud for Interdisciplinary Networked Research and Innovation for Future Work (CoCo)” (CoCo Website, 2023) promotes the transfer of knowledge between HFE research and operational practice in companies about artificial intelligence and is funded by the German Federal Ministry of Education and Research (BMBF). To support our theory development, the CoCo project serves as a demonstrator to illustrate the interrelation between the different disciplines. The purpose of the demonstrator is to reveal the necessity of each discipline, namely HFE, psychology, HCI, information science, and adult education, to create successful HCAI implementation at work and, therefore, to show the impact and potential in society. Participants are predominantly transdisciplinary actors in labor research from science, enterprises, unions, education, and intermediaries pursuing an innovative new approach or applying best practices and have joined forces in “Regional Competence Centers for Labor Research.”

It can be derived from the demonstrator that while many companies are interested in implementing AI applications, they shy away from the research and investment effort involved in developing company-specific solutions. Instead, they aim to use proven AI applications. In this case, the importance of human-centered AI design—especially regarding learning facilitation and ethical-social compatibility—is not sufficiently applied (Pokorni et al., 2021). Undesirable consequences usually emerge only after a time delay and are rarely causally associated with AI use. An essential part of the work of the demonstrator is to increase the relevance of human-centered design of AI applications practically and systematically, which will also enhance the role of HFE experts.

At the same time, it is crucial to qualify the workers for the changes in the working society. The examples from the demonstrator show that, when implementing AI systems on a large scale, it is important to develop and establish a broad range of learning opportunities that can be upscaled to diverse target groups. Within the digital transformation, especially marginalized groups are at risk of getting lost in terms of workforce, labor markets, and in a democratic society. There is a need to focus on these marginalized groups when referring to human centeredness. Adult education providers and unions have an enormous role in this challenge because they have expertise in accessing the marginalized.

Human-centered AI is strongly related to HCI and psychology, albeit with different emphases. As HCI has evolved from a mainly

technical field to an interdisciplinary profession, the same can be expected for HCAI. Nevertheless, there are critical challenges to overcome, like explainability, trustworthiness, or unrealistic user expectations. However, technology acceptance models can be used to build trust to accept AI systems and make them more human-centric, thus keeping expectations in line. Ultimately, achieving an explainable AI contributes to the mutual collaboration and interaction of humans and AI. For ethical reasons, AI's instrumental character must always be considered. Furthermore, AI can be made more accessible to stakeholders by implementing gamification, which can increase stakeholder commitment to increased engagement with human-centered AI, especially in the workplace. Finally, the demonstrator acknowledges the importance of ethical concerns by utilizing the human-in-the-loop concept, thus increasing trust.

It can be concluded that the transfer of human-centered research results into practical application is supported by an interdisciplinary approach that combines different ideas, knowledge, and work methods. By displaying the interrelationships between the disciplines, this article reveals in the following section further directions for research as well as concepts and features a (future) theory of HCAI should entail.

### 6.3. Setting foundations for a human-centered interdisciplinary AI theory: synergistic human-AI symbiosis theory (SHAST)

In order to advance the research field regarding human-centered AI, a conceptualization of a Synergistic Human-AI Symbiosis Theory (SHAST) has been started. SHAST takes into account that the optimal deployment of artificial intelligence in the workplace needs the establishment of a symbiotic relationship between humans and AI systems, drawing upon the expertise of five distinct disciplines: human factors engineering (HFE), human-computer interaction (HCI), psychology, information science, and adult education. Based on the present findings, all five disciplines appear to be necessary to successfully implement HCAI in the workplace. SHAST envisions a future where AI systems and humans collaborate synergistically to achieve unprecedented levels of productivity and wellbeing. The five disciplines presented here are predestined to contribute to the future of HCAI, as they incorporate fundamental components of human centeredness and address important fields of action for its creation and implementation.

SHAST posits that AI should be designed to augment human capabilities, foster seamless interactions, and ensure ethical practices. HFE is the foundational pillar, advocating for AI systems that enhance human potential while preserving human autonomy, decision-making, and overall work performance. HCI, another elementary bedrock, focuses on user-centric design, creating seamless and intuitive interactions between humans and AI, fostering realistic user expectations, and minimizing friction in collaborating with AI technologies that intuitively adapt to user needs and preferences. Psychology's role in SHAST revolves

TABLE 2 Framework with minimum fields of action for the disciplines.

Discipline	Minimum fields of action
HFE	• Balanced workload
	• Enhancing human capabilities
	• Human control
	• Social and ethical responsibility
Information science	• Pragmatics of information
	• Consider contexts and information behavior
	• Information literacy
Adult education	• Include low-skilled workers and workers' unions
	• Tailored learning and teaching in courses and the workplace
HCI and Psychology	• Technology acceptance and adoption
	• Human autonomy and control
	• Realistic user expectations
	• Explainable AI
	• Trust

around cultivating user trust, achieved through transparent AI design and explainable algorithms. Information science considers the pragmatic side of information, ensuring that humans can effectively and efficiently use information systems, for example, by increasing their information literacy. Finally, adult education plays a critical part in SHAST by cultivating digital literacy and ensuring that individuals possess the skills to navigate AI-powered environments, fostering learning opportunities for a workforce to engage with AI technologies for innovation and productivity effectively, minimizing disparities, and enabling broad participation.

SHAST proposes that human-AI symbiosis can be achieved through an interplay of these disciplines, resulting in AI technologies that empower individuals, enhance collaboration, and create a sustainable and equitable future. In fact, SHAST is based on a framework that is presented in Table 2. The outcomes from theory and practice show minimum fields of action for a successful HCAI implementation.

If fundamental aspects from the framework of the five disciplines are missing for the implementation and application of HCAI, then this will result in severe consequences and challenges (Stephanidis et al., 2019; Garibay et al., 2023; Xu et al., 2023). More precisely, it is postulated that when a discipline is not adequately considered, there may be a failure to comply with the minimum fields of action, and implementation may not be successful.

It remains open for discussion and empirical research on which disciplines would further be needed to support the development of a widely applicable theory of human-centered AI. The article aims to convincingly present the normative, theoretical, and methodological concepts from human factors and ergonomics (HFE), psychology, human-computer interaction, information science, and adult education and why they are considered critical building blocks for HCAI and SHAST.

## 7. Conclusions for human-centered AI at work from theory and practice

In the last section, the article summarizes how it contributes to the ongoing scientific discussion on HCAI. Conclusions are drawn for HCAI at work by leveraging insights from disciplines focusing mainly on individuals (psychology), technology (HCI), work (HFE), or work context as one research field amongst others (information science and adult learning). From the fundamental disciplinary considerations and the current experiences and observations from the demonstrator, theoretical implications are derived to bring HCAI in line with today's demands of workers and companies to reflect human centeredness when dealing with the complexity of information, data, and decisions. Furthermore, the article also highlights the relevant internal logic of the individual disciplines and reveals possible mutual complementarity. It mainly argues that, although human-centered AI is a popular concept across disciplines today (Capel and Brereton, 2023), successful HCAI and its design are in strong need for an interdisciplinary approach, as all disciplines conceptualize “their humans” differently in their views and methodological approaches.

Besides the differences in detail, our comparison of the different disciplinary approaches to conceptualizing HCAI and bringing it into practice has revealed important similarities. All presented disciplines have the human at their cores — the development of human-centered AI systems is therefore deeply connected to aspects considered central to humans and that cannot be substituted with machines, such as learning and constructing information. Technology, along with its AI systems, is considered a means for human development—therefore putting the human in a superior position, which also has to be guaranteed by certain regulations to prevent AI from overruling human decision-making. The disciplines also widely agree that human-centricity can only be achieved if struggles with the use of technology are minimized. In the end, it is always a human being in a company who will use technology. Depending on the degree of how much the AI design and AI implementation process reflects the users, including their skills, trust, and their work tasks — the users will cope or fail with technology within their workplace. Hence, raising the issue of humans practically coping or failing with technology use in the workplace as a common theoretical and practical problem of human centeredness might open scientific discourse between disciplines and practice in companies (Nowotny et al., 2001).

This article is based on collaborative scholarly inquiry and a review of discipline-specific literature on the phenomenon of HCAI and joint reflection against the background of the demonstrator used in our research to study HCAI. This led to the findings and arguments in this article, which are summarized in Table 2. The article combines perspectives from five disciplines—however, we are still in the first step, i.e., conceptualization, of a five-phase model for theory building in the applied sciences (Swanson and Chermack, 2013). The need to include further or remove disciplines will become obvious further down the road of the Theory-Research-Practice Cycle that is followed and will include operationalization, confirmation, application, and refinement (Swanson and Chermack, 2013) of the concept we have come up with. Hence, this research is not finished but at

the beginning of understanding what might be critical for HCAI at work.

The first results of the conceptualization stage concerned with successful HCAI implementation at work and with the common focal points of the five disciplines are the founding pillars for a Synergistic Human-AI Symbiosis Theory (SHAST), which answers the research question of what is critical for HCAI at work. It has become apparent, however, that determining the critical aspects of HCAI at work, enabling a symbiotic relationship between humans and AI systems, and describing their interplay theoretically is a complex task. Whether the challenge does not rather require investigation of more minimum fields of action, more disciplinary backgrounds like, e.g., knowledge management, economics, process management, simulation theory, operation research, philosophy, history, sociology, and many more, or a transdisciplinary approach (Defila and Di Giulio, 2018) involving all stakeholders in the research, development, and implementation process of HCAI at work remains an open question. It has been argued that HCAI and human centeredness will benefit from considering many different scientific perspectives.<sup>1</sup> Furthermore, in today's world, disciplinary boundaries are fluid and increasingly blurred, or new disciplines are formed that encompass various aspects of other classical disciplines. We are aware of this, but due to a simplified discussion, we limit ourselves to a few disciplines and outline them, knowing that these boundaries are artificial.

The results from the demonstrator and our collaborative inquiry have shown that the disciplines covered in this article have many similar (and some different) concepts and methods in their portfolios that are most likely critical for successfully implementing HCAI at work. However, a broader interdisciplinary discussion and research are needed for a complete view of the topic.

## Data availability statement

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding authors.

## Author contributions

AM: Conceptualization, Writing—original draft, Writing—review & editing, Funding acquisition, Supervision, Methodology, Project administration, Validation, Visualization. CB-S: Conceptualization, Writing—original draft, Writing—review & editing, Methodology, Project administration, Validation, Visualization. MB: Conceptualization, Writing—original draft, Writing—review & editing, Methodology, Project administration, Validation, Visualization. IP: Conceptualization, Writing—original draft, Writing—review & editing, Methodology, Project administration, Validation, Visualization.

<sup>1</sup> <https://www.frontiersin.org/research-topics/50257/human-centered-ai-at-work-common-ground-in-theories-and-methods>

## Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. The CoCo project is funded by the Federal Ministry of Education and Research-grant numbers 02L19C000, 02L19C001, and 02L19C002.

## Acknowledgments

We are grateful for the valuing and precise feedback of the reviewers. It helped us to improve the present article and to elaborate on its line of argument.

## References

- ACRL (2016). *Framework for Information Literacy for Higher Education*. Chicago, IL: Association of College and Research Libraries.
- Alheit, P. (2021). Biographicity as “mental grammar” of postmodern life. *Eur. J. Res. Educ. Learn. Adults* 12, 81–94. doi: 10.3384/rela.2000-7426.ojs1845
- Amershi, S., Weld, D., Vorvoreanu, M., Fourney, A., Nushi, B., Collisson, P., et al. (2019). “Guidelines for human-AI interaction,” in *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems CHI '19* (New York, NY: ACM), 1–13.
- Anandiadou, K., and Claro, M. (2009). *21st Century Skills and Competences for New Millennium Learners in OECD Countries, OECD Education Working Papers, No. 41*. Washington, DC: OECD Publishing.
- Anderson, T., and Rivera-Vargas, P. (2020). A critical look at educational technology from a distance education perspective. *Digital Educ. Rev.* 37, 208–229. doi: 10.1344/der.2020.37.208-229
- Arbeiter, J. (2023). Acting between professional rules and financial resources in the field of work-oriented basic education. *Zeitschrift für Weiterbildungsforschung* 46, 87–102. doi: 10.1007/s40955-023-00240-2
- Autor:innengruppe Bildungsberichterstattung (2022). *Bildung in Deutschland 2022. Ein Indikatorengeprägter Bericht mit einer Analyse zum Bildungspersonal*. Bielefeld: wbv Media.
- Bagozzi, R. P. (2007). The legacy of the technology acceptance model and a proposal for a paradigm shift. *J. Assoc. Inf. Syst.* 8, 244–254. doi: 10.17705/1jais.00122
- Bates, M. J. (2017). “Information behavior,” in *Encyclopedia of Library and Information Sciences. 4th Edn*, eds. J. McDonald and M. Levine-Clark (Boca Raton: CRC Press), 2074–2085.
- Bauer, J., Festner, D., Gruber, H., Harteis, C., and Heid, H. (2004). The effects of epistemological beliefs on workplace learning. *J. Workpl. Learn.* 16, 284–292. doi: 10.1108/13665620410545561
- Baumeister, H., Garatva, P., Pryss, R., Ropinski, T., and Montag, C. (2023). Digitale Phänotypisierung in der Psychologie - ein Quantensprung in der psychologischen Forschung? *Psychol. Rundsch.* 74, 89–106. doi: 10.1026/0033-3042/a000609
- Baxter, G., and Sommerville, I. (2011). Socio-technical systems: from design methods to systems engineering. *Interacting Comput.* 23, 4–17. doi: 10.1016/j.intcom.2010.07.003
- Bender, E., and Kollé, A. (2020). “Climbing towards NLU—on meaning, form, and understanding in the age of data,” in *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, Seattle, WA: Association for Computational Linguistics, 5185–5198.
- Bennett, D., Metatla, O., Roudaut, A., and Mekler, E. D. (2023). “How does HCI understand human agency and autonomy?” in *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems CHI '23* (New York, NY: Association for Computing Machinery), 1–18.
- Bermúdez, J. P., Nyrup, R., Deterding, S., Mougnot, C., Moradbakhti, L., You, F., et al. (2023). “What is a subliminal technique? An ethical perspective on AI-driven influence,” in *Proceedings of the IEEE ETHICS-2023 Conference* (West Lafayette, IN: IEEE), 1–9.
- BMAS and BMBF (2021). *Nationale Weiterbildungsstrategie*. Available online at: <https://www.bmas.de/SharedDocs/Downloads/DE/Aus-Weiterbildung/nws->

## Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

[fortfuhrung-und-weiterentwicklung.pdf?\\_\\_blob=publicationFile&v=3](#) (accessed June 29, 2023).

BMBF (2022). *Weiterbildungsverhalten in Deutschland 2020. Ergebnisse des Adult Education Survey—AES-Trendbericht*. Bonn: BMBF.

Böhle, F. (2009). “Weder Rationale Reflexion noch Präreflexive Praktik,” in *Erfahrungsgelitet-Subjektivierendes Handeln, Handeln unter Unsicherheit*, eds. F. Böhle and M. Wehrich (Wiesbaden: Springer VS), 203–230.

Bostrom, N., and Yudkowsky, E. (2013). “The ethics of artificial intelligence,” in *Cambridge Handbook of Artificial Intelligence*, eds. K. Frankish, M. Willian, and W. M. Ramsey (Cambridge: University Press), 316–334.

Bräuer, P., and Mazarakis, A. (2022a). “Alexa, can we design gamification without a screen?—Implementing cooperative and competitive audio-gamification for intelligent virtual assistants. *Comput. Hum. Behav.* 135, 107362. doi: 10.1016/j.chb.2022.107362

Bräuer, P., and Mazarakis, A. (2022b). How to design audio-gamification for language learning with amazon Alexa? a long-term field experiment. *Int. J. Hum. Comput. Int.* 12, 1–18. doi: 10.1080/10447318.2022.2160228

Braun, M. (2017). “Arbeit 4.0: Der gesunde Mensch in der digitalisierten Arbeitswelt,” in *Handbuch der Arbeitsmedizin*, eds. D. Nowak, and S. Letzel (Landsberg: Ecomed), 1–24.

Brynjolfsson, E., and McAfee, A. (2012). *Race Against the Machine: How the Digital Revolution is Accelerating Innovation, Driving Productivity, and Irreversibly Transforming Employment and the Economy*. Lexington MA: Digital Frontier Press.

Büchmann, O. (2020). Künstliche Intelligenz und Ethik—ein ungleiches Paar? *Wirtschaftsinformatik Manage.* 12, 206–215. doi: 10.1365/s35764-020-00256-0

Capel, T., and Brereton, M. (2023). “What is human-centered about human-centered AI? A map of the research landscape,” in *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems CHI '23* (New York, NY: Association for Computing Machinery), 1–23. doi: 10.1145/3544548.3580959

Case, D. O., and Given, L. M. (2016). *Looking for Information: A Survey of Research on Information Seeking, Needs, and Behavior, 4th Edn*. Bingley: Emerald Group Publishing.

Chai, C. S., Lin, P. Y., Jong, M. S. Y., Dai, Y., Chiu, T. K., and Huang, B. (2020). “Factors influencing students’ behavioral intention to continue artificial intelligence learning,” in *2020 International Symposium on Educational Technology (ISET)*. Piscataway, NJ: IEEE.

Choung, H., David, P., and Ross, A. (2023). Trust in AI and its role in the acceptance of AI technologies. *Int. J. Hum. Comput. Int.* 39, 1727–1739. doi: 10.1080/10447318.2022.2050543

Clemmensen, T. (2006). Whatever happened to the psychology of human-computer interaction? *Inf. Technol. People* 19, 121–151. doi: 10.1108/09593840610673793

CoCo Website (2023). *Connect & Collect: KI-Gestützte Cloud für die Interdisziplinäre vernetzte Forschung und Innovation für die Zukunftsarbeit*. Available online at: <https://www.coco-projekt.de/> (accessed February 4, 2023).

Council of the European Union (2021). *Council Resolution on a New European Agenda for Adult Learning 2021-2030 (2021/C 504/02)*.

Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Q.* 21, 319–340. doi: 10.2307/249008

- Defila, R., and Di Giulio, A. (2018). *Transdisziplinär und Transformativ Forschen*. Wiesbaden: Springer Fachmedien.
- Dehnpostel, P. (2022). *Betriebliche Bildungsarbeit. Kompetenzbasierte Berufs- und Weiterbildung in digitalen Zeiten*. Baltmannsweiler: Schneider Verlag Hohengehren.
- Del Giudice, M., Scuotto, V., Orlando, B., and Mustilli, M. (2023). Toward the human-centered approach: a revised model of individual acceptance of AI. *Hum. Resour. Manag. Rev.* 33, 100856. doi: 10.1016/j.hrmr.2021.100856
- Deterding, S., Dixon, D., Khaled, R., and Nacke, L. (2011). "From Game Design Elements to Gamefulness: Defining "Gamification." in *Proceedings of the 15th International Academic MindTrek Conference: Envisioning Future Media Environments* (London: ACM Press), 9–15.
- Dollhausen, K., and Lattke, S. (2020). "Organisation und Organisationsformen wissenschaftlicher Weiterbildung," in *Handbuch wissenschaftliche Weiterbildung*, eds. W. Jütte and M. Rohs (Wiesbaden: Springer VS), 99–121.
- Dyckhoff, H. (2006). *Produktionstheorie, 5th Edn*. Berlin: Springer.
- Elbeshausen, S. (2023). "C 10 Modellierung von Benutzer\*innen, Kontextualisierung, Personalisierung," in *Grundlagen der Informationswissenschaft, 7th Edn*, eds. R. Kuhlen, D. Lewandowski, W. Semar and C. Womser-Hacker (Berlin: De Gruyter Saur), 467–476.
- European Union (2021). *Council Resolution on a New European Agenda for Adult Learning 2021-2030 (2021/C 504/02)*. Official Journal of the European Union. Available online at: [https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32021G1214\(01\)](https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32021G1214(01)) (accessed August 28, 2023).
- Floridi, L., Cows, J., Beltrametti, M., Chatila, R., Chazerand, P., Dignum, V., et al. (2018). AI4People—An ethical framework for a good AI society: opportunities, risks, principles, and recommendations. *Mind Mach* 28, 689–707. doi: 10.1007/s11023-018-9482-5
- Garibay, O. O., Winslow, B., Andolina, S., Antona, M., Bodenschatz, A., Coursaris, C., et al. (2023). Six human-centered artificial intelligence grand challenges. *Int. J. Hum. Comput. Int.* 39, 391–437. doi: 10.1080/10447318.2022.2153320
- Gericke, E. (2017). Why returning to VET? – Results of a qualitative comparative study about English and German Mechatronics. *International Journal for Res. Vocat. Educ. Train.* 4, 206–225. doi: 10.13152/IJRVET.4.3.2
- Golmann, R., Hagmann, D., and Loewenstein, G. (2017). Information avoidance. *J. Econ. Lit.* 55, 96–135. doi: 10.1257/jel.20151245
- Goppold, M., and Frenz, M. (2020). Lernen im Prozess der Arbeit: Entwicklung, Umsetzung und Evaluation einer Weiterbildung zur manuellen Montage unter Einsatz von Autorensystemen. *Zeitschrift für Arbeitswissenschaft* 74, 100–116. doi: 10.1007/s41449-020-00202-1
- Greifeneder, E., and Schlebke, K. (2023). "D 1 Information Behaviour. in Grundlagen der Informationswissenschaft," in *Grundlagen der Informationswissenschaft, 7th edn.*, eds. R. Kuhlen, D. Lewandowski, W. Semar and C. Womser-Hacker (Berlin, Boston: De Gruyter Saur), 497–510.
- Griesbaum, J. (2023). "D 8 Informationskompetenz," in *Grundlagen der Informationswissenschaft, 7th Edn.*, eds. R. Kuhlen, D. Lewandowski, W. Semar, and C. Womser-Hacker (Berlin: De Gruyter Saur), 581–594.
- Grote, G., Ryser, C., Wäfler, T., Windischer, A., and Weik, S. (2000). Compass-a method for complementary function allocation in automated work systems. *Int. J. Hum. Comput. Stu.* 52, 267–287. doi: 10.1006/ijhc.1999.0289
- Gust von Loh, S. (2008). Wissensmanagement und Informationsbedarfsanalyse in kleinen und mittleren Unternehmen: Teil 2: Wissensmanagement in KMU. *Inf.-Wissenschaft Praxis* 59, 127–135.
- Habayeb, A. (2022). *Explainable AI Isn't Enough; We Need Understandable AI*. Techopedia. Available online at: <https://www.techopedia.com/explainable-ai-isnt-enough-we-need-understandable-ai/2/34671> (accessed June 24, 2023).
- Hacker, W. (2018). *Menschengerechtes Arbeiten in der Digitalisierten Welt*. Zürich: VDF.
- Haghighat, P., Nguyen, T., Valizadeh, M., Arvan, M., Parde, N., Kim, M., et al. (2023). Effects of an intelligent virtual assistant on office task performance and workload in a noisy environment. *Appl. Ergon.* 109, 103969. doi: 10.1016/j.apergo.2023.103969
- Harteis, C. (2022). "Research on workplace learning in times of digitalisation," in *Research Approaches on Workplace Learning: Insights from a Growing Field*, eds. C. Harteis, D. Gijbels, and E. Kynndt (Cham: Springer International Publishing) 415–428. doi: 10.1007/978-3-030-89582-2
- Hartel, J. (2019). *Turn, Turn, Turn, Proceedings of the Tenth International Conference on Conceptions of Library and Information Science. Ljubljana (Slovenia), June 16th-19th*. <https://informationr.net/ir/24-4/colis/colis1901.html> (accessed June 29, 2023).
- Henrichs, N. (2014). "Sozialisation der information. Zum aufgabenspektrum der informationswissenschaft," in *Menschsein im Informationszeitalter. Informationswissenschaft mit Leidenschaft und Missionarischem Eifer*, ed. N. Henrichs (Glückstadt: Verlag Werner Hülsbusch).
- Höök, K. (2000). Steps to take before intelligent user interfaces become real. *Interact. Comput.* 12, 409–426. doi: 10.1016/S0953-5438(99)00006-5
- Huchler, N. (2022). Komplementäre Arbeitsgestaltung. Grundrisse eines Konzepts zur Humanisierung der Arbeit mit KI. *Zeitschrift für Arbeitswissenschaft* 76, 158–175. doi: 10.1007/s41449-022-00319-5
- Huchler, N., Adolph, L., Andre, E., Bauer, W., Reißner, N., Müller, N., et al. (2020). *Kriterien für die Mensch-Maschine-Interaktion bei KI. Ansätze für die menschengerechte Gestaltung in der Arbeitswelt-Whitepaper aus der Plattform Lernende Systeme*. München: ACATECH.
- IEA (2000). *What is Ergonomics?* Available online at: <https://iea.cc/about/what-is-ergonomics/> (accessed July 11, 2023).
- ILO and IEA (2021). *Principles and Guidelines for (HFE) Design and Management of Work Systems*. Geneva: Joint Document by ILO and IEA.
- Infratest dimap (2023). *KI ist eher Risiko als Chance, nicht jedoch für Junge und Bessergebildete. Umfragen Anal.-Infratest Dimap*. Available online at: <https://www.infratest-dimap.de/umfragen-analysen/bundesweit/umfragen/aktuell/kritischer-ausblick-auf-deutsch-tuerkische-beziehungen-bei-wiederwahl-erdogans/> (accessed May 12, 2023).
- Ingwersen, P., and Järvelin, K. (2005). *The Turn: Integration of Information Seeking and Retrieval in Context*. Cham: Springer Science and Business Media.
- Ipsos (2022). *Opinions About AI Vary Depending on Countries' Level of Economic Development*. Ipsos. Available online at: <https://www.ipsos.com/en/global-opinions-about-ai-january-2022> (accessed May 13 2023).
- ISO 6385 (2016). *Ergonomics Principles in the Design of Work Systems*.
- ISO 9241-110 (2020). *Ergonomics of Human-System Interaction - Part 110: Interaction Principles*.
- ISO 9241-2 (1992). *Ergonomic Requirements for Office Work With Visual Display Terminals (VDTs) - Part 2: Guidance on Task Requirements*.
- ISO 9241-210 (2019). *Ergonomics of Human-System Interaction - Part 210: Human-Centred Design for Interactive Systems*.
- Jetter, H. C. (2023). "D 3 Mensch-Computer-Interaktion, Usability und User Experience," in *Grundlagen der Informationswissenschaft, 7th Edn*, eds. R. Kuhlen, D. Lewandowski, W. Semar, and C. Womser-Hacker (Berlin, Boston: De Gruyter Saur), 525–534.
- Joint Research Center (JRC) (2022). *DigComp 2.2: The Digital Competence Framework for Citizens-With New Examples of Knowledge, Skills and Attitudes*. EU Publications. doi: 10.2760/115376
- Jones, W. (2010). No knowledge but through information. *First Monday* 15, 3602. doi: 10.5210/fm.v15i9.3062
- Kao, W.-K., and Huang, Y.-S. (2023). Service robots in full-and limited-service restaurants: extending technology acceptance model. *J. Hosp. Tour. Manag.* 54, 10–21. doi: 10.1016/j.jht.2022.11.006
- Käpplinger, B., and Lichte, N. (2020). The lockdown of physical co-operation touches the heart of adult education: a delphi study on immediate and expected effects of COVID-19. *Int. Rev. Educ.* 66, 777–795. doi: 10.1007/s11159-020-09871-w
- Karwowski, W., Szopa, A., and Soares, M. (2021). *Handbook of Standards and Guidelines in Human Factors and Ergonomics*. Boca Raton, WA: CRC Press.
- Kosinski, M., Bachrach, Y., Kohli, P., Stillwell, D., and Graepel, T. (2014). Manifestations of user personality in website choice and behaviour on online social networks. *Mach. Learn.* 95, 357–380. doi: 10.1007/s10994-013-5415-y
- Kraus, K. (2008). "Does employability put the german 'vocational order' at risk? An analysis from the perspective of earning oriented pedagogy," in *Work, Education and Employability*, eds. P. Gonon, K. Kraus, J. Oelkers, and S. Stolz Dernbach (Bern: Peter Lang), 55–81.
- Krishnan, C., Gupta, A., Gupta, A., and Singh, G. (2022). "Impact of artificial intelligence-based chatbots on customer engagement and business growth," in *Deep Learning for Social Media Data Analytics Studies in Big Data.*, eds. T.-P. Hong, L. Serrano-Estrada, A. Saxena, and A. Biswas (Cham: Springer International Publishing), 195–210.
- Kuhlen, R. (2004). *Information Grundlagen der praktischen Information und Dokumentation, 5th Edn*. München: Saur, 3–20.
- Kurni, M., Mohammed, M. S., and Srinivasa, K. G. (2023). "AI-enabled gamification in education," in *A Beginner's Guide to Introduce Artificial Intelligence in Teaching and Learning*, eds. M. Kurni, M. S. Mohammed, and S. K G (Cham: Springer), 105–114.
- Laato, S., Tiainen, M., Najmul Islam, A. K. M., and Mäntymäki, M. (2022). How to explain AI systems to end users: a systematic literature review and research agenda. *Int. Res.* 32, 1–31. doi: 10.1108/INTR-08-2021-0600
- Lammers, A., Lukowski, F., and Weis, K. (2022). The relationship between works councils and firms' further training provision in times of technological change. *Br. J. Ind. Relat.* 61, 392–424. doi: 10.1111/bjir.12710
- Landgrebe, J., and Smith, B. (2022). *Why Machines Will Never Rule the World: Artificial Intelligence Without Fear*. Abingdon: Routledge.

- Langer, M., Oster, D., Speith, T., Hermanns, H., Kästner, L., Schmidt, E., et al. (2021). What do we want from explainable artificial intelligence (XAI)?—A stakeholder perspective on Xai and a conceptual model guiding interdisciplinary xai research. *Artif. Int.* 296, 103473. doi: 10.1016/j.artint.2021.103473
- Leichtmann, B., Hinterreiter, A., Humer, C., Streit, M., and Mara, M. (2023a). Explainable artificial intelligence improves human decision-making: results from a mushroom picking experiment at a public art festival. *Int. J. Hum. Comput. Int.* 12, 1–18. doi: 10.1080/10447318.2023.2221605
- Leichtmann, B., Humer, C., Hinterreiter, A., Streit, M., and Mara, M. (2023b). Effects of explainable artificial intelligence on trust and human behavior in a high-risk decision task. *Comput. Hum. Behav.* 139, 107539. doi: 10.1016/j.chb.2022.107539
- Lenzen, M. (2019). *Künstliche Intelligenz: Was sie kann & was uns erwartet*. München: Beck.
- Lewandowski, D., and Womser-Hacker, C. (2023). “D 6 information seeking behaviour” in *Grundlagen der Informationswissenschaft, 7th Edn*, eds. R. Kuhlen, D. Lewandowski, W. Semar, and C. Womser-Hacker (Berlin: De Gruyter Saur), 553–566.
- Lloyd, A. (2013). “Building information resilient workers: the critical ground of workplace information literacy. What have we learnt?” in *Worldwide Commonalities and Challenges in Information Literacy Research and Practice: European Conference on Information Literacy, ECIL 2013*, Istanbul: Springer International Publishing, 219–228.
- Luger, E., and Sellen, A. (2016). ““Like having a really bad PA”: the gulf between user expectation and experience of conversational agents,” in *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems CHI '16*. New York, NY: Association for Computing Machinery (ACM), 5286–5297.
- Marsden, N., and Haag, M. (2016). “Stereotypes and politics: reflections on personas,” in *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems CHI '16*. New York, NY: Association for Computing Machinery (ACM), 4017–4031.
- Mazarakis, A. (2013). Like diamonds in the sky—how feedback can boost the amount of available data for learning analytics. *Int. J. Technol. Enhanc. Learn.* 5, 107–116. doi: 10.1504/IJTEL.2013.059083
- Mazarakis, A. (2021). Gamification reloaded: current and future trends in gamification science. *i-com* 20, 279–294. doi: 10.1515/icom-2021-0025
- McCarthy, J. (2007). *What is Artificial Intelligence? Stanford University*. Available online at: <https://www-formal.stanford.edu/jmc/whatisai.pdf> (accessed August 29, 2023).
- McLean, S., Read, G. J., Thompson, J., Baber, C., Stanton, N. A., and Salmon, P. M. (2021). The risks associated with artificial general intelligence: a systematic review. *J. Exp. Theor. Artif. Int.* 24, 1–15. doi: 10.1080/0952813X.2021.1964003
- Merriam, S. B., and Bierema, L. L. (2013). *Adult Learning: Linking Theory and Practice*. London: John Wiley & Sons.
- Middleton, L., Hall, H., Muir, L., and Raeside, R. (2018). The interaction between people, information and innovation: information literacy to underpin innovative work behaviour in a Finnish organization. *Proc. Assoc. Inf. Sci. Technol.* 55, 367–376. doi: 10.1002/pr2.2018.14505501040
- Miller, T. (2019). Explanation in artificial intelligence: insights from the social sciences. *Artif. Int.* 267, 1–38. doi: 10.1016/j.artint.2018.07.007
- Ng, D. T. K., Leung, J. K. L., Chu, K. W. S., and Qiao, M. S. (2021a). AI literacy: definition, teaching, evaluation and ethical issues. *Proc. Assoc. Inf. Sci. Technol.* 58, 504–509. doi: 10.1002/pr2.487
- Ng, D. T. K., Leung, J. K. L., Chu, S. K. W., and Qiao, M. S. (2021b). Conceptualizing AI literacy: an exploratory review. *Comput. Educ. Artif. Int.* 2, 100041. doi: 10.1016/j.caeai.2021.100041
- Nowotny, H., Scott, P. B., and Gibbons, M. (2001). *Re-Thinking Science. Knowledge and the Public in an Age of Uncertainty*. Cambridge, MA: Polity.
- OECD (2023). *Equity and Inclusion in Education: Finding Strength through Diversity*. Paris: OECD Publishing.
- OECD. (2019). *Artificial Intelligence in Society*. Paris: OECD Publishing.
- Pirolli, P., and Card, S. (1999). Information foraging. *Psychol. Rev.* 106, 643–675. doi: 10.1037/0033-295X.106.4.643
- Pokorni, B., Braun, M., and Knecht, C. (2021). *Menschzentrierte KI-Anwendungen in der Produktion. Praxis-erfahrungen und Leitfaden zu Betrieblichen Einführungsstrategien. Projektbericht im Fortschrittszentrum “Lernende Systeme”*. Stuttgart: Fraunhofer IAO.
- Raftopoulos, M., and Hamari, J. (2023). “Artificial intelligence in the workplace: implementation challenges and opportunities,” in *Proceedings of the Americas Conference on Information Systems (AMCIS) 2023*. Cancun: AMCIS.
- Rammert, W. (2009). “Hybride Handlungsträgerschaft: Ein soziotechnisches Modell verteilten Handelns,” in *Intelligente Objekte*, eds. O. Herzog and T. Schildhauer (Berlin: Springer), 23–33.
- Rauch, W. (1988). *Was ist Informationswissenschaft?* Graz: Kienreich.
- Reddy, M. C., and Jansen, B. J. (2008). A model for understanding collaborative information behavior in context: a study of two healthcare teams. *Inf. Proc. Manage.* 44, 256–273. doi: 10.1016/j.ipm.2006.12.010
- Riege, A. (2005). Three-dozen knowledge-sharing barriers managers must consider. *J. Knowl. Manag.* 9, 18–35. doi: 10.1108/13673270510602746
- Rohs, M., and Ganz, M. (2015). MOOCs and the claim of education for all: a disillusion by empirical data. *Int. Rev. Res. Open Distrib. Learn.* 16, 1–19. doi: 10.19173/irrodl.v16i6.2033
- Schlick, C., Bruder, R., and Luczak, H. (2018). *Arbeitswissenschaft, 4th Edn*. Berlin, Heidelberg: Springer Vieweg.
- Schnabel, D. (2017). *Kompetenzen für die Arbeitswelt Von Heute und Morgen: 21st Century Skills and Beyond*. Available online at: [https://hochschulforumdigitalisierung.de/de/blog/kompetenzen-21st-century-skills#:~:sim\\$%text=Das%2021st%20Century%20Skills%20Modell%20deckt%20vier%20Kompetenzfelder,Kritisches%20Denken%204%20Flexibilit%C3%A4t%2C%20Ambiguit%C3%A4tstoleranz%2C%20Eigenmotivation%2C%20Selbst%C3%A4ndiges%20Arbeiten](https://hochschulforumdigitalisierung.de/de/blog/kompetenzen-21st-century-skills#:~:sim$%text=Das%2021st%20Century%20Skills%20Modell%20deckt%20vier%20Kompetenzfelder,Kritisches%20Denken%204%20Flexibilit%C3%A4t%2C%20Ambiguit%C3%A4tstoleranz%2C%20Eigenmotivation%2C%20Selbst%C3%A4ndiges%20Arbeiten) (accessed June 29, 2023).
- Schön, D. (1983). *The Reflective Practitioner: How Professionals Think in Action*. New York, NY: Basic Books.
- Shannon, C. (2001). A mathematical theory of communication. *ACM SIGMOBILE Mob. Comput. Commun. Rev.* 5, 3–55. doi: 10.1145/584091.584093
- Shneiderman, B. (2020). Human-centered artificial intelligence: three fresh ideas. *AIS Trans. Hum.-Comput. Interact.* 12, 109–124. doi: 10.17705/1thci.00131
- Shneiderman, B. (2022). *Human-Centered AI*. New York, NY: Oxford University Press.
- Sohn, K., and Kwon, O. (2020). Technology acceptance theories and factors influencing artificial intelligence-based intelligent products. *Telemat. Inform.* 47, 101324. doi: 10.1016/j.tele.2019.101324
- Spitzley, H. (1980). *Wissenschaftliche Betriebsführung, REFA-Methodenlehre und Neuorientierung der Arbeitswissenschaft*. Frankfurt: Bund.
- Stephanidis, C., Salvendy, G., Antona, M., Chen, J. Y. C., Dong, J., Duffy, V. G., et al. (2019). Seven HCI grand challenges. *Int. J. Hum. Comput. Int.* 35, 1229–1269. doi: 10.1080/10447318.2019.1619259
- Stock, W., and Stock, M. (2015). *Handbook of Information Science*. Berlin: De Gruyter Saur.
- Swanson, R. A., and Chermack, T. J. (2013). *Theory Building in Applied Disciplines*. San Francisco, CA: Berrett-Koehler Publishers.
- Tan, D. Y., and Cheah, C. W. (2021). Developing a gamified AI-enabled online learning application to improve students’ perception of university physics. *Comput. Educ. Artif. Intell.* 2, 100032. doi: 10.1016/j.caeai.2021.100032
- Tang, P. M., Koopman, J., Mai, K. M., De Cremer, D., Zhang, J. H., Reynders, P., et al. (2023). No person is an island: unpacking the work and after-work intelligence. *J. Appl. consequences of interacting with artificial Psychol.* 14, 1–24. doi: 10.1037/apl00.01103
- Tocchetti, A., Corti, L., Brambilla, M., and Celino, I. (2022). EXP-crowd: a gamified crowdsourcing framework for explainability. *Front. Artif. Intell.* 5, 826499. doi: 10.3389/frai.2022.826499
- Touretzky, D., Gardner-McCune, C., Martin, F., and Seehorn, D. (2019). Envisioning AI for K-12: what should every child know about AI? *Proc. AAAI Conf. Artif. Int.* 33, 9795–9799. doi: 10.1609/aaai.v33i01.33019795
- Travis, T. A. (2017). From the classroom to the boardroom: the impact of information literacy instruction on workplace research skills. *Educ. Lib.* 34, 19–31. doi: 10.26443/el.v34i2.308
- Ulich, E. (2011). *Arbeitspsychologie, 7th Edn*. Zürich: VDF Hochschulverlag.
- UNESCO (2013). *Global media and Information Literacy (MIL) Assessment Framework: Country Readiness and Competencies*. Paris: UNESCO.
- UNESCO (2015). *Global Network of Learning Cities*. Guiding Document. Hamburg: UNESCO Institute for Lifelong Learning.
- UNESCO (2019). *4th Global Report on Adult Learning and Education*. Hamburg: UNESCO Institute for Lifelong Learning.
- Venkatesh, V., Morris, M. G., Davis, G. B., and Davis, F. D. (2003). User acceptance of information technology: toward a unified view. *MIS Q.* 12, 425–478. doi: 10.2307/30036540
- Wahde, M., and Virgolin, M. (2023). DAISY: an implementation of five core principles for transparent and accountable conversational AI. *Int. J. Hum. Comput. Int.* 39, 1856–1873. doi: 10.1080/10447318.2022.2081762
- Watkins, B. J., and Tisdell, E. J. (2006). Negotiating the labyrinth from margin to center: adult degree program administrators as program planners within Higher Education Institutions. *Adult Educ. Q.* 56, 134–159. doi: 10.1177/0741713605283433
- Wersig, G. (1974). *Information-Kommunikation-Dokumentation*. Darmstadt: Wissenschaftliche Buchgesellschaft.
- Westphal, M., Vössing, M., Satzger, G., Yom-Tov, G. B., and Rafaei, A. (2023). Decision control and explanations in human-AI collaboration:

improving user perceptions and compliance. *Comput. Hum. Behav.* 144, 107714. doi: 10.1016/j.chb.2023.107714

Whorf, B. L. (1963). *Sprache, Denken, Wirklichkeit. Beiträge zur Metalinguistik und Sprachphilosophie*. Hamburg: Rowohlt.

Wilkins, U., Reyes, C. C., Treude, T., and Kluge, A. (2021). *Understandings and Perspectives of Human-Centered AI-a Transdisciplinary Literature Review. In Bericht zum 67. Arbeitswissenschaftlichen Kongress*. Dortmund: GfA-Press.

Wilkins, U., Süße, T., and Voigt, B.-F. (2014). "Umgang mit Paradoxien von Industrie 4.0—Die Bedeutung reflexiven Arbeitshandelns," in *Industrie 4.0—Wie intelligente Vernetzung und kognitive Systeme unsere Arbeit verändern*, eds. W. Kersten, H. Koller, and H. Lödding (Berlin: GITO), 199–210.

Wilson, J. R. (2014). Fundamentals of systems of ergonomics/human factors. *Appl. Ergon.* 45, 5–13. doi: 10.1016/j.apergo.2013.03.021

Wilson, T. D. (1999). Models in information behaviour research. *J. Document.* 55, 249–270. doi: 10.1108/EUM0000000007145

Wilson, T. D. (2000). Human information behavior. *Inf. Sci.* 3, 49–55. doi: 10.28945/576

Xu, W., Dainoff, M. J., Ge, L., and Gao, Z. (2023). Transitioning to human interaction with ai systems: new challenges and opportunities for HCI professionals to enable human-centered AI. *Int. J. Hum. Comput. Int.* 39, 494–518. doi: 10.1080/10447318.2022.2041900

Zipf, G. K. (1949). *Human Behavior and the Principle of Least Effort*. San Diego, CA: Hafner Publishing Company.