What is Human-Centered about Human-Centered AI? A Map of the Research Landscape

Tara Capel
Queensland University of Technology
t.capel@qut.edu.au

Margot Brereton
Queensland University of Technology
m.brereton@qut.edu.au

ABSTRACT

The application of Artificial Intelligence (AI) across a wide range of domains comes with both high expectations of its benefits and dire predictions of misuse. While AI systems have largely been driven by a technology-centered design approach, the potential societal consequences of AI have mobilized both HCI and AI researchers towards researching human-centered artificial intelligence (HCAI). However, there remains considerable ambiguity about what it means to frame, design and evaluate HCAI. This paper presents a critical review of the large corpus of peer-reviewed literature emerging on HCAI in order to characterize what the community is defining as HCAI. Our review contributes an overview and map of HCAI research based on work that explicitly mentions the terms 'human-centered artificial intelligence' or 'human-centered machine learning' or their variations, and suggests future challenges and research directions. The map reveals the breadth of research happening in HCAI, established clusters and the emerging areas of *Interaction with AI* and *Ethical AI*. The paper contributes a new definition of HCAI, and calls for greater collaboration between AI and HCI research, and new HCAI constructs.

CCS CONCEPTS

• **Human-centered computing** → Human computer interaction (HCI); HCI theory, concepts and models.

KEYWORDS

human-centered artificial intelligence, human-centered machine learning, artificial intelligence, machine learning, critical review

ACM Reference Format:
Tara Capel and Margot Brereton. 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), April 23–28, 2023, Hamburg, Germany*. ACM, New York, NY, USA, 22 pages. https://doi.org/10.1145/3544548.3580959

1 INTRODUCTION

As AI systems become more widespread, concerns grow about their embedded values, the goodness of their decisions, whom they benefit and whom they disadvantage [290]. Widely deployed in everyday applications such as spam filtering, text prediction, credit card fraud detection, credit scoring, search engines, news trends, market segmentation and advertising, insurance, loan qualification, and so on, they make decisions with social consequences, often on our personal and trace data that we generate every day [48]. Concerns arise from the inscrutability of the models used, embedded bias (particularly against minority groups) due to the historic data available and algorithmic choices, privacy issues, out-of-control machines, human rights challenges, and illusions of meaning that can be generated [48, 269]. There are also environmental costs due to the massive amounts of processing and the research opportunity cost of focusing on generating understanding from copious amounts of scraped data over alternative approaches [29, 269] – for example, the efforts invested into large language models could be utilized to revitalize smaller languages.

The term human-centered AI is increasing in use, motivating a sense that AI is to serve the people and in response to growing concerns about AI's potential to exploit and mislead. But 'human-centered AI' means many different things to different people. The human might be the subject of AI algorithmic study, the user of AI products, an agent in the design of the AI system itself and so on. Alternatively, HCAI might be invoked as an aspirational term, much like 'sustainable mining' in the resources sector or 'trusted autonomy', with considerable debate as to how and whether it can be achieved.

While there is clearly much at stake, there are already various and broad definitions of what it means to frame, design, and evaluate HCAI. As such, this review looks at the large corpus of literature emerging on HCAI to develop an understanding of what the community is defining as HCAI. We present and discuss the results of a critical review of peer-reviewed conference and journal articles that have been published up until July 2022. Our review included papers that explicitly used the term 'human-centered artificial intelligence' or 'human-centered machine learning' or their variations. Our aim is to provide an overview of the current state of HCAI, explore the claim to human-centeredness within these works, and investigate how human-centeredness affects the interaction between the human and the AI and its resulting impacts. Our intention is to create a map overview of the field to assist researchers to locate and differentiate their work, to identify gaps and opportunities in research and to assist beginning researchers to understand the nuances in the term HCAI. This paper is intended as an initial step in mapping the landscape of HCAI research. We advocate for further exploration in this space to include work that relates to the topic of HCAI but does not explicitly mention the term.

We begin this paper with a historical overview of the term HCAI as it has been introduced in seminal works in the field. We then describe the method of our critical review of papers. This led us to

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.
CHI '23, April 23–28, 2023, Hamburg, Germany
© 2023 Copyright held by the owner/author(s). Publication rights licensed to ACM.
ACM ISBN 978-1-4503-9421-5/23/04…$15.00
https://doi.org/10.1145/3544548.3580959

map the evolving field of HCAI in its current state as we interpreted it from the literature. Maps help to see the relationships between the various approaches, methods, and tools [249].

2 HISTORICAL OVERVIEW OF AI AND HCAI

The definitions of both AI and HCAI are varied and evolving in part due to the technical evolution of AI techniques and the corresponding evolution of HCI. Defined broadly, "AI comprises any technique that enables computers to mimic human behavior and reproduce or excel over human decision-making to solve complex tasks independently or with minimal human intervention" [140, 237].

2.1 Early Knowledge-Based AI, Debates about Context, and the Emergence of a Situated HCI Paradigm

Early AI techniques employed in the 1980s relied on a knowledge-based approach in which computers could automatically reason upon knowledge statements and logical inference rules coded into formal languages [39, 112]. However, the paradigm was limited by the difficulty of explicating knowledge in detail, especially tacit knowledge [46]. HCI scholars have argued this is because human reasoning is embodied, situated, in a social context and involves actions, often improvised, in the world, the complexity of which formal models cannot replicate [39, 84, 281]. Moreover, context itself is rarely delineable, stable and separable from activity so that it can be readily defined as a form of information to be encoded, the dominant positivist view. An alternative, phenomenological view prevalent in HCI research is that context arises from activities, is a relational property that holds between objects and activities, with the scope of relevant contextual features defined dynamically [85]. This latter view reflects a shift in HCI research towards a more situated paradigm [122], asking how machines can be designed as effective resources for human situated action, rather than trying to automate human reasoning. However, new techniques in machine learning (ML) have revolutionized AI.

2.2 Machine Learning, Increasing Computing Power and Data Availability, and Humans-in-the-Loop

Rather than codifying and programming knowledge into computers, machine learning "seeks to automatically learn meaningful relationships and patterns from examples and observations" [140], thereby automating the task of analytic model building through inferential statistics [38]. Machine learning has been fueled by new programming frameworks, access to the needed computing power, and data availability. By learning from previous examples and extracting regularities from massive (high-dimensional) datasets, machine learning can help to produce reliable and repeatable classifications to inform decisions [140].

In terms of human involvement, supervised learning requires human effort to label datasets to train algorithms to make the right predictions. In forms of human-in-the-loop and interactive machine learning [97], humans may be involved in classifying training, testing, tuning/correcting, and validating machine learning algorithms. By contrast, in unsupervised learning, algorithms are not given any labels and find structure in the input and memorize data in their own ways, by creating features in multiple layers of an artificial neural network [112].

2.3 Deep Learning and Human Interpretability

"Deep learning" [112] uses highly scalable algorithms to build models which are more complex and difficult for machine learning algorithms (like linear regression, logistic regression, decision trees etc.) to model. Using large, unstructured data, such as recorded voice, images, and text, it offers powerful applications such as voice recognition, autonomous vehicle operation and text generation. However, it is difficult to know why the underlying artificial neural network constitutes itself in the way that it does, what features of data it thinks are important, and the basis for its decisions, since there is no clear mapping to parameters which have a physical meaning or interpretation. Hence, the black box nature of deep learning raises concerns of accuracy and bias.

2.4 The "Contextualized Human"

By operating over vast quantities of data, collected from the real world, video data and from social contexts such as social media, machine learning and deep learning have access to much richer contexts than early forms of symbolic AI, and the computing power and statistical techniques to derive meaning from it. Due to digitally recorded aspects of context in massive and rich datasets, human reasoning, interaction and language can be mimicked and simulated. As Blackwell [39] notes, this "reduces the contextualized human to a machine-like source of interaction data. Rather than cognition that is not situated, our new concern should be interaction that is not humane". Blackwell [39] argues the need for "humane interaction", identifying concerns ranging from authorship, fair attribution, reward, self-determination and control. Its richness notwithstanding, context gleaned from sensor data, video, social media etc. is still highly selective, reduced and different from the social and embodied perceptions of people in the real world. Thus, Blackwell [39] further argues the need for improved conceptual constructs that can be used to account for a new designed relationship between user intentions and inferred models.

2.5 From AI to HCAI research

While AI research has sought to replicate and replace human performance, or simulate and emulate humans and their behavior (e.g., conversational agents, humanoid robots etc.), its increased sophistication has fueled an interest in HCAI, which takes aspects of the human user/partner/operator, their values and agency into account.

In response to this shift towards HCAI, international bodies such as the European Union have started to address the implications of AI within society, advocating for the creation of human-centered AI. Stanford University, UC Berkeley and MIT have established HCAI research institutes, advocating for AI that is humanistic and ethical, and that does not replace humans, but rather enhances them [266–269, 313]. For example, Stanford Institute for Human-Centered AI advocates for three overarching objectives that HCAI research and design should follow: "to technically reflect the depth characterized by human intelligence; to improve human capabilities rather than replace them; and to focus on AI's impact on humans" [313]. In line with these objectives, Xu [313] proposed an initial HCAI framework that includes three main components: "1) ethically aligned design, which creates AI solutions that avoid discrimination, maintain fairness and justice, and do not replace humans; 2) technology that fully reflects human intelligence, which further enhances AI technology to reflect the depth characterized by human intelligence; and 3) human factors design to ensure that AI solutions are explainable, comprehensible, useful, and usable". The overarching aim of Xu's [313] framework is to "promote a comprehensive approach, ultimately providing people with safe, efficient, healthy and satisfying HCAI solutions".

HCAI has also been argued by Shneiderman [269] to be focused on "amplifying, augmenting and enhancing human performance in ways that make systems reliable, safe and trustworthy". Shneiderman [269] proposes his own HCAI framework aimed at encouraging designers and researchers to question and consider the nature of both automation and autonomy: "1) design for high levels of human control and high levels of computer automation so as to increase human performance, 2) understand the situations in which full human control or full computer control are necessary, and 3) avoid the dangers of excessive human control or excessive computer control". Shneiderman's [266] definition of HCAI also emphasizes user experience design by placing the human at the center of design thinking, where "researchers and developers for HCAI systems focus on measuring human performance and satisfaction, valuing customer and consumer needs and ensuring meaningful human control". However, by examining the locus and distribution of control, and how AI could be used in increasing human performance, one could argue that Shneiderman's perspective on HCAI stems from a more engineering systems approach. A design-centered approach might be more tentative about the role of AI, further consider what constitutes good performance, and examine multiple ways of creatively using AI to enhance effectiveness, exploration and creativity. As an example, in the context of radiology, it might seem obvious that good performance constitutes a high percentage likelihood of a correct identification of a pathology, but there are important considerations around the implications of an incorrect diagnosis, false positives and negatives, the long-term deskilling of radiologists, healthcare costs, etc. Overall, the best performance may involve many diverging factors and how to achieve that will be arguable.

Xu's [313] approach to HCAI is more focused on the role that HCI professionals should play in its design. Xu [313] advocates for those within HCI to take a leading role by providing explainable, comprehensible, useful and usable AI, and encourages them to proactively participate in AI research and design in order to integrate methods between these two fields, increase their influence, and gain AI knowledge. However, when Yang et al. [314] investigated how UX designers could effectively work with machine learning, they discovered that designers found it more effective to work in collaboration with data scientists, rather than become machine learning experts themselves [314]. Shneiderman [266–269] advocates for HCAI systems to be designed through user-centered participatory design methods which engage a diverse range of stakeholders. The design of HCAI systems has been approached through various methods, including participatory design, focus groups, interviews, usability studies and observations [11, 190, 279, 303]. There is also work reporting that the design of the interaction with the AI itself is HCAI, in that the interface accompanying the AI is human-centered as it is explainable and interpretable by humans [5, 7, 16, 31] or that a human is engaged when the AI needs assistance or supervision [1, 25, 34, 37].

This also raises questions about what it means to evaluate HCAI with some researchers evaluating it based on human performance and satisfaction [266–269], how often the human needs to intervene [5], or how well AI designed through a human-centered design approach performs in particular contexts [26].

There remains ambiguity around HCAI with various and broad definitions of what it means to frame, design and evaluate HCAI. As such, this review looks at the large corpus of literature emerging on HCAI in order to develop an understanding of what the community is defining as HCAI.

3 THE REVIEW APPROACH

There is a large corpus of literature emerging on human-centered AI within a wide range of disciplines. The focus of this critical review was to understand the claim to human-centeredness in these works and how this claim to human-centeredness relates to the interaction with AI.

3.1 Search and Selection Process and Results

This review looks at the fields of Human-Centered Artificial Intelligence (HCAI) and Human-Centered Machine Learning (HCML). As such, we have chosen to include papers that explicitly self-identify as engaging with and contributing to human-centered AI and human-centered ML. We conducted a search in the Association for Computing Machinery (ACM) Digital Library, IEEE Explore, ScienceDirect, Springer Link and Scopus. The search included any result that was published any time up until July 1 2022, that either used the term "Human-Centered AI", "Human-Centered Artificial Intelligence", "Human-Centred AI", "Human-Centred Artificial Intelligence", "Human-Centered ML", "Human-Centered Machine Learning", "Human-Centred ML", or "Human-Centred Machine Learning", using the "OR" logic operator.

This resulted in a total of 2,357 initial items. We then read all abstracts to establish if the paper fit within the focus of our critical review and applied two criteria for keeping a paper in this initial review round: 1) that our key terms appeared in the title, keywords or main text of the article, and 2) that the papers were peer reviewed full papers. Reasons for rejecting a paper included:

- Where the key terms searched for were found outside of the title, abstract or main text of the article. For example, if the authors were associated with a Human-Centered Artificial Intelligence institute or lab, but the paper itself did not include any of the above keywords.
- Papers that were not peer-reviewed full papers. For example, extended abstracts, workshop proposals, pre-prints, posters, book chapters, Ph.D. dissertations, and theses.
- Papers not written in English.

This brought the publication total to 431. These papers were read with a focus on how they related to our critical review in terms of their claim to human-centeredness. For example, some papers broadly advocated for HCAI as part of their planned future work but did not address human-centeredness in the research described, hence they were rejected during this second review round. Other reasons papers were rejected at this point included papers where HCAI was raised as a challenge and where the main contribution was a review. These review papers differed from the review contributed within this paper in that they tended to analyze a specific area of HCAI or HCML. For example, enhancing trust in machine learning models, or HCAI in autonomous ship systems or healthcare. In the end, the remaining 257 papers were read in full and built the basis for the map and analysis we present in this paper.

The aim of this project was to understand what is meant when people use the term Human-Centered AI or Human-Centered ML. As such, we engaged with work that explicitly used the terms 'human-centered artificial intelligence' or 'human-centered machine learning' or their variations. We acknowledge that there are many papers which relate to the topic of HCAI and HCML that do not explicitly mention either term and that these will therefore have been missed from our analysis. This work is intended as an initial step in mapping the landscape of HCAI and HCML research and we advocate for further exploration in this space.

3.2 Analysis

Our analysis aims to understand the claim to human-centeredness in the research articles and how this claim to human-centeredness relates to the interaction with AI. Both authors conducted the initial five steps of thematic analysis [42], this included familiarization, initial coding, theme search, theme review, naming and definition. This involved reading through each of the papers of the final corpus with an aim to understand the human-centered nature or claim to human-centeredness of each. We then coded and thematized each paper based on that understanding of what their human-centered approach was. Some papers did not need to be read in detail as the approach to human-centeredness was apparent, whereas others required a more extended treatment. We met regularly to discuss papers and requested a second read of papers if there was uncertainty. We utilized a Miro board in order to write notes, code and theme each paper, which meant the naming process was done visually and iteratively, with both authors discussing where on the map the papers belonged.

A third, independent reviewer, was also engaged in complete coding capturing anything of relevance to the research question, "what is human-centered about human-centered AI?". This activity resulted in the explicit representation of the codes directly mapping research papers' content to provisional sub-themes, for example human rights mapped to papers that explore the issues of trust, privacy, racial discrimination etc. In the last phase only the two authors were involved in the conceptualization of the visual thematic HCAI map to derive the overarching themes of 'Ethical AI', 'Human Teaming with AI', 'Explainable and Interpretable AI', and 'Human-Centered Approach to Designing and Evaluating AI'. It is important to note that some papers cover more than one quadrant which was captured through the Miro board with papers duplicated across themes tagged as such. Of the 257 final papers, 65 were double counted and 5 were triple counted.

While the workshops proposals were not included within the final corpus or included on the HCAI map, they were used to help guide our analysis in terms of validating the overarching themes of interest to human-centered AI and ML research that stemmed from our thematic analysis. The initial CHI workshop on 'Human-Centered Machine Learning' intended to gather the community around the topic in 2016, identifying key research questions in the application of human-centered approaches of machine learning [110]. These were concerned with the role of the human in machine learning systems, usability challenges, how to design and evaluate machine learning systems with a human-centered approach, how human-centered machine learning is applied in various domains, how it could support creative work, and how domains such as big data analytics could be democratized. This was then followed up with a workshop in 2019 with the aim to further explore emerging areas and recent advancements in human-centered machine learning in order to articulate an updated agenda for human-centered research in machine learning [225].

Since the 2016 CHI workshop, there have been further related workshops that addressed the broad research area of HCAI and HCML [198, 199], as well as more focused areas such as explainable and interpretable AI [8, 86, 90, 94, 108, 115], humans teaming with AI [21, 77, 92, 124, 141, 227, 320], human-centered approaches for designing and evaluating AI [12, 204, 285], and ethical issues pertaining to AI and ML [13, 20, 22, 27, 50, 107, 113, 118, 176, 226, 243]. There are also workshops that have brought together two or more areas including 'Co-Designing AI Futures', which addressed both design and ethics of AI and ML [172], and 'Generative AI and HCI' [200] and 'Human-AI Co-Creation with Generative Models' [307] which addressed both humans teaming with AI and human-centered approaches for designing and evaluating AI.

We note here that Braun and Clarke's [42] reflexive thematic analysis approach rejects the idea of themes 'emerging' from data. Rather, it is the researchers' active process that generates and conceptualizes them. We acknowledge that the thematic analysis and mapping is our interpretation of the claims to human-centered AI laid out in the literature.

While thematic analysis seeks to generate themes and subthemes that describe the data, it does not have a tradition of attempting to lay the themes out on a map, by determining axes in order to visually separate out and relate the themes, although visual depictions of relations are sometimes presented. Since our research question is to describe a field, it makes sense to try to depict that field visually. Thus, we continued our iterative and inductive process toward considering what axes would help to spatially visualize the field of research, using the method of Sanders [249]. After describing the themes that constitute our findings below, at the beginning of the discussion, we elaborate on how we chose axes to visualize the field.

4 WHAT IS HUMAN-CENTERED AI?

Through our analysis and mapping, we found four major areas of research under the umbrella of HCAI. These are Explainable and Interpretable AI, Human-Centered Approaches to Design and Evaluate AI, Humans Teaming with AI, and Ethical AI. These are depicted on the map in Figure 1. A fifth newer area emerging in the center of the map is Interaction with AI, which we return to in the discussion. The four major research areas are described below.

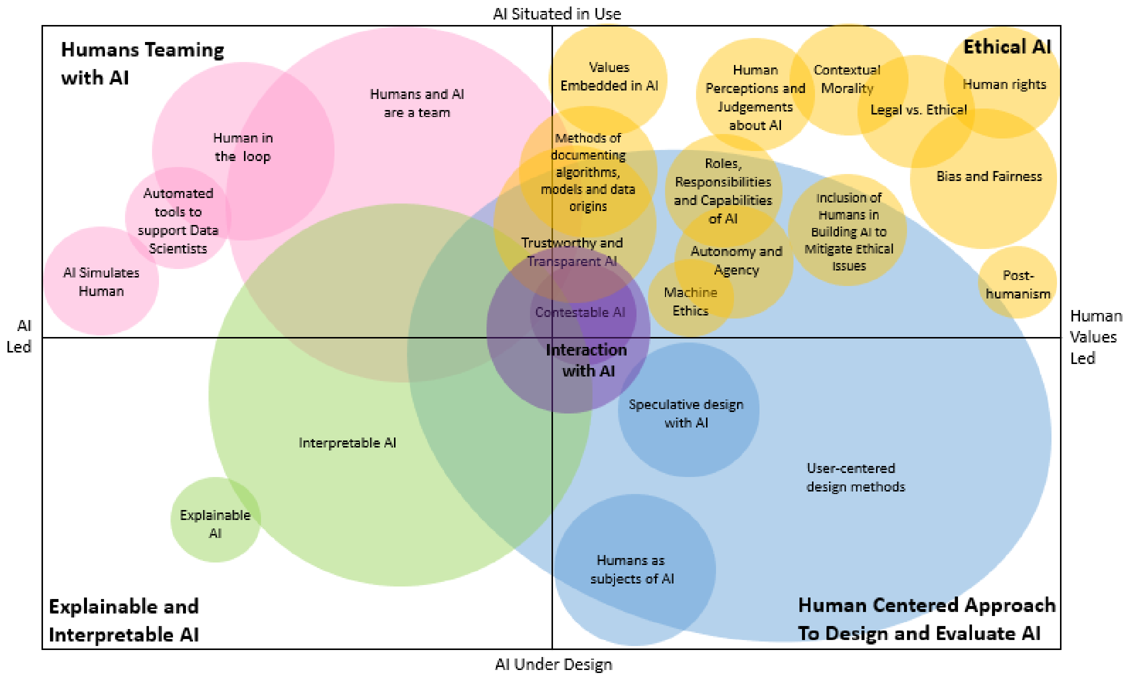

Figure 1: Map of the field of HCAI. The bubbles reflect the relative number of papers found in each area. The color coding indicates belonging to the major research areas.

4.1 Explainable and Interpretable AI

Explainable and interpretable AI includes a range of tools, methods and frameworks which aid a human in understanding the decisions or predictions made by the AI. This research area is in response to the 'black box' nature of AI models which make it unclear how or why an AI arrived at its decision or prediction, sometimes even to the person who built the AI. Explainability is used to both understand a model's behavior and improve a model's performance.

While explainability and interpretability are widely recognized as important, there is no one definition of what it means for AI to be explainable or interpretable. There are diverse ways in which a model could be interpreted, and explanations need to be tailored depending on the context and audience, for example data scientists [132] or domain experts [326]. Our mapping of the literature in this area revealed an increasing level of human agency within HCAI from systems being able to form an explanation [23, 103, 213, 231], towards ensuring that people are able to interpret and comprehend those explanations [5, 7, 16, 32, 44, 65, 75, 78–81, 89, 95, 120, 127, 132, 133, 136, 142, 145, 146, 150, 151, 158, 159, 170, 174, 186, 187, 196, 197, 207, 209, 221, 224, 235, 236, 246, 257, 262, 264, 280, 282, 283, 286, 288, 294, 318, 325, 326], to people being able to interact with and contest those explanations through asking "What if?" questions in the context of recommending someone for a home loan [274] and directly manipulating the underlying constraints considered by the system [297]. This emerging field is known as contestable AI.

Within our corpus, around 20% of the total papers fall into the category of interpretable AI. This is representative of the move towards HCAI, where the human-centered interaction with the AI stems from the human's ability to interpret a system's decision or prediction, and not just the system's ability to explain its decision or prediction. The importance of interpretability was particularly prevalent in contexts such as child welfare screening [326], education [5, 7], medicine and healthcare [95, 196, 246], finance [80, 186], fraud detection in digital retailing and banking [65], manufacturing [158, 236], media [170, 264], autonomous vehicles [280, 294], air traffic management [159], facilities management [75], and understanding game play patterns [142].

Interpretable AI was researched, designed and/or evaluated in relation to data-driven technologies that were either making a recommendation, decision, prediction, or classification where an interpretable explanation is required. For example, identifying fake media (deepfakes) [170], recommending news articles for people to read [264], diagnosing pathologies such as coronary artery disease [246], understanding patterns of behavior related to student success and risk [7], and financial investment recommendations [80].

Research within HCAI also argues that AI that is both explainable and interpretable makes the AI more transparent and can impact people's perceptions and judgements of AI, particularly with respect to trust [264].

Ongoing challenges identified are that explanations are very particular to the people who need them in their context of use, the level of expertise the person has around machine learning models, and their domain expertise. This has implications for both explanation design, 'tailorability', interaction design and for workforce training and quality assurance (e.g., in domains such as radiation therapy). The extent to which explanations and recommendations depend on the dataset is typically not discussed in detail and might benefit

Table 1: Human-centered approaches to design and evaluation of AI found in the literature survey

Specific Methods	Broader Design Approaches	Field Study Methods	Evaluation with End Users
Design probes: [279, 303] Design workshops: [11, 216, 277, 282] Gameplay: [154] Prototyping: [190, 259, 278] Wizard of Oz: [297] Design workbook: [36] Design toolkits: [102] Data enabled design: [206] Personas: [135] Vignette experiment: [116] Eye tracking: [272]	Participatory design: [165] Co-design: [15, 19, 31, 117] Research through design: [195, 256] Iterative design process: [7, 87, 119, 132, 196, 203, 273, 318, 326] Speculative design: [89]	Field study: [114, 145, 148, 321] Narrative study: [292] Mixed methods: [2, 25, 32, 49, 61, 95, 262, 296] Interview study: [75, 78, 211, 312]	[26, 28, 44, 53, 62, 72, 103, 104, 111, 156, 162, 164, 175, 180, 187, 194, 231, 248, 253, 254, 261, 287, 299, 300, 306, 308, 322–324]

from further investigation into how relations between explanations and datasets can be shown.

4.2 Human-Centered Approach to Designing and Evaluating AI

A human-centered approach to designing and evaluating AI involves the use of human-computer interaction methods and tools in the design and evaluation of AI systems. The overarching claim is that by engaging with a human-centered process in the design and evaluation, the AI that is designed is inherently human-centered.

Within our corpus, 49% of the papers approached the design and evaluation of AI through a wide variety of human-centered design methods, or advocated for such [18, 66, 134, 244, 268]. These approaches encompassed specific methods, broader design approaches, field study methods and evaluation with end users as shown in Table 1.

Human-centered methods were applied in a wide variety of AI contexts which demonstrates not only the vast application of AI, but the desire for greater human agency in these areas. The health and medical domains featured strongly. In many cases AI is used to assist with individual treatment or diagnosis by mobilizing historical data of others. Medical and health research included medical imaging and diagnosis [26, 51, 95, 117, 134, 291], clinical assessment of Multiple Sclerosis [196], care within intensive care units [148], post-operative behavior [206], assessing cognitive health [321], identifying depression through social media [239], emotion recognition [63], pregnancy [209], type 1 diabetes management [19], nutrition [103], fitness instruction [104], physical rehabilitation [164], assisted living [91], visual impairment [32, 195], deaf and hard of hearing [114], health communication [253], work performance management [216], improving workability of aged workers [24], and contact tracing [119].

HCI methods were used in contexts to support state organisations to make decisions about individuals e.g., child welfare [61, 62, 277, 326], and policing and recidivism [116, 262]. In education much of the work involved assessing student performance and risk [7, 11, 175, 194, 303, 306, 324]. HCI methods were also applied in complex real-time decision-making contexts, which hitherto have been largely AI led. e.g., autonomous vehicles [28, 208], autonomous systems [125, 268], robotics [156], public transportation [296], and on-demand food donation transportation [165]. There was a significant use of HCI research to apply language models e.g., text classification [231], multilingual interfaces [25], coding qualitative research [60, 99], creation of presentation slides for data scientists [322], and language translation [244].

HCI is beginning to be seen in critical and large areas such as fraud detection [65], cybersecurity [318], finance and investment decisions [80, 254], fact checking in news or misinformation [205], and news dissemination during crisis [180]. HCI methods were also seen in arts, entertainment, and sport e.g., sound design [259], music and art [260], fashion [144], gaming [4, 142], football [287], interactive AI for public spaces [173], storytelling [36], urban exploration [297], as well as marine ecology [217], facilities management [74] and career day organisation [17].

The outcomes of some of this research in this area were not specifically an AI or ML system, but guidelines or principles for designing them. These included designing for human-AI interaction [10, 80, 238, 263, 309], for co-creative AI [173, 260], for decision support [65, 69], for the design process [67, 251], for AI-powered services [305], for analytic tasks [144] and for trustworthy AI [252].

Research also investigated how humans contribute to data labelling [202, 208, 217], and how AI and ML could be incorporated into the research and design process [60, 91, 100, 109, 155]. For example, using machine learning to support qualitative coding in social science [60], and creating visual tools for qualitative data analysis that uses a human-in-the-loop to enable the tool to learn [155].

Within this human-centered approach to design and evaluate AI, humans were sometimes the subject of the AI systems being designed (e.g., in dance or radiology), but the claim to human centeredness stemmed from users (e.g., dancers, radiologists, pregnant women) being involved in the conceptualization of the AI system [4, 17, 51, 63, 98, 142, 167, 209, 239, 291].

Notably, most papers researching AI use in healthcare involved clinical specialists as users [51, 291], but not patients themselves, with the notable exception of a study looking at the preferences and expectations expressed by pregnant women in regards to AI [209]. The study found pregnant women would welcome an intelligent solution that is able to prevent a high-risk pregnancy by providing emotional support, as long as it is responsible, trustworthy, useful, user-centered, safe and personalized.

More recently human-centered design approaches are also embracing using AI as a design material in more speculative design research methods. Benjamin et al. [30] "propose *thingly uncertainty* to capture the capacity of ML-driven artefacts to be uncertain about the world, and thereby generating and adapting to a wide continuum of relations to other things, their datafied environment and people". How to design with AI and all its foibles, and how to make AI accessible as a design material to more people who do not know the technical details of AI is of increasing interest [314].

4.2.1 Humans as Subjects of AI.. Ten papers were themed "humans as subjects of AI" [4, 17, 51, 63, 98, 142, 167, 209, 239, 291] because they modelled humans, often their individual behavior or biology, in order to provide feedback and assistance. Human-centered design processes tended to be used to establish the foundations of the research, although not in all cases. For example, Saha et al. [239] applied new ways of thinking about person-centered approaches in human-centric, context-aware, and social sensing applications requiring personalized attributes. Calisto et al. [51] established foundations of research via a human-centered design process and following guidelines for human-AI interaction in the design of a human-centric AI assistant to aid radiologists in breast image classification.

4.2.2 Challenges for Human-Centered Approaches in AI.. There are ongoing challenges in terms of empowering people in designing with AI and ML that extend beyond being a participant or being a subject of AI research. One key challenge is conveying what AI can do and how it works to participants in such a way that they can usefully critique and envision applications. A series of papers explored how best to support designers [33, 86, 125, 278, 298, 310, 311, 314, 324] and non-experts (those not trained in AI) [205, 247, 316] in working and designing with AI. While there has been important work done in this area, there are ongoing challenges in terms of AI and ML education, and democratization [247, 316] to empower people in the design and use of AI systems. While there is awareness through social media, page rank algorithms, recommender systems and auto text completion about rudimentary aspects of AI, by learning from many people's actions, ML systems present us with a higher order of "sociomaterial complexity" [41]. Understanding the potential for how this will unfold, edge cases and outliers, how data might change the system over time, and how to engage people in long term use, custodianship and evaluation of AI needs to be researched.

4.3 Humans Teaming with AI

Humans teaming with AI posits that by working together, both the AI and the human can perform better and enhance their capabilities more than either could achieve by working alone. Human-machine team systems are usually evaluated through a lens of performance (the humans or the AI or both) and/or the satisfaction of the human with the performance of the AI system.

Our mapping of the literature revealed that the relationship between the AI and the human within this team is situated across a spectrum with humans becoming progressively more centered within the team, from the AI simulating a human, to a human being in the loop, to a human-AI collaborative team.

4.3.1 Humans and AI are a Team. Research that considered humans and AI as a team did so largely through a lens of collaboration. Within our corpus, around 20% of the total papers researched these human-AI collaborative teams in various capacities and domains [3, 14, 23, 24, 36, 49, 58, 61, 64, 66, 76, 82, 100, 116, 117, 128, 139, 143, 149, 163, 164, 179, 181, 184, 189, 201, 205, 217, 218, 220, 223, 254, 255, 260, 266–270, 275, 276, 287, 289, 293, 299, 302, 308, 317–319, 322]. Unlike in section 4.2, teaming was not prominent in medical fields, but was more prevalent in areas of healthcare and rehabilitation e.g., patient-centered healthcare [3], online health communities [302], mental health treatment [64], physical rehabilitation [164], and stroke rehabilitation [163]. Teaming did not emerge in autonomous vehicle work but was seen in areas of complex decision making such as piloted aircraft systems [76], power system control centers for energy transition [184], infrastructure assessment [149], plant maintenance [128], cybersecurity [318], and business process management [24]. Teaming research is seen in contexts to support state organisations to make decisions about individuals in areas that require significant caution, such as policing and recidivism [116], child welfare [61], and student performance within education [317].

Teaming emerged within creative endeavors where humans and AI engaged in varieties of co-creation, for example sound design [259], ideation [179], concert performance [319], music [293], and storytelling [36]. Teaming was seen in a variety of language areas where people could combine human meaning interpretation together with AI techniques e.g., fact checking [205], journalism [289], unstructured text analysis [270], code documentation [299], translating source code into different programming languages [308], and creation of presentation slides for data scientists [322]. Teaming was also seen in a variety of other areas such as digitization of geographic regions [189], marine ecology [217], football data analytics [287], robotics [82], financial advice [254], and design education and feedback [139].

Some of the examples of the work where human-AI collaboration and teaming was the claim to human-centeredness included pursuing a goal of hybrid intelligence [128, 218], designing conversational agents to work in collaboration with a human in areas such as finance and health [3, 254], exploring the use of AI-based screening tools to reduce bias in areas such as child welfare [61], and designing interfaces between advice-giving machines and advice-receiving human decision-makers in the context of "bailing and jailing" [116]. These are areas in which the complex individual context of application needs to be carefully considered against the weight of AI advice derived from algorithmic learning on datasets.

4.3.2 Humans in the Loop. Human-in-the-loop (HITL) is an area of AI that is aimed at leveraging both AI and the human in the creation and ongoing use of machine learning models [1, 4, 25, 34, 37, 93, 129, 133, 137, 155, 162, 203, 219, 273, 284]. It is closely related to human-machine teaming, but usually has a slightly narrower scope for human action and is more AI led.

Humans may be involved in model creation through labeling data, training the model, correcting erroneous classifications, evaluating, tuning etc. Their use of the model predictions can also form further input data for the continued evolution of the model, for example: I want to listen to a song I will like after this one. Humans both help train the model and continue to help evolve the AI through use and implicit or explicit feedback. Early interactive AI [97] involved humans training the model through interaction with model outputs, requiring the AI to do automatic feature selection with fast classifier training-times. HITL work raises important considerations about how AI and human input can be combined in terms of how AI systems work, where and when it is best for humans to contribute, and how AI computing might be reconfigured. For example, recent work explored decentralizing fake news detection through swarm learning in order to effectively protect user privacy and involve user feedback in the loop of fake news detection learning models [83].

Models created through a HITL process inherently require a human to interact with the model in some capacity, as such there is a natural claim to human-centeredness in these works. Note however that some regulatory approval processes, particularly in medicine, may prevent continuing evolution of a model through human feedback in order to ensure the model is known, understood and can be reliably approved. Contexts for HITL included pre-processing data [34], wastewater-based epidemiology [1], creating multilingual interfaces [25], agriculture and forestry [137], recognizing facial expressions [284], gaming [4], and ecology [219].

The ways in which the human interacted with the model varied across papers, for example in some work the human was in a supervisory role monitoring the decisions of the AI [1], in other work human expert knowledge was leveraged to bring in "experience and conceptual understanding to the AI pipeline" [137]. Another approach was to have AI embedded in conversational systems to continuously monitor and adapt to human users to make the systems language more 'human-like' [25].

HITL and human-machine teaming are closely related. Papers through their own terminology self-select as being HITL or human-machine teaming. Human-machine teaming papers tend to describe a broader role for the human in a wider context than do HITL papers.

4.3.3 *Automated Tools to Support Data Scientists.*

Within our corpus, two papers specifically focused on the relationship between data scientists and automated tools, also referred to as Auto-ML [300, 312]. These Auto-ML tools are aimed at leveraging the latest machine learning techniques to support data science projects with tasks such as pre-processing data, feature engineering and model selection. At their core, both works' claim to human-centeredness was based on the interaction between the Auto-ML systems and the data scientists who use them, arguing that Auto-ML tools should be aimed at supporting a partnership between the tool and the human, rather than focusing on full automation of the process.

4.3.4 *AI Simulates Human.*

AI designed to simulate human processes also appeared in our corpus of data, having been labeled by its authors as human-centered AI [59, 71, 175]. This work aims to mimic human processes or have the technology replace the human, but rarely considers human agency. Our initial research question of what is human-centered about human-centered AI arose in part because we often encountered work and grant applications of this sort that were described as 'human-centered' research. Examples of the work in this section included attempting to simulate the biological processes that occur in human vision as a means of detecting and classifying disease on a tomato leaf [59], envisioning a scenario where a system designed to evaluate student's learning performances would replace the need for a teacher within that process [175], and emulating a human decision-making process in the context of emergencies based on the abilities of the human mind [71]. This work sits at the far left of the map of HCAI in Figure 1 for this reason. Although it is not human-centered, it may benefit from a human-centered approach.

4.3.5 *Challenges for Human Machine Teaming.*

This mapping from the AI simulating a human, to a human being in the loop, to a human-AI collaborative team is indicative of the move towards human-centered AI, with research taking more consideration around how the human could be more centered and given more agency within the human-AI team. Ongoing challenges in this space include considerations around what is the appropriate level of human and machine contributions within these teams, where the strengths of AI and humans complement each other, developing the competencies and capabilities of both. Other challenges relate to fields that are not represented in HITL and HMT research, such as medicine, autonomous vehicles and real-time decision-making systems, that may invoke complex moral dilemmas in complex situations of application.

4.4 Ethical AI

Ethical AI [6, 9, 35, 40, 43, 45, 47, 52, 54–57, 62, 68, 70, 73, 80, 87, 94, 96, 101, 105, 106, 123, 126, 130, 133, 138, 152, 153, 157, 160, 161, 165, 166, 168, 169, 171, 177, 178, 180, 182, 183, 188, 191–193, 207, 210, 212, 214, 215, 222, 228–230, 232–234, 240–242, 245, 250, 258, 264–269, 292, 294, 295, 301, 304] is AI that seeks accountability regarding fundamental human values and rights, and advocates for more transparent design of AI. Its overarching claim to human-centeredness is that it considers the rights and values of the people who are working with the AI or impacted by the AI, particularly within sensitive contexts. Unlike the other three themes that stemmed from our analysis, work situated within ethical AI broached a wide range of topics with no saturation in any one area. In more established areas such as HITL and interpretable AI, humans tend to have a consistent role of being the one the interpretation is designed for, or the one who helps label, train, reclassify, use and tune the AI model. In the emerging space of ethical AI, which considers more broadly the scope of humane interaction [39], the human role is being considered from many new and different perspectives. Here we will give a brief overview of those topics as they relate to the ethical AI agenda within HCAI.

4.4.1 *Values Embedded in AI..*

Papers within this theme advocate that HCAI can achieve value alignment by embodying human values, rather than just supporting them, through stakeholder engagement in the design process [62, 157, 165, 228, 232, 250]. For example, Komatsu et al. [157] worked with a group of journalists to understand their professional values and how these might be supported and/or undermined by AI. They found that values such as truth, impartiality and originality were important to journalists, and that those values should be embodied in systems designed for them.

4.4.2 *Contextual Morality.*

The concept of contextual morality builds on the work exploring the intersection between AI and human values by recognizing that values differ between people and considers the effect of context on user expectations and behavior [229, 292]. Within the review, these considerations are applied to autonomous decision-making systems, such as autonomous vehicles [229], investigating ongoing social concerns around the ethical nature of the decisions autonomous systems may make within evolving contexts such as accidents, for example whether an action prioritizes the life of the driver, third party vehicle, passenger, pedestrian etc.

4.4.3 *Trustworthy and Transparent AI..*

As AI systems permeate a broader range of application domains and work in collaboration with humans, there are important considerations for both trustworthy and transparent AI. One essential element is disclosing how system data was gathered and how the algorithms and models were designed from that data [126]. AI systems can also behave in unexpected and inconsistent ways, which can impact the trustworthiness of the system to the user, therefore it is also important for AI systems to explain their outputs in a way that is both understandable and interpretable to the user [47, 80, 133, 294]. Transparent AI also encourages data scientists to reflect on their positionality in context with other stakeholders, as well as the models that also embody a perspective, so that their decisions can be less biased and more transparent [52]. The relationship between trustworthy and transparent AI remains inconsistent. While transparency is important for building trust in AI, it does not always make a system more trustworthy, particularly in instances where the system is incorrect, or the explanation does not align with expectations.

4.4.4 *Methods of Documenting Algorithms, Models, and Data Origins.*

Transparency can be achieved through data or model cards, documentation of the data and who is represented within it, as well as how the model was made, assumptions that were made, how the model behaves in relation to diverse groups of people, and how well the model performs in relation to those groups [45, 55, 68, 105, 191, 222]. The focus on effective documentation of data among Ethical AI researchers arises from the impact that data quality has on AI models, and the problem that in AI research, "Everyone wants to do the model work, not the data work" [240].

4.4.5 *Human Perceptions and Judgements about AI..*

Research that investigated human perceptions and judgements about AI had a particular focus on trust and acceptance of AI systems, both from the perspective of data scientists and non-experts. This work explored how AI can be designed, built, and implemented in a way that is explainable, interpretable, and transparent so it leads to a greater sense of trust and acceptance of AI [87, 152, 166, 180, 264, 304].

4.4.6 *Roles, Responsibilities, and Capabilities of AI..*

Research into human roles and responsibilities when using AI systems is helping to create methods and frameworks to determine the capabilities people need in an AI-infused world, and what or who is accountable for the behavior and outcome of an AI system [9, 73, 183]. This work supports trustworthy and transparent AI and is needed to ensure uptake of AI systems. Medical fields such as radiology and radiation therapy have started to introduce AI education to clinicians, and are concerned with workforce training, user interaction, explainability and critical reasoning, protocols for data sharing for clinical use, and quality assurance.

4.4.7 *Autonomy and Human Agency.*

Autonomy and human agency consider how AI systems can support or hinder human autonomy [161], for example consenting to your data being used in training systems. Work within this theme also addressed the impacts on loss of autonomy [212], including demotivating people, damaging mental health, undermining interpersonal relationships, and reducing self-efficacy.

4.4.8 *Machine Ethics.*

Machine ethics is concerned with the way in which machines are imbued with ethical considerations, how these ethical considerations adapt over time, and the consequences of those ethical considerations for how the machine behaves towards other machines and humans. It endeavors to create AI that is guided by ethical principles both from a top-down approach where the machine applies a set of pre-defined ethical considerations as well as a bottom-up approach where it can learn new considerations [57].

4.4.9 *Inclusion of Humans in Building AI to Mitigate Ethical Issues.*

Work in this area argues for the inclusion of humans in building AI as a potential means of mitigating ethical issues [168, 234]. Of importance are those who will be impacted by the deployment of an AI system, particularly in the development of requirements for AI to be ethical [234].

4.4.10 *Legal vs Ethical.*

Ethical AI is not limited to what is permissible by law and often goes beyond legal requirements. Work in this area considers the interplay between ethical and legal approaches to AI, particularly around the regulation of AI [182], and advocates for the progression of legal frameworks in the construction of AI to keep up with ethical implications [54].

4.4.11 *Bias and Fairness.*

Work within bias and fairness explores the implications of bias in AI systems and how AI can reinforce societal bias through data and algorithms, with a goal to create AI systems that are fairer and more accurate [40, 258]. This work also looks beyond bias towards investigating power differentials to explore the root causes of bias beyond data and algorithms [188].

4.4.12 *Human Rights.*

Human rights HCI research explores the impacts of AI on people's human rights, particularly people's right to privacy in contexts such as health and surveillance, and the impacts of biased data and algorithms within contexts where privacy is paramount [70, 210, 265].

4.4.13 *Posthumanism.*

Posthumanism addresses the human supremacist rhetoric in discussions around AI ethics and policy, which privileges human interest over others (such as endangered species, forests etc.) as a matter of public policy. Work in this area that engages with HCAI explores the blurry dichotomy between humans and machine agents [94].

4.4.14 *Challenges for Ethical AI..*

Our analysis revealed considerations of trust, transparency, autonomy, agency, privacy, morality and non-discrimination threaded throughout many of the above topics. Big challenges relate to fair resource deployment, ethical and transparent data use, and morality of time critical AI systems deployed in the wild. Inscrutability of models from massive data are a concern, being hard to check for embedded biases, risking cementing in historical, sexist, and racist perspectives. They may be used to create illusions of meaning and misinformation [29]. AI tends to be developed for the Global North, to meet corporate business goals, which may sideline some issues and populations and leave them vulnerable to misapplications of AI.

5 DISCUSSION

5.1 The Mapping

The literature review found significant breadth in what is considered by the community to be human-centered AI, from research led by a strong technical focus to human values focused research, and from design focused research to deeply contextual investigations of use.

We presented the research on a map, as maps are useful for showing the complexity and change in the research landscape. The underlying landscape of the map only changes as major forces affect it [249], such as the emergence of new research areas, large shifts in technology, or new societal movements. The research map shows the dominant areas of research in HCAI to date.

We debated and reflected as we iterated potential dimensions of the map, seeking to find and name axes of the HCAI research space in a way that would "bring clarity and light to the landscape" [249]. There were different candidates for axes on which to distribute the major research areas. One obvious one, chosen as the horizontal axis, reflected the extent to which the research focus is on understanding and foregrounding human values and concerns themselves (*human values led*) versus the research focus foregrounding the AI or algorithmic techniques themselves in service of AI products that are more accountable, intelligible, and scrutable (*AI led*). The middle ground reflects research that seeks to bring together in some deeper way both human values and AI development. The horizontal axis is thus a primary differentiator; however, we debated other potential axes. Other axis candidates included generalizability vs specificity of the AI model; 'agency' vs 'Agency,' the small 'a' reflecting individual agency, choice, and awareness in the moment of use, the large 'A' reflecting agency over time due to a broader and deeper consideration of factors impacting groups and societies. Another potentially revealing scale ran from individual use through to group, organizational and societal impact. Another potential differentiator was contextual richness – the extent to which context of use is foregrounded in all its specificity, which tends to be revealing of the kinds of reductions and choices made in AI design.

We found the clearest way to map out the space of the research was to add a second vertical dimension reflecting the spectrum of whether AI was *under design* or *situated in use*. This axis allows expression of whether AI is being studied in use in society to reveal its flaws, limitations, and challenges, or whether it is being designed. AI designed in the lab would be represented at the lower end of the axis whereas AI being 'designed in use,' informed by much fuller consideration of context of use and intertwined with consultations with users, inhabits the middle of the axis. Design-after-design [88], the notion of how users appropriate and alter finished designs to their own ends also inhabits this middle area of the axis. The top of the vertical axis reflects AI being studied in use, to inform design, policy, advocacy, or lawmaking or to critique design from a use standpoint, but where the focus is understanding use rather than design itself. Thus, the vertical axis represents the extent to which research is embedded in use or under design, and the center is an interesting place inhabited by design work that is enmeshed within rich contexts.

Naturally, not all research is described neatly by one label or fits on only one area of the map, however the map was the clearest way that we found to convey the landscape of research. We underscore that higher dimensional maps are possible. For example, our preferred third axis candidate would be generality versus specificity of the AI model. Specific models are easier to check, more likely to involve humans-in-the-loop of labelling, training, testing, verifying etc., and the scope of model use is narrower. General models, such as dominant language models, potentially have much greater applicability to mimic humans, create deepfakes, and to be applied far from their intended contexts of use or their data sources. While any combination of axes and further dimensions are possible, we chose two axes that seemed to best reveal human-centered concerns.

5.2 How the Major Research Areas Inhabit the Map

The four major research areas making up the quadrants of the map are (I) *Explainable and Interpretable AI*, (ii) *Human-Centered Design Methods*, (iii) *Human-AI Teaming*, and a large agglomeration of research areas that emerged as distinct in our review but together make up what we call (iv) *Ethical AI*. Finally, emerging in the middle is a research area we refer to as *Interaction with AI*, which encompasses designing contestable AI systems.

We begin with the emergent middle area of *Interaction with AI*. This research explicitly addresses the need to understand how people will interact with inferred models in embodied and situated contexts. Although there is some overlap with *Human-AI Teaming* approaches (including interactive AI and HITL), *Human-AI Teaming* approaches tend to define the context in which human input is sought to closely align with the original system intentions. In a departure, *Interaction with AI* research investigates how people are using AI systems creatively, how they can contest AI systems, and how AI system use in the wild can inform design, in ways hitherto unimagined by the AI creators. Given the growth of artificial general intelligence, with applications that, for example, generate text and images in response to text requests, with little understanding of the user's context of use, this area is likely to grow quickly.

Research into AI use in situated and embodied contexts can uncover how to restore agency to users [39] that may overtly seem to have been lost to AI models derived from digital "data contexts". Research can work out all the questions people ask and workarounds that they make to accommodate the AI model and determine protocols for how AI should be used in the world. Research can uncover user needs in order to inform AI design, and workplace training, to ensure informed use of systems with embedded AI.

Another goal for research in *Interaction with AI* is better interaction. Harper's [121] discussion of text autocorrect systems points out that while spelling words is well supported, choosing the right word in the context, in the larger "unit of action" of making sentences is not. What is required is not explanations, but good interaction design informed by HCI. He argues HCI can help to create "abstractions that cohere both the user and the application(s) in ways that lets them work hand in hand" [121]. This is similar to Blackwell's [39] call for improved conceptual constructs that can be used to account for a new designed relationship between user intentions and inferred models.

Vaccaro et al. [290] argue that designing for contestability in systems can assist in "surfacing values, aligning system design and use with context, and building legitimacy". For example, people should be able to submit their own scenario, see how the AI performs, argue their case, and decide the merits of the system. That is, interacting with AI should acknowledge and embrace human reasoning styles, which are also made on the basis of incomplete and selective data [39]. Interaction design of AI can find ways to highlight potential ethical questions, consequences, and accountability through provocations in the design.

Research in this middle ground of *Interaction with AI* demands multidisciplinary skilled teams that draw together technical, design and human-centered researchers, which is perhaps why it has been slower to emerge but has great potential and is in ascendance.

At the far left of the horizontal axis (*AI led -> Human values led*) we see AI led research that seeks to emulate or replace human capabilities, simulating human processes such as vision, or human-decision making. This is largely technical research that does not engage in considerations of human agency. It is called human-centered because humans are the subject of AI research, but many may feel this is a misnomer. Although inspired by humans, it does not necessarily consider their agency, but we argue it should in future. In broad brush strokes, moving from left bottom fanning out to the right and the top, we might consider research as being *on people, for people, and/or with people.*

Within the quadrant of research called *Explainable and Interpretable AI*, explainable AI is largely technical research that seeks to represent and explain AI decision making processes for people, with varying degrees of human inclusion in understanding how to best represent explanations for people. Interpretable AI research, which focuses on ensuring human interpretability of AI decisions and explanations, is shown further along the human values led horizontal axis. The corpus of interpretable AI work found in our survey was much larger than the corpus of explainable AI work, probably because interpretable AI tends to have a more human-centered focus, whereas much explainable AI work would not include human-centered keywords or make human-centered claims to identify itself for inclusion in this review. Thus, only a subset of the large field of explainable AI, (that which identifies itself as human-centered), is captured here and shown on the map.

Human-AI Teaming inhabits the upper left area of the map, representing approaches that seek to complement the computational power of AI with human skills and judgment. *Human-AI Teaming* is an evolution in some respects of human-in-the-loop methods (also shown) and early forms of interactive AI. The research tends to be context specific, hence it is placed in the upper part of the map. Human-in-the-loop approaches and early interactive AI tend to be framed as AI problems, with people assisting the AI through labeling data, correcting erroneous classifications, and evaluating and tuning the model. Human-machine or *Human-AI Teaming* methods seek greater synergies beyond putting the human-in-the-loop, with humans guiding machines/AI by problem framing, questioning, labeling, information seeking, dialogue and mitigating against machine failures such as data bias, model complexity, rare cases, and exceptions. Machines/AI in turn can facilitate human exploration and critical thinking about large-scale data and mitigate against human failings such as bias, limited attention fatigue and affect. Designing effective human-AI teaming demands both attention to context, human concerns, human skills, and interaction design, as well as AI techniques. High levels of human control and high levels of automation can complement each other. The extent to which the contextual and human aspects are included determines how far towards the top and middle of the diagram a human-AI/machine teaming project would sit.

It is worth clarifying that the purpose of human-AI teaming is to make the best use of both human and AI capabilities, rather than the human simply being called upon to do what the AI cannot yet manage in an AI led project i.e., step 4 in the 5 levels of vehicle autonomy [232]. For example, step 4 in the quest for vehicle autonomy involves a driver monitoring an autonomous vehicle to take over if it fails. However, humans are notably bad at monitoring monotonous processes in which they play no active part, and their bodies are not primed to attend and react in the way they would if they were themselves driving. As such, the solely AI-led framing of many autonomous vehicle projects fails to embrace human competencies and agency from the outset, and thus few were identified in our review. Human-AI teaming projects aim to make the best of both human and AI capabilities, enabling both humans and the AI system to learn and develop. They consider what should and should not be automated. The more a project balances these human and AI capabilities, the further from the left and more towards the middle they sit on the map. Marres [185] identifies as problematic that engineers often create their own space in the world to work on technical problems, constraining the role of humans and the "interactional frame", which then limits exploration of alternative futures. Such approaches fall off the far left of the map.

Human-Centered Design research is shown in the lower right area of the map, reflecting its human values led and design led focus. It researches how AI can be designed to align with human needs and aspirations. People may be involved as informants, participants, or partners in research. Informed by the fields of participatory design and co-design [271], projects include those who will potentially be impacted by design in the design process. This is done from both a moral and pragmatic standpoint. From a moral standpoint, those who will be impacted should be consulted. From a pragmatic standpoint, they are experts in their own lives, able to offer detailed contextual insights and directions as to what is needed. Many participatory design methods such as excursions (visiting functioning system installations), scenarios, future workshops and games have been applied to understand what AI offers as well as to imagine AI as part of their future activities [41].

To date, speculative design and interaction design with AI are less well represented in human-centered design research, in part due to the struggle that interaction designers face in working with AI as

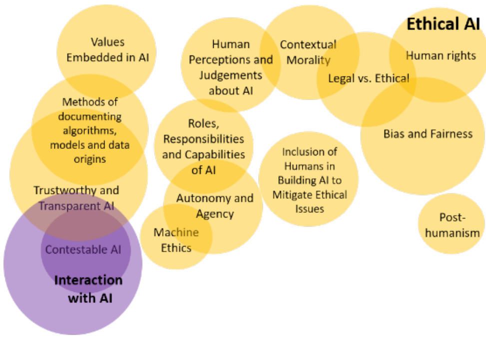

Figure 2: Ethical AI research has sprouted examining many aspects of power, bias, agency, transparency, governance, and morality.

a design material [314, 315]. However, exciting work is emerging in speculative design [30] and in contestable AI [131]. Interaction design is investigating how people interact with AI systems, asking how the AI is represented in the user interface [224], and beginning to ask how people can compose AI for themselves.

Ethical AI as shown in Figure 2 represents an agglomeration of research to understand values embedded in AI projects, issues of ethics, fairness, and power. Succinctly provocative, Kalluri [147] states "Don't ask if artificial intelligence is good or fair, ask how it shifts power – Those who could be exploited by AI should be shaping its projects". *Ethical AI* projects seek to understand and reveal the contextual details of how AI is made and how it impacts people. It advocates for and demonstrates transparent processes in the construction of AI models – documenting how data is gathered, what it comprises, who is represented within it, assumptions and mechanisms behind models, how they represent different user groups and perform in relation to those groups etc., thus revealing the implicit values and decisions made. It also advocates for consideration of how resources are deployed. e.g., the extent to which resources are deployed to generate language models for dominant languages versus endangered languages, where the low number of speakers and materials on which algorithms can be trained demands novel HCAI approaches. With its focus on the human experience of AI, situated in the real world, *Ethical AI* is both human values led and deeply contextual, thus it inhabits the top right area of the map. Much of this research points to the need for more innovative and considered design work. Posthuman and more than human approaches to AI are also seen in this space.

5.3 Reflections on the Landscape

There is significant breadth in what is considered by the community to be human-centered AI, from AI led to human values led research, and from design focused to contextually focused research. Despite all the human-centered design and ethical AI research represented here, it is extremely common at the AI led end of research (beyond HCAI research mapped here) to ignore human agency and user involvement altogether, or to only consider human user involvement to evaluate a final system already conceived by technologists. The notion of early investigation into contexts and use with potential users and those impacted to frame studies is alien to many who produce AI, even though AI is being embedded in domains as diverse as agriculture, manufacturing, transportation, government, finance, security, law enforcement and healthcare. Corporate technical AI researchers and engineers hold a great deal of power when designing systems that are becoming more central to our lives, embedding their own values into the design of algorithms. Ethical AI projects seek to examine and reveal these values.

A critique might similarly be levelled at HCI that while AI is altering the basis of computing, HCI has not paid enough attention to learning how new AI systems work and their consequences and possibilities for new interaction design [121]. It is important to look at both the system's and the user's point of view [41, 121]. The way in which users and their use context and AI are both intertwined and decoupled to best effect is the core topic of HCAI.

Within the field of human-centered AI research, the greatest challenge and opportunity is to engage greater collaboration across the spectrum of research, engaging technical, design, human-centered and critical ethical research, to amplify and accelerate endeavors to create exemplar processes and systems. This is no small challenge as it demands conversations from many different perspectives and viewpoints within HCAI. Researchers need to be willing to learn and understand the language and methods of other areas, as well as the values and motivations that underlie research in that area.

AI led research is dominated by the language and values of inferential statistics and a mindset to eke out value from big data and algorithms to apply to unique and often personal problems and situations at different places and times, sometimes in high stakes and dynamic situations. The detailed context of application is less of a focus. In contrast, human-centered research pays attention to personal and social experience and context and how people use, appropriate and experience technologies in context. It is undeniable that AI and HCI need each other and that HCAI research can benefit from stronger collaborations across fields and efforts to understand each other's work and values. Such collaborations can yield new conceptual constructs that account for human relations with AI, guide coherent interface design [121], and reflect new designed relationships between user intentions and inferred models [39].

Beyond this, HCAI research can seek to reach out to influence the broader AI research and business communities and domain stakeholders. Domain experts can highlight considerations around values and potential consequences that may not be obvious to AI designers, giving them greater influence. Given the multiplicity of interests, varieties of users, business interests and domain interests at play in any situation, this is no small challenge, but worth noting in closing commentary on this landscape.

5.4 A Definition of HCAI that Reflects the Research Landscape and Agenda

Returning to the early definitions of HCAI in section 2.5, we offer a revised definition that reflects the research landscape; it emphasizes the importance of ethics and the ability to interact with AI to understand, guide and contest it. Early definitions emphasize the betterment of humans but leave open who decides what is good for humans and who benefits. Incorporating ethics into the definition of HCAI focusses attention on the requirement for ethical practice to consider all who may be impacted (including other species), as any ethics panel would insist. By incorporating interaction, consideration of actual use is foregrounded. We therefore define Human-Centered Artificial Intelligence as follows:

> Human-Centered Artificial Intelligence utilizes data to empower and enable its human users, while revealing its underlying values, biases, limitations, and the ethics of its data gathering and algorithms to foster ethical, interactive, and contestable use.

How a system reveals itself is a matter of judgement and interaction design. What makes sense depends on the context of use. However, the emphasis in HCAI work on interpretability, human-AI teaming, ethics, human-centered design and interaction makes clear that HCAI systems should aim to partner with people in ways that foster their agency and awareness.

6 CONCLUSION

The aim of this project has been to understand what is meant when people use the term Human-Centered AI. Our paper aims to be a research primer and summary for beginning researchers, students, and practitioners, as well as a tool for reflection for those established in the field. Our map of the domain reveals the diversity of research on axes from *AI led* to *human values-led* and from *AI under design* to *AI situated in use*. The major research areas identified are Explainable and Interpretable AI, Human-Machine Teaming, Human-Centered Design, and Ethical AI. Emerging fields are Designing Interaction with AI, more-than-human considerations, and research across the board that has a greater infusion of ethical AI, transparency, and greater consideration of harmful impacts, bias, and discrimination. Based on the research landscape and agenda, we offer a new definition of HCAI that emphasizes the importance of ethics and the ability to interact with AI to understand, guide and contest it.

HCAI needs greater collaboration between AI and HCI researchers to develop new conceptual constructs that account for human relations with AI, guide more coherent interface design, and reflect relationships between user intentions and inferred models. We hope the map might inform researchers about the breadth of research happening in HCAI, gaps in the way research projects are formulated, areas of HCAI that might be embraced to strengthen a team and project, and that it may lead to research into new HCAI constructs and methods.

ACKNOWLEDGMENTS

We thank our reviewers for their constructive feedback on our paper and our national research council and university for supporting this research. This project is funded by an Australian Research Council Discovery Project grant. Project ID ARC DP200103582 Human-machine teaming: designing synergistic learning of humans and machines. It is also supported with funds from QUT (Queensland University of Technology) Centre for Data Science.

REFERENCES

[1] Omar M Abdeldayem, Areeg M Dabbish, Mahmoud M Habashy, Mohamed K Mostafa, Mohamed Elhefnawy, Lobna Amin, Eslam G Al-Sakkari, Ahmed Ragab, and Eldon R Rene. 2022. Viral outbreaks detection and surveillance using wastewater-based epidemiology, viral air sampling, and machine learning techniques: A comprehensive review and outlook. *Science of The Total Environment* 803 (Jan. 2022), 149834-149834. DOI: https://doi.org/10.1016/j.scitotenv.2021.149834

[2] Patricia Abril-Jimenez, Maria Fernanda Cabrera-Umpierrez, Sergio Gonzalez, Rosa Carreton, Ginger Claassen, and Maria Teresa Arredondo Waldmeyer. 2022. Design of Human-Centered Adaptive Support Tools to Improve Workability in Older Workers. A Field of Research of Human-Centered AI. In *International Conference on Human-Computer Interaction (HCII '22)*. Springer, Cham, 177-187. https://doi.org/10.1007/978-3-031-05028-2_11

[3] Achini Adikari, Daswin de Silva, Harsha Moraliyage, Damminda Alahakoon, Jiahui Wong, Mathew Gancarz, Suja Chackochan, Bomi Park, Rachel Heo, and Yvonne Leung. 2022. Empathic conversational agents for real-time monitoring and co-facilitation of patient-centered healthcare. *Future Generation Computer Systems* 126 (Jan 2022), 318-329. https://doi-org/10.1016/j.future.2021.08.015~

[4] Sabbir Ahmad, Andy Bryant, Erica Kleinman, Zhaoqing Teng, Truong-huy D. Nguyen, and Magy Seif El-Nasr. 2019. Modeling individual and team behavior through spatio-temporal analysis. In *Proceedings of the Annual Symposium on Computer-Human Interaction in Play (CHI PLAY '19)*. Springer, Barcelona, Spain, 601-612. https://doi.org/10.1145/3311350.3347188

[5] Junaid Akhtar. 2020. An interactive multi-agent reasoning model for sentiment analysis: a case for computational semiotics. *Artificial Intelligence Review* 53, 6 (Nov. 2019), 3987-4004. DOI: https://doi.org/10.1007/s10462-019-09785-6

[6] Ramya Akula and Ivan Garibay. 2021. Ethical AI for Social Good. In *International Conference on Human-Computer Interaction (HCII 2021)*, Springer, Cham, 369-380. https://doi.org/10.1007/978-3-030-90963-5_28~

[7] Ahmed Al-Doulat, Nasheen Nur, Alireza Karduni, Aileen Benedict, Erfan Al-Hossami, Mary Lou Maher, Wenwen Dou, Mohsen Dorodchi, and Xi Niu. 2020. Making sense of student success and risk through unsupervised machine learning and interactive storytelling. In *International Conference on Artificial Intelligence in Education (AIED 2020)*. Springer, Ifrane, Morocco, 3-15. https://doi.org/10.1007/978-3-030-52237-7_1~

[8] Cecilia Ovesdotter Alm, Alberto Alvarez, José Font, Antonios Liapis, Thomas Pederson, and Johan Salo. 2020. Invisible AI-driven HCI Systems – When, Why and How. In *Proceedings of the 11th Nordic Conference on Human-Computer Interaction: Shaping Experiences, Shaping Society (NordiCHI '20)*. Association for Computing Machinery, New York, NY, USA, Article 130, 1–3. https://doi.org/10.1145/3419249.3420099

[9] Steven Alter. 2022. Agent Responsibility Framework for Digital Agents: Roles and Responsibilities Related to Facets of Work. In *International Conference on Business Process Modeling, Development and Support, International Conference on Evaluation and Modeling Methods for Systems Analysis and Development (BPMDS 2022, EMMSAD 2022)*, Springer, Cham 237-252. https://doi.org/10.1007/978-3-031-07475-2_16

[10] Saleema Amershi, Dan Weld, Mihaela Vorvoreanu, Adam Fourney, Besmira Nushi, Penny Collisson, Jina Suh, Shamsi Iqbal, Paul N. Bennet, Kori Inkpen, Jaime Teevan, Ruth Kikin-Gil, and Eric Horvitz. 2019. Guidelines for human-AI interaction. In *Proceedings of the 2019 CHI conference on human factors in computing systems (CHI '19)*. Springer, Glasgow, Scotland, UK, 1-13. https://doi.org/10.1145/3290605.3300233

[11] Renate Andersen, Anders I. Mørch, and Kristina Torine Litherland. 2022. Collaborative learning with block-based programming: investigating human-centered artificial intelligence in education. *Behaviour & Information Technology* 41, 9 (June 2022), 1-18. DOI: https://doi-org/10.1080/0144929X.2022.2083981

[12] Tariq Osman Andersen, Francisco Nunes, Lauren Wilcox, Elizabeth Kaziunas, Stina Matthiesen, and Farah Magrabi. 2021. Realizing AI in Healthcare: Challenges Appearing in the Wild. In *Extended Abstracts of the 2021 CHI Conference on Human Factors in Computing Systems (CHI EA '21)*. Association for Computing Machinery, New York, NY, USA, Article 108, 1–5. https://doi.org/10.1145/3411763.3441347

[13] Plamen P Angelov, Nuria Oliver, Adrian Weller, Manuel Rodriguez, Isabel Valera, Silvia Chiappa, Hoda Heidari, and Niki Kilbertus. 2019. Workshop on Human-Centric Machine Learning. Retrieved from https://nips.cc/Conferences/2019/Schedule?showEvent$=$13203.

[14] Obinna Anya and Hissam Tawfik. 2018. An "Awareness" environment for clinical decision support in e-health. In *International Conference on Bioinformatics and Biomedical Engineering (IWBBIO 2018)*. Springer, Granada, Spain, 456-467. https://doi-org/10.1007/978-3-319-78759-6_41~

[15] Alvaro Aranda-Muñoz, Ulrika Florin, Yuji Yamamoto, Yvonne Eriksson, and Kristian Sandström. 2022. Co-Designing with AI in Sight. *Proceedings of the Design Society* 2 (May 2022), 101-110. https://doi.org/10.1017/pds.2022.11~

[16] Andre Artelt, Fabian Hinder, Valerie Vaquet, Robert Feldhans, and Barbara Hammer. 2022. Contrasting Explanations for Understanding and Regularizing Model Adaptations. *Neural Processing Letters* (May 2022), 1-25. https://doi.org/10.1007/s11063-022-10826-5~

[17] Martin Atzmueller, Cicek Güven, Parisa Shayan, Spyroula Masiala, Rick Mackenbach, and Werner Liebregts. 2019. Observing and Modeling User Behavior on Socio-Spatial Interaction Networks: Conformance, Exceptions, and Anomalies.

In *2019 First International Conference on Transdisciplinary AI (TransAI).* IEEE, Laguna Hills, CA, United States, 27-34. https://doi.org/10.1109/TransAI46475.2019.00013~
[18] Jan Auernhammer. 2020. Human-centered AI: The role of Human-centered Design Research in the development of AI, in Boess, S., Cheung, M. and Cain, R. (eds.), *Synergy - DRS International Conference 2020,* 11-14 August, Held online. https://doi.org/10.21606/drs.2020.282~
[19] Amid Ayobi, Katarzyna Stawarz, Dmitri Katz, Paul Marshall, Taku Yamagata, Raul Santos-Rodriguez, Peter Flach, and Aisling Ann O'Kane. 2021. Co-Designing Personal Health? Multidisciplinary Benefits and Challenges in Informing Diabetes Self-Care Technologies. *Proc. ACM Hum.-Comput. Interact.* 5, CSCW2, Article 457 (October 2021), 26 pages. https://doi.org/10.1145/3479601
[20] Daria Baidakova, Fabio Casati, Alexey Drutsa, and Dmitry Ustalov. 2020. Crowd Science Workshop: Remoteness, Fairness, and Mechanisms as Challenges of Data Supply by Humans for Automation. Retrieved from https://neurips.cc/virtual/2020/public/workshop_16111.html.
[21] Matthias Baldauf, Peter Fröhlich, Shadan Sadeghian, Philippe Palanque, Virpi Roto, Wendy Ju, Lynne Baillie, and Manfred Tscheligi. 2021. Automation Experience at the Workplace. In *Extended Abstracts of the 2021 CHI Conference on Human Factors in Computing Systems (CHI EA '21).* Association for Computing Machinery, New York, NY, USA, Article 89, 1–6. https://doi.org/10.1145/3411763.3441332
[22] Gagan Bansal, Alison Marie Smith-Renner, Zana Buçinca, Tongshuang Wu, Kenneth Holstein, Jessica Hullman, and Simone Stumpf. 2022. Workshop on Trust and Reliance in AI-Human Teams (TRAIT). In *Extended Abstracts of the 2022 CHI Conference on Human Factors in Computing Systems (CHI EA '22).* Association for Computing Machinery, New York, NY, USA, Article 116, 1–6. https://doi.org/10.1145/3491101.3503704
[23] Gagan Bansal, Tongshuang Wu, Joyce Zhou, Raymond Fok, Besmira Nushi, Ece Kamar, Marco Tulio Ribeiro, and Daniel Weld. 2021. Does the Whole Exceed its Parts? The Effect of AI Explanations on Complementary Team Performance. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems (CHI '21).* Association for Computing Machinery, New York, NY, USA, Article 81, 1–16. https://doi.org/10.1145/3411764.3445717
[24] Thomas Baudel, Manon Verbockhaven, Victoire Cousergue, Guillaume Roy, and Rida Laarach. 2021. ObjectivAIze: Measuring Performance and Biases in Augmented Business Decision Systems. In *IFIP Conference on Human-Computer Interaction (INTERACT 2021).* Springer, Cham, Bari, Italy, 300-320. https://doi.org/10.1007/978-3-030-85613-7_22~
[25] Anshul Bawa, Pranav Khadpe, Pratik Joshi, Kalika Bali, and Monojit Choudhury. 2020. Do Multilingual Users Prefer Chat-bots that Code-mix? Let's Nudge and Find Out! *Proc. ACM Hum.-Comput. Interact.* 4, CSCW1, Article 41 (May 2020), 23 pages. https://doi.org/10.1145/3392846
[26] Emma Beede, Elizabeth Baylor, Fred Hersch, Anna Iurchenko, Lauren Wilcox, Paisan Ruamviboonsuk, and Laura M. Vardoulakis. 2020. A Human-Centered Evaluation of a Deep Learning System Deployed in Clinics for the Detection of Diabetic Retinopathy. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems (CHI '20).* Association for Computing Machinery, New York, NY, USA, 1–12. https://doi.org/10.1145/3313831.3376718
[27] Ahmad Beirami, Emily Black, Krishna Gummadi, Hoda Heidari, Baharan Mirzasoleiman, Meisam Razaviyayn, and Joshua Williams. 2021. Workshop on Responsible AI. Retrieved from https://iclr.cc/virtual/2021/workshop/2132.
[28] Thierry Bellet, Aurélie Banet, Mari Petiot, Bertrand Richard, and Joshua Quick. 2020. Human-Centered AI to Support an Adaptive Management of Human-Machine Transitions with Vehicle Automation. *Information* 12, 1, Article 13 (Dec. 2020), 18 pages. https://doi.org/10.3390/info12010013~
[29] Emily M. Bender, Timnit Gebru, Angelina McMillan-Major, and Shmargaret Shmitchell. 2021. On the Dangers of Stochastic Parrots: Can Language Models Be Too Big?. In *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency (FAccT '21).* Association for Computing Machinery, New York, NY, USA, 610–623. https://doi.org/10.1145/3442188.3445922~
[30] Jesse Josua Benjamin, Arne Berger, Nick Merrill, and James Pierce. 2021. Machine Learning Uncertainty as a Design Material: A Post-Phenomenological Inquiry. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems (CHI '21).* Association for Computing Machinery, New York, NY, USA, Article 171, 1–14. https://doi.org/10.1145/3411764.3445481~
[31] Jesse Josua Benjamin, Christoph Kinkeldey, Claudia Müller-Birn, Tim Korjakow, and Eva-Maria Herbst. 2022. Explanation Strategies as an Empirical-Analytical Lens for Socio-Technical Contextualization of Machine Learning Interpretability. *Proc. ACM Hum.-Comput. Interact.* 6, GROUP, Article 39 (January 2022), 25 pages. https://doi.org/10.1145/3492858
[32] Cynthia L. Bennett, Daniela K. Rosner, and Alex S. Taylor. 2020. The Care Work of Access. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems (CHI '20).* Association for Computing Machinery, New York, NY, USA, 1-15. https://doi.org/10.1145/3313831.3376568~
[33] Emil Bergström, and Pontus Wärnestål. 2022. Exploring the Design Context of AI-Powered Services: A Qualitative Investigation of Designers' Experiences with Machine Learning. In *International Conference on Human-Computer Interaction (HCII '22).* Springer, Cham, 3-21. https://doi.org/10.1007/978-3-031-05643-7_1
[34] Jürgen Bernard, Marco Hutter, Heiko Reinemuth, Hendrik Pfeifer, Christian Bors, and Jörn Kohlhammer. 2019. Visual-Interactive Preprocessing of Multivariate Time Series Data. *Computer Graphics Forum* 38, 3 (July 2019) 401-412. https://doi.org/10.1111/cgf.13698~
[35] Steve J. Bickley and Benno Torgler. 2022. Cognitive architectures for artificial intelligence ethics. *AI & SOCIETY* (June 2022) 1-19. https://doi.org/10.1007/s00146-022-01452-9~
[36] Oloff C. Biermann, Ning F. Ma, and Dongwook Yoon. 2022. From Tool to Companion: Storywriters Want AI Writers to Respect Their Personal Values and Writing Strategies. In *Designing Interactive Systems Conference (DIS '22).* Association for Computing Machinery, New York, NY, USA, 1209–1227. https://doi.org/10.1145/3532106.3533506
[37] Sergiu Bilc, Adrian Groza, George Muntean, and Simona Delia Nicoara. 2021. Interleaving Automatic Segmentation and Expert Opinion for Retinal Conditions. *Diagnostics* 12, 1, Article 22 (Dec. 2021) 14 pages. https://doi.org/10.3390/diagnostics12010022~
[38] Christopher M. Bishop. 2006. *Pattern Recognition and Machine Learning.* Springer, New York, NY.
[39] Alan F. Blackwell. 2015. Interacting with an inferred world: the challenge of machine learning for humane computer interaction. In *Proceedings of The Fifth Decennial Aarhus Conference on Critical Alternatives (CA '15).* Aarhus University Press, Aarhus N, 169–180. https://doi.org/10.7146/aahcc.v1i1.21197~
[40] Brandon M. Booth, Louis Hickman, Shree Krishna Subburaj, Louis Tay, Sang Eun Woo, and Sidney K. D'Mello. 2021. Bias and Fairness in Multimodal Machine Learning: A Case Study of Automated Video Interviews. In *Proceedings of the 2021 International Conference on Multimodal Interaction (ICMI '21).* Association for Computing Machinery, New York, NY, USA, 268–277. https://doi.org/10.1145/3462244.3479897
[41] Tone Bratteteig and Guri Verne. 2018. Does AI make PD obsolete? exploring challenges from artificial intelligence to participatory design. In *Proceedings of the 15th Participatory Design Conference: Short Papers, Situated Actions, Workshops and Tutorial - Volume 2 (PDC '18).* Association for Computing Machinery, New York, NY, USA, Article 8, 1–5. https://doi.org/10.1145/3210604.3210646
[42] Virginia Braun and Victoria Clarke. 2006. Using thematic analysis in psychology.*Qualitative Research in Psychology*3, 2 (July 2008), 77-101. https://doi.org/10.1191/1478088706qp063oa
[43] Simone Borsci, Ville V. Lehtola, Francesco Nex, Michael Ying Yang,Ellen-Wien Augustijn, Leila Bagheriye, Christoph Brune, Ourania Kounadi, Jamy Li, Joao Moreira, Joanne Van Der Nagel, Bernard Veldkamp, Duc V. Le, Mingshu Wang, Fons Wijnhoven, Jelmer M. Wolterink, and Raul Zurita-Milla. 2022. Embedding artificial intelligence in society: looking beyond the EU AI master plan using the culture cycle. *AI & society* (Jan. 2022) 1-20. https://doi.org/10.1007/s00146-021-01383-x~
[44] Clara Bove, Jonathan Aigrain, Marie-Jeanne Lesot, Charles Tijus, and Marcin Detyniecki. 2022. Contextualization and Exploration of Local Feature Importance Explanations to Improve Understanding and Satisfaction of Non-Expert Users. In *27th International Conference on Intelligent User Interfaces (IUI '22).* Association for Computing Machinery, New York, NY, USA, 807–819. https://doi.org/10.1145/3490099.3511139
[45] Karen L. Boyd. 2021. Datasheets for Datasets help ML Engineers Notice and Understand Ethical Issues in Training Data. *Proc. ACM Hum.-Comput. Interact.* 5, CSCW2, Article 438 (October 2021), 27 pages. https://doi.org/10.1145/3479582
[46] Erik Brynjolfsson and Andrew Mcafee. 2017. Artificial Intelligence, For Real. *Harvard Business Review* 1 (July 2017), 1-31.
[47] P. Burggräf, J. Wagner,and T.M. Saßmannshausen. 2020. Sustainable interaction of human and artificial intelligence in cyber production management systems. In *Congress of the German Academic Association for Production Technology (WGP '20).* Springer, Berlin, Heidelberg, 508-517. https://doi.org/10.1007/978-3-662-62138-7_51
[48] Jenna Burrell. 2016. How the machine 'thinks': Understanding opacity in machine learning algorithms. *Big data & society* 3, 1 (Jan. 2016), 205395171562251. https://doi.org/10.1177/2053951715622512
[49] Federico Cabitza, Andrea Campagner, and Carla Simone. 2021. The need to move away from agential-AI: Empirical investigations, useful concepts and open issues. *International Journal of Human-Computer Studies* 155 (Nov. 2021), 102696. https://doi-org/10.1016/j.ijhcs.2021.102696~
[50] Xavier V Caddle, Afsaneh Razi, Seunghyun Kim, Shiza Ali, Temi Popo, Gianluca Stringhini, Munmun De Choudhury, and Pamela J. Wisniewski. 2021. MOSafely: Building an Open-Source HCAI Community to Make the Internet a Safer Place for Youth. In *Companion Publication of the 2021 Conference on Computer Supported Cooperative Work and Social Computing (CSCW '21).* Association for Computing Machinery, New York, NY, USA, 315–318. https://doi.org/10.1145/3462204.3481731
[51] Francisco Maria Calisto, Carlos Santiago, Nuno Nunes, and Jacinto C. Nascimento. 2021. Introduction of human-centric AI assistant to aid radiologists for multimodal breast image classification. *International Journal of Human-Computer Studies* 150 (June 2021) 102607. https://doi-org/10.1016/j.ijhcs.2021.

102607~
[52] Scott Allen Cambo and Darren Gergle. 2022. Model Positionality and Computational Reflexivity: Promoting Reflexivity in Data Science. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems (CHI '22)*. Association for Computing Machinery, New York, NY, USA, Article 572, 1–19. https://doi.org/10.1145/3491102.3501998
[53] Andrea Campagner, Federico Sternini, and Federico Cabitza. 2022. Decisions are not all equal. Introducing a utility metric based on case-wise raters' perceptions. *Computer Methods and Programs in Biomedicine* 221 (June 2022) 106930. https://doi-org/10.1016/j.cmpb.2022.106930~
[54] Margarita Robles Carrillo. 2020. Artificial intelligence: From ethics to law. *Telecommunications Policy* 44, 6 (July 2020) 101937. https://doi.org/10.1016/j.telpol.2020.101937~
[55] Andreas Cebulla, Zygmunt Szpak, Catherine Howell, Genevieve Knight, and Sazzad Hussain. 2022. Applying ethics to AI in the workplace: the design of a scorecard for Australian workplace health and safety. *AI & society* (May 2022) 1-17. https://doi.org/10.1007/s00146-022-01460-9~
[56] Joana Cerejo and Miguel Carvalhais. 2020. The lens of Anticipatory Design under AI-driven Services. In *International Conference on Digital Design & Communication (DIGICOM 2020)* ATAS: PT & ES, Barcelos, Portugal, 345-357.
[57] Rémy Chaput, Jérémy Duval, Olivier Boissier, Mathieu Guillermin, and Salima Hassas. 2021. A Multi-Agent Approach to Combine Reasoning and Learning for an Ethical Behavior. In *Proceedings of the 2021 AAAI/ACM Conference on AI, Ethics, and Society (AIES '21)*. Association for Computing Machinery, New York, NY, USA, 13–23. https://doi.org/10.1145/3461702.3462515
[58] Angelos Chatzimparmpas, Rafael M. Martins, Kostiantyn Kucher, and Andreas Kerren. 2020. StackGenVis: Alignment of data, algorithms, and models for stacking ensemble learning using performance metrics. *IEEE Transactions on Visualization and Computer Graphics* 27, 2 (Oct. 2020) 1547-1557. https://doi.org/10.1109/TVCG.2020.3030352~
[59] Hsing-Chung Chen, Agung Mulyo Widodo, Andika Wisnujati, Mosiur Rahaman, Jerry Chun-Wei Lin, Liukui Chen, and Chien-Erh Weng. 2022. AlexNet convolutional neural network for disease detection and classification of tomato leaf. *Electronics* 11, 6 (2022) 951. https://doi.org/10.3390/electronics11060951
[60] Nan-Chen Chen, Margaret Drouhard, Rafal Kocielnik, Jina Suh, and Cecilia R. Aragon. 2018. Using Machine Learning to Support Qualitative Coding in Social Science: Shifting the Focus to Ambiguity. *ACM Trans. Interact. Intell. Syst.* 8, 2, Article 9 (June 2018), 20 pages. https://doi.org/10.1145/3185515~~
[61] Hao-Fei Cheng, Logan Stapleton, Anna Kawakami, Venkatesh Sivaraman, Yanghuidi Cheng, Diana Qing, Adam Perer, Kenneth Holstein, Zhiwei Steven Wu, and Haiyi Zhu. 2022. How Child Welfare Workers Reduce Racial Disparities in Algorithmic Decisions. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems (CHI '22)*. Association for Computing Machinery, New York, NY, USA, Article 162, 1–22. https://doi.org/10.1145/3491102.3501831
[62] Hao-Fei Cheng, Logan Stapleton, Ruiqi Wang, Paige Bullock, Alexandra Chouldechova, Zhiwei Steven Wu, and Haiyi Zhu. 2021. Soliciting Stakeholders' Fairness Notions in Child Maltreatment Predictive Systems. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems (CHI '21)*. Association for Computing Machinery, New York, NY, USA, Article 390, 1–17. https://doi.org/10.1145/3411764.3445308
[63] Woan-Shiuan Chien, Huang-Cheng Chou, and Chi-Chun Lee. 2021. Self-assessed Emotion Classification from Acoustic and Physiological Features within Small-group Conversation. In *Companion Publication of the 2021 International Conference on Multimodal Interaction (ICMI '21 Companion)*. Association for Computing Machinery, New York, NY, USA, 230–239. https://doi.org/10.1145/3461615.3485411
[64] Prerna Chikersal, Danielle Belgrave, Gavin Doherty, Angel Enrique, Jorge E. Palacios, Derek Richards, and Anja Thieme. 2020. Understanding Client Support Strategies to Improve Clinical Outcomes in an Online Mental Health Intervention. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems (CHI '20)*. Association for Computing Machinery, New York, NY, USA, 1–16. https://doi.org/10.1145/3313831.3376341
[65] Douglas Cirqueira, Markus Helfert, and Marija Bezbradica. 2021. Towards design principles for user-centric explainable AI in fraud detection. In *International Conference on Human-Computer Interaction (HCII '21)*. Springer, Cham, 21-40. https://doi.org/10.1007/978-3-030-77772-2_2
[66] Nancy Cooke, Mustafa Demir, and Lixiao Huang. 2020. A framework for human-autonomy team research. In *International Conference on Human-Computer Interaction (HCII '20)*. Springer, Cham, 134-146. https://doi.org/10.1007/978-3-030-49183-3_11
[67] Joseph Corneli, Simon Holland, Alison Pease, Paul Mulholland, Dave Murray-Rust, Theodore Scaltsas, and Alan Smaill. 2018. Patterns of Design. In *Proceedings of the 23rd European Conference on Pattern Languages of Programs (EuroPLoP '18)*. Association for Computing Machinery, New York, NY, USA, Article 22, 1–11. https://doi.org/10.1145/3282308.3282331
[68] Anamaria Crisan, Margaret Drouhard, Jesse Vig, and Nazneen Rajani. 2022. Interactive Model Cards: A Human-Centered Approach to Model Documentation. In *2022 ACM Conference on Fairness, Accountability, and Transparency (FAccT '22)*. Association for Computing Machinery, New York, NY, USA, 427–439. https://doi.org/10.1145/3531146.3533108
[69] Stefan Cronholm and Hannes Göbel. 2022. Design Principles for Human-Centred AI. In *2022 European Conference on Information Systems (ECIS '22)*. Timisoara, Romania, 18–24.
[70] Rita Cucchiara and Matteo Fabbri. 2022. Fine-grained Human Analysis under Occlusions and Perspective Constraints in Multimedia Surveillance. *ACM Trans. Multimedia Comput. Commun. Appl.* 18, 1s, Article 32 (February 2022), 23 pages. https://doi.org/10.1145/3476839
[71] Syed Nasir Danial, Jennifer Smith, Brian Veitch, and Faisal Khan. 2019. On the realization of the recognition-primed decision model for artificial agents. *Human-centric Computing and Information Sciences* 9, 1 (Oct. 2019) 1-38. DOI: https://doi.org/10.1186/s13673-019-0197-2
[72] Subhajit Das, Dylan Cashman, Remco Chang, and Alex Endert. 2019. Gaggle: Visual Analytics for Model Space Navigation. In *Proceedings of Graphics Interface (GI '20)*. University of Toronto, 137-147. https://doi.org/10.20380/GI2020.15
[73] Mehdi Dastani and Vahid Yazdanpanah. 2022. Responsibility of AI Systems. *AI & SOCIETY* (June 2022) 1-10. https://doi.org/10.1007/s00146-022-01481-4
[74] Helmut Degen, Christoph J. Budnik, Kunal Chitre, and Andrew Lintereur. 2021. How to Explain It to Facility Managers? A Qualitative, Industrial User Research Study for Explainability. In *International Conference on Human-Computer Interaction (HCII '21)*. Springer, Cham, 401-422. https://doi.org/10.1007/978-3-030-90963-5_31
[75] Helmut Degen and Stavroula Ntoa. 2021. From a workshop to a framework for human-centered Artificial Intelligence. In *International Conference on Human-Computer Interaction (HCII '21)*. Springer, Cham, 166-184. https://doi.org/10.1007/978-3-030-77772-2_11
[76] Mustafa Demir, Nathan J. McNeese, and Nancy J. Cooke. 2019. The evolution of human-autonomy teams in remotely piloted aircraft systems operations. *Frontiers in Communication* 4, 50, Article 50 (Sep. 2019), 12 pages. DOI: https://doi.org/10.3389/fcomm.2019.00050
[77] Arturo Deza, Joshua Peterson, Apurva Ratan Murty, and Thomas Griffiths. 2022. Shared Visual Representations in Human and Machine Intelligence. Retrieved from https://www.svrhm.com/.
[78] Shipi Dhanorkar, Christine T. Wolf, Kun Qian, Anbang Xu, Lucian Popa, and Yunyao Li. 2021. Who needs to know what, when?: Broadening the Explainable AI (XAI) Design Space by Looking at Explanations Across the AI Lifecycle. In *Designing Interactive Systems Conference 2021 (DIS '21)*. Association for Computing Machinery, New York, NY, USA, 1591–1602. https://doi.org/10.1145/3461778.3462131
[79] Natalia Díaz-Rodríguez, Alberto Lamas, Jules Sanchez, Gianni Franchi, Ivan Donadello, Siham Tabik, David Filliat, Policarpo Cruz, Rosana Montes, and Francisco Herrera. 2022. EXplainable Neural-Symbolic Learning (X-NeSyL) methodology to fuse deep learning representations with expert knowledge graphs: The MonuMAI cultural heritage use case. *Information Fusion* 79 (March 2022), 58-83. https://doi.org/10.48550/arXiv.2104.11914
[80] Murat Dikmen and Catherine Burns. 2022. The effects of domain knowledge on trust in explainable AI and task performance: A case of peer-to-peer lending. *International Journal of Human-Computer Studies* 162 (June 2022), 102792. https://doi.org/10.1016/j.ijhcs.2022.102792
[81] Liya Ding. 2018. Human Knowledge in Constructing AI Systems — Neural Logic Networks Approach towards an Explainable AI. *Procedia Comput. Sci.* 126, C (2018), 1561–1570. https://doi.org/10.1016/j.procs.2018.08.129
[82] Stephane Doncieux, Raja Chatila, Sirko Straube, and Frank Kirchner. 2022. Human-centered AI and robotics. *AI Perspectives* 4, 1 (Jan. 2022), 1-14. https://doi.org/10.1186/s42467-021-00014-x
[83] Xishuang Dong, Shouvon Sarker, and Lijun Qian. Integrating Human-in-the-loop into Swarm Learning for Decentralized Fake News Detection. In *2022 International Conference on Intelligent Data Science Technologies and Applications (IDSTA)*. IEEE, 46-53. https://doi.org/10.48550/arXiv.2201.02048
[84] Paul Dourish. 2001. Where the action is: the foundations of embodied interaction. MIT press, London, England.
[85] Paul Dourish. 2003. What we talk about when we talk about context. *Personal and ubiquitous computing* 8, 1 (Dec 2003), 19-30. https://doi-org/10.1007/s00779-003-0253-8
[86] Graham Dove, Kim Halskov, Jodi Forlizzi, and John Zimmerman. 2017. UX Design Innovation: Challenges for Working with Machine Learning as a Design Material. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems (CHI '17)*. Association for Computing Machinery, New York, NY, USA, 278–288. https://doi.org/10.1145/3025453.3025739
[87] Jaimie Drozdal, Justin Weisz, Dakuo Wang, Gaurav Dass, Bingsheng Yao, Changruo Zhao, Michael Muller, Lin Ju, and Hui Su. 2020. Trust in AutoML: exploring information needs for establishing trust in automated machine learning systems. In *Proceedings of the 25th International Conference on Intelligent User Interfaces (IUI '20)*. Association for Computing Machinery, New York, NY, USA, 297–307. https://doi.org/10.1145/3377325.3377501
[88] Pelle Ehn. 2008. Participation in design things. In *Proceedings of the Tenth Anniversary Conference on Participatory Design 2008 (PDC '08)*. Indiana University,

USA, 92–101.
[89] Upol Ehsan, Q. Vera Liao, Michael Muller, Mark O. Riedl, and Justin D. Weisz. 2021. Expanding Explainability: Towards Social Transparency in AI systems. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems (CHI '21).* Association for Computing Machinery, New York, NY, USA, Article 82, 1–19. https://doi.org/10.1145/3411764.3445188
[90] Upol Ehsan, Philipp Wintersberger, Q. Vera Liao, Martina Mara, Marc Streit, Sandra Wachter, Andreas Riener, and Mark O. Riedl. 2021. Operationalizing Human-Centered Perspectives in Explainable AI. In *Extended Abstracts of the 2021 CHI Conference on Human Factors in Computing Systems (CHI EA '21).* Association for Computing Machinery, New York, NY, USA, Article 94, 1–6. https://doi.org/10.1145/3411763.3441342
[91] Haroon Elahi, Aniello Castiglione, Guojun Wang, and Oana Geman. 2021. A human-centered artificial intelligence approach for privacy protection of elderly App users in smart cities. *Neurocomputing* 444 (July 2021), 189-202. https://doi.org/10.1016/j.neucom.2020.06.149
[92] Luba Elliott, Sander Dieleman, Rebecca Fiebrink, Jesse Engel, Adam Roberts, and Tom White. 2018. Second Workshop on Machine Learning for Creativity and Design. Retrieved from https://nips.cc/Conferences/2018/Schedule?showEvent$=$10924.
[93] Christos Emmanouilidis, Sabine Waschull, Jos Bokhorst, and Hans Wortmann. 2021. Human in the AI Loop in Production Environments. In *IFIP International Conference on Advances in Production Management Systems (APMS '21).* Springer, Cham, 331-342. https://doi.org/10.1007/978-3-030-85910-7_35
[94] Daniel Estrada. 2020. Human supremacy as posthuman risk. *The Journal of Sociotechnical Critique* 1, 1 (May, 2020), 1-40. https://doi.org/10.25779/j5ps-dy87
[95] Theodore Evans, Carl Orge Retzlaff, Christian Geißler, Michaela Kargl, Markus Plass, Heimo Müller, Tim-Rasmus Kiehl, Norman Zerbe, and Andrea Holzinger. 2022. The explainability paradox: Challenges for xAI in digital pathology. *Future Generation Computer Systems* 133 (Aug. 2022), 281-296. https://doi.org/10.1016/j.future.2022.03.009
[96] Gregory Ewing and Ibrahim Demir. 2021. An ethical decision-making framework with serious gaming: a smart water case study on flooding. *Journal of Hydroinformatics* 23, 3 (March 2021), 466-482. https://doi.org/10.2166/hydro.2021.097~
[97] Jerry Alan Fails and Dan R. Olsen. 2003. Interactive machine learning. In *Proceedings of the 8th international conference on Intelligent user interfaces (IUI '03).* Association for Computing Machinery, New York, NY, USA, 39–45. https://doi.org/10.1145/604045.604056
[98] Sarah Fdili Alaoui, Jules Françoise, Thecla Schiphorst, Karen Studd, and Frederic Bevilacqua. 2017. Seeing, Sensing and Recognizing Laban Movement Qualities. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems (CHI '17).* Association for Computing Machinery, New York, NY, USA, 4009–4020. https://doi.org/10.1145/3025453.3025530~~
[99] Bruce Ferwerda, Marko Tkalcic, and Panagiotis Germanakos. 2022. Sixth HUMANIZE Workshop on Transparency and Explainability in Adaptive Systems Through User Modeling Grounded in Psychological Theory: Summary. In *27th International Conference on Intelligent User Interfaces (IUI '22 Companion).* Association for Computing Machinery, New York, NY, USA, 10–12. https://doi.org/10.1145/3490100.3511167
[100] Jessica L. Feuston and Jed R. Brubaker. 2021. Putting Tools in Their Place: The Role of Time and Perspective in Human-AI Collaboration for Qualitative Analysis. *Proc. ACM Hum.-Comput. Interact.* 5, CSCW2, Article 469 (October 2021), 25 pages. https://doi.org/10.1145/3479856~~
[101] Francesca Foffano, Teresa Scantamburlo and Atia Cortés. 2022. Investing in AI for social good: an analysis of European national strategies. *AI & Society* (May, 2022), 1-22. https://doi-org/10.1007/s00146-022-01445-8
[102] Jules Françoise, Baptiste Caramiaux, and Téo Sanchez. 2021. Marcelle: Composing Interactive Machine Learning Workflows and Interfaces. In *The 34th Annual ACM Symposium on User Interface Software and Technology (UIST '21).* Association for Computing Machinery, New York, NY, USA, 39–53. https://doi.org/10.1145/3472749.3474734~~
[103] Krzysztof Z. Gajos and Lena Mamykina. 2022. Do People Engage Cognitively with AI? Impact of AI Assistance on Incidental Learning. In *27th International Conference on Intelligent User Interfaces (IUI '22).* Association for Computing Machinery, New York, NY, USA, 794–806. https://doi.org/10.1145/3490099.3511138~~
[104] Andrew Garbett, Ziedune Degutyte, James Hodge, and Arlene Astell. 2021. Towards Understanding People's Experiences of AI Computer Vision Fitness Instructor Apps. In *Designing Interactive Systems Conference 2021 (DIS '21).* Association for Computing Machinery, New York, NY, USA, 1619–1637. https://doi.org/10.1145/3461778.3462094~~
[105] Timnit Gebru, Jamie Morgenstern, Briana Vecchione, Jennifer Wortman Vaughan, Hanna Wallach, Hal Daumé III, and Kate Crawford. 2021. Datasheets for datasets. *Commun. ACM* 64, 12 (December 2021), 86–92. https://doi.org/10.1145/3458723
[106] Ilina Georgieva, Claudio Lazo, Tjerk Timan, and Anne Fleur van Veenstra. 2022. From AI ethics principles to data science practice: a reflection and a gap analysis based on recent frameworks and practical experience. *AI and Ethics* 2,4 (Jan. 2022), 697-711. https://doi-org /10.1007/s43681-021-00127-3
[107] Deepti Ghadiyaram, Abhimanyu Dubey, Angelina Wang, Laurens van der Maaten, Olga Russakovsky, Judy Hoffman, Dhruv Mahajan, and Han Zhao. 2021. Responsible Computer Vision. Retrieved from https://sites.google.com/view/rcv-cvpr2021.
[108] Robert Gianni, Santtu Lehtinen and Mika Nieminen. 2022. Governance of Responsible AI: From Ethical Guidelines to Cooperative Policies. *Frontiers in Computer Science* 4, Article 873437 (May 2022), 17 pages. https://doi.org/10.3389/fcomp.2022.873437
[109] Marco Gillies. 2019. Understanding the Role of Interactive Machine Learning in Movement Interaction Design. *ACM Trans. Comput.-Hum. Interact.* 26, 1, Article 5 (February 2019), 34 pages. https://doi.org/10.1145/3287307
[110] Marco Gillies, Rebecca Fiebrink, Atau Tanaka, Jérémie Garcia, Frédéric Bevilacqua, Alexis Heloir, Fabrizio Nunnari, Wendy Mackay, Saleema Amershi, Bongshin Lee, Nicolas d'Alessandro, Joëlle Tilmanne, Todd Kulesza, and Baptiste Caramiaux. 2016. Human-Centred Machine Learning. In *Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems (CHI EA '16).* Association for Computing Machinery, New York, NY, USA, 3558–3565. https://doi.org/10.1145/2851581.2856492
[111] Mariia Golovianko, Svitlana Gryshko, Vagan Terziyan, and Tuure Tuunanen, T. 2022. Responsible cognitive digital clones as decision-makers: A design science research study. *European Journal of Information Systems* (May 2022), 1-23. https://doi.org/10.1080/0960085X.2022.2073278
[112] Ian Goodfellow, Yoshua Bengio, and Aaron Courville. 2016. *Deep Learning.* MIT Press, Cambridge, Massachusetts.
[113] Ian Goodfellow, Tim Hwang, Bryce Goodman, and Mikel Rodriguez. 2017. Machine Deception. Retrieved from https://nips.cc/Conferences/2017/Schedule?showEvent$=$8763.
[114] Steven M. Goodman, Ping Liu, Dhruv Jain, Emma J. McDonnell, Jon E. Froehlich, and Leah Findlater. 2021. Toward User-Driven Sound Recognizer Personalization with People Who Are d/Deaf or Hard of Hearing. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 5, 2, Article 63 (June 2021), 23 pages. https://doi.org/10.1145/3463501~~
[115] Mark Graus, Bruce Ferwerda, Marko Tkalcic, and Panagiotis Germanakos. 2021. Fifth HUMANIZE workshop on Transparency and Explainability in Adaptive Systems through User Modeling Grounded in Psychological Theory: Summary. In *26th International Conference on Intelligent User Interfaces - Companion (IUI '21 Companion).* Association for Computing Machinery, New York, NY, USA, 21–23. https://doi.org/10.1145/3397482.3450708
[116] Nina Grgić-Hlača, Christoph Engel, and Krishna P. Gummadi. 2019. Human Decision Making with Machine Assistance: An Experiment on Bailing and Jailing. *Proc. ACM Hum.-Comput. Interact.* 3, CSCW, Article 178 (November 2019), 25 pages. https://doi.org/10.1145/3359280~~
[117] Hongyan Gu, Jingbin Huang, Lauren Hung, and Xiang 'Anthony' Chen. 2021. Lessons Learned from Designing an AI-Enabled Diagnosis Tool for Pathologists. *Proc. ACM Hum.-Comput. Interact.* 5, CSCW1, Article 10 (April 2021), 25 pages. https://doi.org/10.1145/3449084~~
[118] Dylan Hadfield-Menell, Jacob Steinhardt, David Duvenaud, David Krueger, and Anca Dragan. 2017. Aligned Artificial Intelligence. Retrieved from https://nips.cc/Conferences/2017/Schedule?showEvent$=$8794.
[119] Kaely Hall, Dong Whi Yoo, Wenrui Zhang, Mehrab Bin Morshed, Vedant Das Swain, Gregory D. Abowd, Munmun De Choudhury, Alex Endert, John Stasko, and Jennifer G Kim. 2022. Supporting the Contact Tracing Process with WiFi Location Data: Opportunities and Challenges. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems (CHI '22).* Association for Computing Machinery, New York, NY, USA, Article 76, 1–14. https://doi.org/10.1145/3491102.3517703~~
[120] Henry Han, Wentian Li, Jiacun Wang, Guimin Qin, and Xianya Qin. 2022. Enhance Explainability of Manifold Learning. *Neurocomputing* 500 (Aug. 2022), 877-895. https://doi-org/10.1016/j.neucom.2022.05.119
[121] Richard H.R. Harper. 2019. The Role of HCI in the Age of AI. *International Journal of Human–Computer Interaction* 35, 15 (June 2019), 1331-1344. https://doi-org/10.1080/10447318.2019.1631527
[122] Steve Harrison, Deborah Tatar, and Phoebe Sengers. 2007. The three paradigms of HCI. In *Alt. Chi. Session at the SIGCHI Conference on human factors in computing systems (CHI '07).* San Jose, California, USA, 1-18.
[123] Hongmei He, John Gray, Angelo Cangelosi, Qinggang Meng, T. Martin McGinnity, and Jörn Mehnen. 2021. The challenges and opportunities of human-centred AI for trustworthy robots and autonomous systems. In *2020 3rd International Conference on Intelligent Robotic and Control Engineering (IRCE).* IEEE, 68-74. https://doi.org/10.1109/IRCE50905.2020.9199244
[124] Behnam Hedayatnia, Rahul Goel, Shereen Oraby, Abigail See, Chandra Khatri, Y-Lan Boureau, Alborz Geramifard, Marilyn Walker, and Dilek Hakkani-Tur. 2020. Human in the Loop Dialogue Systems. Retrieved from https://neurips.cc/virtual/2020/workshop/16125.
[125] Jennifer Heier. 2021. Design Intelligence - Taking Further Steps Towards New Methods and Tools for Designing in the Age of AI. In *Artificial Intelligence in HCI: Second International Conference, AI-HCI 2021, Held as Part of the 23rd HCI*

International Conference, HCII 2021, Virtual Event, July 24–29, 2021, Proceedings. Springer-Verlag, Berlin, Heidelberg, 202–215. https://doi.org/10.1007/978-3-030-77772-2_13

[126] Marco Hellmann, Diana C. Hernandez-Bocanegra, and Jürgen Ziegler. 2022. Development of an Instrument for Measuring Users' Perception of Transparency in Recommender Systems. *system* 12 (March 2022) 7.

[127] Thomas Herrmann. 2022. Promoting Human Competences by Appropriate Modes of Interaction for Human-Centered-AI. In *Artificial Intelligence in HCI: 3rd International Conference, AI-HCI 2022, Held as Part of the 24th HCI International Conference, HCII 2022,* Virtual Event, June 26 – July 1, 2022, Proceedings. Springer-Verlag, Berlin, Heidelberg, 35–50. https://doi.org/10.1007/978-3-031-05643-7_3

[128] Thomas Herrmann. 2020. Socio-Technical Design of Hybrid Intelligence Systems – The Case of Predictive Maintenance. In *Artificial Intelligence in HCI: First International Conference, AI-HCI 2020, Held as Part of the 22nd HCI International Conference, HCII 2020,* Copenhagen, Denmark, July 19–24, 2020, Proceedings. Springer-Verlag, Berlin, Heidelberg, 298–309. https://doi.org/10.1007/978-3-030-50334-5_20

[129] Thomas Herrmann and Sabine Pfeiffer. 2022. Keeping the organization in the loop: a socio-technical extension of human-centered artificial intelligence. *AI & Society* (Feb. 2022), 1-20. https://doi-org/10.1007/s00146-022-01391-5~

[130] Hendrik Heuer and Andreas Breiter. 2020. How Fake News Affect Trust in the Output of a Machine Learning System for News Curation. In *Disinformation in Open Online Media: Second Multidisciplinary International Symposium, MISDOOM 2020,* Leiden, The Netherlands, October 26–27, 2020, Proceedings. Springer-Verlag, Berlin, Heidelberg, 18–36. https://doi.org/10.1007/978-3-030-61841-4_2

[131] Tad Hirsch, Kritzia Merced, Shrikanth Narayanan, Zac E. Imel, and David C. Atkins. 2017. Designing Contestability: Interaction Design, Machine Learning, and Mental Health. In *Proceedings of the 2017 Conference on Designing Interactive Systems (DIS '17).* Association for Computing Machinery, New York, NY, USA, 95–99. https://doi.org/10.1145/3064663.3064703~~

[132] Fred Hohman, Andrew Head, Rich Caruana, Robert DeLine, and Steven M. Drucker. 2019. Gamut: A Design Probe to Understand How Data Scientists Understand Machine Learning Models. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems (CHI '19).* Association for Computing Machinery, New York, NY, USA, Paper 579, 1–13. https://doi.org/10.1145/3290605.3300809~~

[133] Andreas Holzinger. 2021. The next frontier: AI we can really trust. In *Joint European Conference on Machine Learning and Knowledge Discovery in Databases (ECML PKDD 2021).* Springer, Cham, 427-440. https://doi.org/10.1007/978-3-030-93736-2_33

[134] Andreas Holzinger, Matthias Dehmer, Frank Emmert-Streib, Rita Cucchiara, Isabelle Augenstein, Javier Del Ser, Wojciech Samek, Igor Jurisica, and Natalia Díaz-Rodríguez. 2022. Information fusion as an integrative cross-cutting enabler to achieve robust, explainable, and trustworthy medical artificial intelligence. *Inf. Fusion* 79, C (Mar 2022), 263–278. https://doi.org/10.1016/j.inffus.2021.10.007

[135] Andreas Holzinger, Michaela Kargl, Bettina Kipperer, Peter Regitnig, Markus Plass, and Heimo Müller. 2022. Personas for Artificial Intelligence (AI) An Open Source Toolbox. *IEEE Access* 10 (Jan. 2022), 23732-23747. https://doi.org/10.1109/ACCESS.2022.3154776~

[136] Andreas Holzinger, Michael Kickmeier-Rust, and Heimo Müller. 2019. KANDINSKY Patterns as IQ-Test for Machine Learning. In *Machine Learning and Knowledge Extraction: Third IFIP TC 5, TC 12, WG 8.4, WG 8.9, WG 12.9 International Cross-Domain Conference, CD-MAKE 2019*, Canterbury, UK, August 26–29, 2019, Proceedings. Springer-Verlag, Berlin, Heidelberg, 1–14. https://doi.org/10.1007/978-3-030-29726-8_1

[137] Andreas Holzinger, Anna Saranti, Alessa Angerschmid, Carl Orge Retzlaff, Andreas Gronauer, Vladimir Pejakovic, Francisco Medel-Jimenez, Theresa Krexner, Christoph Gollob, and Karl Stampfer. 2022. Digital Transformation in Smart Farm and Forest Operations Needs Human-Centered AI: Challenges and Future Directions. *Sensors* 22, 8, Article 3043 (April, 2022), 35 pages. https://doi.org/10.3390/s22083043

[138] Md Naimul Hoque, Bhavya Ghai, and Niklas Elmqvist. 2022. DramatVis Personae: Visual Text Analytics for Identifying Social Biases in Creative Writing. In *Designing Interactive Systems Conference (DIS '22).* Association for Computing Machinery, New York, NY, USA, 1260–1276. https://doi.org/10.1145/3532106.3533526

[139] Ajit Jain, Andruid Kerne, Nic Lupfer, Gabriel Britain, Aaron Perrine, Yoonsuck Choe, John Keyser, and Ruihong Huang. 2021. Recognizing creative visual design: multiscale design characteristics in free-form web curation documents. In *Proceedings of the 21st ACM Symposium on Document Engineering (DocEng '21).* Association for Computing Machinery, New York, NY, USA, Article 18, 1–10. https://doi.org/10.1145/3469096.3469869

[140] Christian Janiesch, Patrick Zschech, and Kai Heinrich. 2021. Machine learning and deep learning. *Electronic Markets* 31, 3 (April, 2021), 685-695. https://doi.org/10.1007/s12525-021-00475-2~~

[141] Natasha Jaques, Edward Hughes, Jakob Foerster, Noam Brown, Kalesha Bullard, and Charlotte Smith. 2021. Cooperative AI. Retrieved from https://www.cooperativeai.com/workshop/neurips-2021.

[142] Nithesh Javvaji, Casper Harteveld, and Magy Seif El-Nasr. 2020. Understanding Player Patterns by Combining Knowledge-Based Data Abstraction with Interactive Visualization. In *Proceedings of the Annual Symposium on Computer-Human Interaction in Play (CHI PLAY '20).* Association for Computing Machinery, New York, NY, USA, 254–266. https://doi.org/10.1145/3410404.3414257~~

[143] Theodore Jensen. 2021. Disentangling Trust and Anthropomorphism Toward the Design of Human-Centered AI Systems. In *Artificial Intelligence in HCI: Second International Conference, AI-HCI 2021, Held as Part of the 23rd HCI International Conference, HCII 2021,* Virtual Event, July 24–29, 2021, Proceedings. Springer-Verlag, Berlin, Heidelberg, 41–58. https://doi.org/10.1007/978-3-030-77772-2_3

[144] Youngseung Jeon, Seungwan Jin, and Kyungsik Han. 2021. FANCY: Human-centered, Deep Learning-based Framework for Fashion Style Analysis. In *Proceedings of the Web Conference 2021 (WWW '21).* Association for Computing Machinery, New York, NY, USA, 2367–2378. https://doi.org/10.1145/3442381.3449833~~

[145] Jinglu Jiang, Surinder Kahai, and Ming Yang. 2022. Who needs explanation and when? Juggling explainable AI and user epistemic uncertainty. *Int. J. Hum.-Comput. Stud.* 165, C (Sep 2022). https://doi.org/10.1016/j.ijhcs.2022.102839

[146] Marina Johnson, Abdullah Albizri, Antoine Harfouche, and Samuel Fosso-Wamba. 2022. Integrating human knowledge into artificial intelligence for complex and ill-structured problems: Informed artificial intelligence. *Int. J. Inf. Manag.* 64, C (Jun 2022). https://doi.org/10.1016/j.ijinfomgt.2022.102479

[147] Pratyusha Kalluri. 2020. Don't ask if artificial intelligence is good or fair, ask how it shifts power. *Nature* 583, 7815 (July 2020), 169–169. https://doi.org/10.1038/d41586-020-02003-2~~

[148] Annika Kaltenhauser, Verena Rheinstädter, Andreas Butz, and Dieter P. Wallach. 2020. "You Have to Piece the Puzzle Together": Implications for Designing Decision Support in Intensive Care. In *Proceedings of the 2020 ACM Designing Interactive Systems Conference (DIS '20).* Association for Computing Machinery, New York, NY, USA, 1509–1522. https://doi.org/10.1145/3357236.3395436~~

[149] Enes Karaaslan, Ulas Bagci, and F. Necati Catbas. 2021. Attention-guided analysis of infrastructure damage with semi-supervised deep learning. *Automation in Construction*125(May 2021), 103634. https://doi-org/10.1016/j.autcon.2021.103634~

[150] Harmanpreet Kaur, Eytan Adar, Eric Gilbert, and Cliff Lampe. 2022. Sensible AI: Re-imagining Interpretability and Explainability using Sensemaking Theory. In *2022 ACM Conference on Fairness, Accountability, and Transparency (FAccT '22).* Association for Computing Machinery, New York, NY, USA, 702–714. https://doi.org/10.1145/3531146.3533135~~

[151] Hassan Khosravi, Simon Buckingham Shum, Guanliang Chen, Cristina Conati, Yi-Shan Tsai, Judy Kay, Simon Knight, Roberto Martinez-Maldonado, Shazia Sadiq, and Dragan Gašević. 2022. Explainable artificial intelligence in education. *Computers and Education: Artificial Intelligence* 3 (May 2022), 100074. https://doi.org/10.1016/j.caeai.2022.100074

[152] Kimon Kieslich, Birte Keller, and Christopher Starke. 2022. Artificial intelligence ethics by design. Evaluating public perception on the importance of ethical design principles of artificial intelligence. *Big Data & Society* 9, 1 (Jan. 2022), 20539517221092956. https://doi.org/10.1177/20539517221092956

[153] Hankyung Kim and Youn-kyung Lim. 2021. Teaching-Learning Interaction: A New Concept for Interaction Design to Support Reflective User Agency in Intelligent Systems. In *Designing Interactive Systems Conference 2021 (DIS '21).* Association for Computing Machinery, New York, NY, USA, 1544–1553. https://doi.org/10.1145/3461778.3462141~~

[154] Moon K. Kim and Ethan Trewhitt. 2022. SatisfAI: A Serious Tabletop Game to Reveal Human-AI Interaction Dynamics. In *Adaptive Instructional Systems: 4th International Conference, AIS 2022, Held as Part of the 24th HCI International Conference, HCII 2022,* Virtual Event, June 26 – July 1, 2022, Proceedings. Springer-Verlag, Berlin, Heidelberg, 174–189. https://doi.org/10.1007/978-3-031-05887-5_13

[155] Gertraud Koch, Chris Biemann, Isabel Eiser, Tim Fischer, Florian Schneider, Teresa Stumpf, and Alejandra Tijerina García. 2022. *D-WISE Tool Suite for the Sociology of Knowledge Approach to Discourse. In Culture and Computing: 10th International Conference, C&C 2022, Held as Part of the 24th HCI International Conference, HCII 2022,* Virtual Event, June 26 – July 1, 2022, Proceedings. Springer-Verlag, Berlin, Heidelberg, 68–83. https://doi.org/10.1007/978-3-031-05434-1_5

[156] Dorothea Koert, Maximilian Kircher, Vildan Salikutluk, Carlo D'Eramo, and Jan Peters. 2020. Multi-channel interactive reinforcement learning for sequential tasks. *Frontiers in Robotics and AI* 7, Article 97 (Sept. 2020), 19 pages. https://doi.org/10.3389/frobt.2020.00097

[157] Tomoko Komatsu, Marisela Gutierrez Lopez, Stephann Makri, Colin Porlezza, Glenda Cooper, Andrew MacFarlane, and Sondess Missaoui. 2020. AI should embody our values: Investigating journalistic values to inform AI technology design. In *Proceedings of the 11th Nordic Conference on Human-Computer Interaction: Shaping Experiences, Shaping Society (NordiCHI '20).* Association for Computing Machinery, New York, NY, USA, Article 11, 1–13. https://doi.org/10.1145/3419249.3420105~~

[158] Arzam Kotriwala, Benjamin Klöpper, Marcel Dix, Gayathri Gopalakrishnan, Dawid Ziobro, and Andreas Potschka. 2021. XAI for Operations in the Process Industry-Applications, Theses, and Research Directions. In *Proceedings of the AAAI 2021 Spring Symposium on Combining Machine Learning and Knowledge Engineering (AAAI MAKE 2021).* Palo Alto, California, USA, Article 26, 1-12.

[159] Theocharis Kravaris, Konstantinos Lentzos, Georgios Santipantakis, George A. Vouros, Gennady Andrienko, Natalia Andrienko, Ian Crook, Jose Manuel Cordero Garcia, and Enrique Iglesias Martinez. 2022. Explaining deep reinforcement learning decisions in complex multiagent settings: towards enabling automation in air traffic flow management. *Applied Intelligence* 53 (June 2022), 4063-4098. https://doi-org/10.1007/s10489-022-03605-1

[160] Sebastian Laacke, Regina Mueller, Georg Schomerus and Sabine Salloch. 2021. Artificial Intelligence, Social Media and Depression. A New Concept of Health-Related Digital Autonomy. *The American Journal of Bioethics* 21, 7 (Jan. 2021), 4-20. https://doi-org/10.1080/15265161.2020.1863515

[161] Arto Laitinen and Otto Sahlgren. 2021. AI Systems and Respect for Human Autonomy. *Frontiers in Artificial Intelligence* 4, Article 705164 (Oct. 2021), 14 pages. https://doi.org/10.3389%2Ffrai.2021.705164

[162] Chang-Shing Lee, Mei-Hui Wang, Wen-Kai Kuan, Sheng-Hui Huang, Yi-Lin Tsai, Zong-Han Ciou, Chen-Kang Yang, and Naoyuki Kubota. 2021. BCI-based hit-loop agent for human and AI robot co-learning with AIoT application. *Journal of Ambient Intelligence and Humanized Computing* (Oct. 2021), 1-25. https://doi-org/10.1007/s12652-021-03487-0~

[163] Min Hun Lee, Daniel P. Siewiorek, Asim Smailagic, Alexandre Bernardino, and Sergi Bermúdez i Badia. 2022. Towards Efficient Annotations for a Human-AI Collaborative, Clinical Decision Support System: A Case Study on Physical Stroke Rehabilitation Assessment. In *27th International Conference on Intelligent User Interfaces (IUI '22).* Association for Computing Machinery, New York, NY, USA, 4–14. https://doi.org/10.1145/3490099.3511112~~

[164] Min Hun Lee, Daniel P. Siewiorek, Asim Smailagic, Alexandre Bernardino, and Sergi Bermúdez i Badia. 2021. A Human-AI Collaborative Approach for Clinical Decision Making on Rehabilitation Assessment. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems (CHI '21).* Association for Computing Machinery, New York, NY, USA, Article 392, 1–14. https://doi.org/10.1145/3411764.3445472~~

[165] Min Kyung Lee, Daniel Kusbit, Anson Kahng, Ji Tae Kim, Xinran Yuan, Allissa Chan, Daniel See, Ritesh Noothigattu, Siheon Lee, Alexandros Psomas, and Ariel D. Procaccia. 2019. WeBuildAI: Participatory Framework for Algorithmic Governance. *Proc. ACM Hum.-Comput. Interact.* 3, CSCW, Article 181 (November 2019), 35 pages. https://doi.org/10.1145/3359283~~

[166] Sunok Lee, Minha Lee, and Sangsu Lee. 2022. What If Artificial Intelligence Become Completely Ambient in Our Daily Lives? Exploring Future Human-AI Interaction through High Fidelity Illustrations. *International Journal of Human–Computer Interaction* (June 2022), 1-19. https://doi-org/10.1080/10447318.2022.2080155

[167] Christoph Leitner, Robert Jarolim, Bernhard Englmair, Annika Kruse, Karen Andrea Lara Hernandez, Andreas Konrad, Eric Yung-Sheng Su, Jörg Schröttner, Luke A. Kelly, Glen A. Lichtwark, Markus Tilp and Christian Baumgartner. 2021. A Human-Centered Machine-Learning Approach for Muscle-Tendon Junction Tracking in Ultrasound Images. *IEEE Transactions on Biomedical Engineering* 69, 6 (June 2022), 1920-1930. https://doi.org/10.1109/tbme.2021.3130548

[168] Bruno Lepri, Nuria Oliver, and Alex Pentland. 2021. Ethical machines: The human-centric use of artificial intelligence. *IScience* 24, 3, Article 102249 (March 2021) 17 pages. https://doi.org/10.1016/j.isci.2021.102249~

[169] Jamy Li and Mark Chignell. 2022. FMEA-AI: AI fairness impact assessment using failure mode and effects analysis. *AI and Ethics* 2, 4 (March 2022) 837-850. https://doi-org./10.1007/s43681-022-00145-9

[170] Suk-Young Lim, Dong-Kyu Chae, and Sang-Chul Lee. 2022. Detecting Deepfake Voice Using Explainable Deep Learning Techniques. *Applied Sciences* 12, 8, Article 3926 (April 2022) 14 pages. https://doi.org/10.3390/app12083926

[171] Ji Hun Lim and Hun Yeong Kwon. 2021. A Study on the Modeling of Major Factors for the Principles of AI Ethics. In *DG.O2021: The 22nd Annual International Conference on Digital Government Research (DG.O'21).* Association for Computing Machinery, New York, NY, USA, 208–218. https://doi.org/10.1145/3463677.3463733

[172] Daria Loi, Christine T. Wolf, Jeanette L. Blomberg, Raphael Arar, and Margot Brereton. 2019. Co-designing AI Futures: Integrating AI Ethics, Social Computing, and Design. In *Companion Publication of the 2019 on Designing Interactive Systems Conference 2019 Companion (DIS '19 Companion).* Association for Computing Machinery, New York, NY, USA, 381–384. https://doi.org/10.1145/3301019.3320000

[173] Duri Long, Mikhail Jacob, and Brian Magerko. 2019. Designing Co-Creative AI for Public Spaces. In *Proceedings of the 2019 on Creativity and Cognition (C&C '19).* Association for Computing Machinery, New York, NY, USA, 271–284. https://doi.org/10.1145/3325480.3325504~~

[174] Luca Longo, Randy Goebel, Freddy Lecue, Peter Kieseberg, and Andreas Holzinger. 2020. Explainable artificial intelligence: Concepts, applications, research challenges and visions. In *International Cross-Domain Conference for Machine Learning and Knowledge Extraction (CD-MAKE 2020).* Springer, Cham, 1-16. https://doi.org/10.1007/978-3-030-57321-8_1~

[175] Owen HT Lu, Anna YQ Huang, Danny CL Tsai, and Stephen JH Yang. 2021. Expert-Authored and Machine-Generated Short-Answer Questions for Assessing Students Learning Performance. *Educational Technology & Society* 24, 3 (July 2021), 159-173.

[176] Margaux Luck, Tristan Sylvain, Joseph Paul Cohen, Arsene Fansi Tchango, Valentine Goddard, Aurelie Helouis, Yoshua Bengio, Sam Greydanus, Cody Wild, Taras Kucherenko, Arya Farahi, Jonathan Penn, Sean McGregor, Mark Crowley, Abhishek Gupta, Kenny Chen, Myriam Côté, and Rediet Abebe. 2018. AI for Social Good. Retrieved from https://nips.cc/Conferences/2018/Schedule?showEvent$=$10904.

[177] Jonas Lundberg, Mattias Arvola, and Karljohan Lundin Palmerius. 2021. Human Autonomy in Future Drone Traffic: Joint Human–AI Control in Temporal Cognitive Work. *Frontiers in Artificial Intelligence* 4, Article 704082 (July 2021), 13 pages. https://doi.org/10.3389/frai.2021.704082

[178] Batnasan Luvaanjalba and Bo-chiuan Su. 2022. An Epistemological Analysis of the "Brain in a Vat" Approach for the Philosophy of Artificial Intelligence. In *HCI in Business, Government and Organizations: 9th International Conference, HCIBGO 2022, Held as Part of the 24th HCI International Conference, HCII 2022,* Virtual Event, June 26 – July 1, 2022, Proceedings. Springer-Verlag, Berlin, Heidelberg, 97–111. https://doi.org/10.1007/978-3-031-05544-7_8

[179] Maximilian Mackeprang, Claudia Müller-Birn, and Maximilian Timo Stauss. 2019. Discovering the Sweet Spot of Human-Computer Configurations: A Case Study in Information Extraction. *Proc. ACM Hum.-Comput. Interact.* 3, CSCW, Article 195 (November 2019), 30 pages. https://doi.org/10.1145/3359297~~

[180] Theodora A. Maniou, and Andreas Veglis. 2020. Employing a chatbot for news dissemination during crisis: Design, implementation and evaluation. *Future Internet* 12, 7 (June 2020), 109. https://doi.org/10.3390/fi12070109~

[181] Sotiris Manitsaris, Gavriela Senteri, Dimitrios Makrygiannis, and Alina Glushkova. 2020. Human movement representation on multivariate time series for recognition of professional gestures and forecasting their trajectories. *Frontiers in Robotics and AI* 7, Article 80 (Aug. 2020), 20 pages. https://doi.org/10.3389/frobt.2020.00080

[182] Alessandro Mantelero and Maria Samantha Esposito. 2021. An evidence-based methodology for human rights impact assessment (HRIA) in the development of AI data-intensive systems. *Computer Law & Security Review* 41, Article 105561 (July 2021), 35 pages. https://doi-org/10.1016/j.clsr.2021.105561

[183] Lina Markauskaite, Rebecca Marrone, Oleksandra Poquet, Simon Knight, Roberto Martinez-Maldonado, Sarah Howard, Jo Tondeur, Maarten De Laat, Simon Buckingham Shum, Dragan Gašević and George Siemens. 2022. Rethinking the entwinement between artificial intelligence and human learning: What capabilities do learners need for a world with AI?. *Computers and Education: Artificial Intelligence* 3, Article 100056 (Jan. 2022), 16 pages. https://doi.org/10.1016/j.caeai.2022.100056

[184] Antoine Marot, Adrian Kelly, Matija Naglic, Vincent Barbesant, Jochen Cremer, Alexandru Stefanov, and Jan Viebahn. 2022. Perspectives on Future Power System Control Centers for Energy Transition. *Journal of Modern Power Systems and Clean Energy* 10, 2 (March 2022), 328-344. https://doi.org/10.35833/MPCE.2021.000673

[185] Noortje Marres. 2020. Co-existence or displacement: Do street trials of intelligent vehicles test society?. *The British Journal of Sociology* 71, 3 (Jan. 2020), 537-555. https://doi-org/10.1111/1468-4446.12730~

[186] Nijat Mehdiyev, Constantin Houy, Oliver Gutermuth, Lea Mayer, and Peter Fettke. 2021. Explainable Artificial Intelligence (XAI) Supporting Public Administration Processes–On the Potential of XAI in Tax Audit Processes. In *International Conference on Wirtschaftsinformatik (WI 2021).* Springer, Cham, 413-428. https://doi-org/10.1007/978-3-030-86790-4_28

[187] Silvan Mertes, Tobias Huber, Katharina Weitz, Alexander Heimerl, and Elisabeth André. 2022. GANterfactual—Counterfactual Explanations for Medical Non-experts Using Generative Adversarial Learning. *Frontiers in Artificial Intelligence* 5, Article 825565 (April 2022), 19 pages. https://doi-org/10.3389/frai.2022.825565~

[188] Milagros Miceli, Julian Posada, and Tianling Yang. 2022. Studying Up Machine Learning Data: Why Talk About Bias When We Mean Power? *Proc. ACM Hum.-Comput. Interact.* 6, GROUP, Article 34 (January 2022), 14 pages. https://doi.org/10.1145/3492853~~

[189] Chris J. Michael, Steven M. Dennis, Corey Maryan, Samuel Irving, and Margaret L. Palmsten. 2019. A General Framework for Human-Machine Digitization of Geographic Regions from Remotely Sensed Imagery. In *Proceedings of the 27th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems (SIGSPATIAL '19).* Association for Computing Machinery, New York, NY, USA, 259–268. https://doi.org/10.1145/3347146.3359370~~

[190] Swati Mishra and Jeffrey M Rzeszotarski. 2021. Designing Interactive Transfer Learning Tools for ML Non-Experts. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems (CHI '21).* Association for Computing Machinery, New York, NY, USA, Article 364, 1–15. https://doi.org/10.1145/3411764.3445096~~

[191] Margaret Mitchell, Simone Wu, Andrew Zaldivar, Parker Barnes, Lucy Vasserman, Ben Hutchinson, Elena Spitzer, Inioluwa Deborah Raji, and Timnit Gebru. 2019. Model Cards for Model Reporting. In *Proceedings of the Conference on Fairness, Accountability, and Transparency (FAT* '19)*. Association for Computing Machinery, New York, NY, USA, 220–229. https://doi.org/10.1145/3287560.3287596~~

[192] Jakub Mlynar, Farzaneh Bahrami, André Ourednik, Nico Mutzner, Himanshu Verma, and Hamed Alavi. 2022. AI beyond Deus ex Machina – Reimagining Intelligence in Future Cities with Urban Experts. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems (CHI '22)*. Association for Computing Machinery, New York, NY, USA, Article 370, 1–13. https://doi.org/10.1145/3491102.3517502~~

[193] Judith Molka-Danielsen, Jazz Rasool, and Carl H. Smith. 2022. Design and Deployment Considerations for Ethically Advanced Technologies for Human Flourishing in the Workplace. In *IFIP Working Conference on Human Work Interaction Design (HWID 2021)*. Springer, Cham 101-122. https://doi.org/10.1007/978-3-031-02904-2_5

[194] Cecily Morrison, Edward Cutrell, Martin Grayson, Elisabeth RB Becker, Vasiliki Kladouchou, Linda Pring, Katherine Jones, Rita Faia Marques, Camilla Longden, and Abigail Sellen. 2021. Enabling meaningful use of AI-infused educational technologies for children with blindness: Learnings from the development and piloting of the PeopleLens curriculum. In *The 23rd International ACM SIGACCESS Conference on Computers and Accessibility (ASSETS '21)*. Association for Computing Machinery, New York, NY, USA, Article 2, 1–13. https://doi.org/10.1145/3441852.3471210~~

[195] Cecily Morrison, Edward Cutrell, Martin Grayson, Anja Thieme, Alex Taylor, Geert Roumen, Camilla Longden, Sebastian Tschiatschek, Rita Faia Marques, and Abigail Sellen. 2021. Social Sensemaking with AI: Designing an Open-ended AI Experience with a Blind Child. *In Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems (CHI '21)*. Association for Computing Machinery, New York, NY, USA, Article 396, 1–14. https://doi.org/10.1145/3411764.3445290~~

[196] Cecily Morrison, Kit Huckvale, Bob Corish, Richard Banks, Martin Grayson, Jonas Dorn, Abigail Sellen, and Sân Lindley. 2018. Visualizing Ubiquitously Sensed Measures of Motor Ability in Multiple Sclerosis: Reflections on Communicating Machine Learning in Practice. *ACM Trans. Interact. Intell. Syst.* 8, 2, Article 12 (June 2018), 28 pages. https://doi.org/10.1145/3181670~~

[197] Heimo Müller and Andreas Holzinger. 2021. Kandinsky Patterns. *Artif. Intell.* 300, C (Nov. 2021). https://doi.org/10.1016/j.artint.2021.103546

[198] Michael Muller, Plamen P Angelov, Hal Daumé III, Shion Guha, Q.Vera Liao, Nuria Oliver, and David Piorkowski. 2022. Human Centered AI. Retrieved from https://nips.cc/virtual/2022/workshop/50008.

[199] Michael Muller, Plamen P Angelov, Shion Guha, Marina Kogan, Gina Neff, Nuria Oliver, Manuel Rodriguez, and Adrian Weller. 2021. Human Centered AI. Retrieved from https://nips.cc/Conferences/2021/Schedule?showEvent$=$21841.

[200] Michael Muller, Lydia B Chilton, Anna Kantosalo, Charles Patrick Martin, and Greg Walsh. 2022. GenAICHI: Generative AI and HCI. In *Extended Abstracts of the 2022 CHI Conference on Human Factors in Computing Systems (CHI EA '22)*. Association for Computing Machinery, New York, NY, USA, Article 110, 1–7. https://doi.org/10.1145/3491101.3503719

[201] Michael Muller and Justin Weisz. 2022. Extending a Human-AI Collaboration Framework with Dynamism and Sociality. In *2022 Symposium on Human-Computer Interaction for Work (CHIWORK 2022)*. Association for Computing Machinery, New York, NY, USA, Article 10, 1–12. https://doi.org/10.1145/3533406.3533407

[202] Michael Muller, Christine T. Wolf, Josh Andres, Michael Desmond, Narendra Nath Joshi, Zahra Ashktorab, Aabhas Sharma, Kristina Brimijoin, Qian Pan, Evelyn Duesterwald, and Casey Dugan. 2021. Designing Ground Truth and the Social Life of Labels. In *Proceedings of the 2021 CHI Conference on Ha gne Rödeluman Factors in Computing Systems (CHI '21)*. Association for Computing Machinery, New York, NY, USA, Article 94, 1–16. https://doi.org/10.1145/3411764.3445402~~

[203] Yuri Nakao, Lorenzo Strappelli, Simone Stumpf, Aisha Naseer, Daniele Regoli, and Giulia Del Gamba. 2022. Towards Responsible AI: A Design Space Exploration of Human-Centered Artificial Intelligence User Interfaces to Investigate Fairness. *International Journal of Human–Computer Interaction* (May 2022), 1-27. https://doi-org/10.1080/10447318.2022.2067936~

[204] Iohanna Nicenboim, Elisa Giaccardi, Marie Louise Juul Søndergaard, Anuradha Venugopal Reddy, Yolande Strengers, James Pierce, and Johan Redström. 2020. More-Than-Human Design and AI: In Conversation with Agents. In *Companion Publication of the 2020 ACM Designing Interactive Systems Conference (DIS' 20 Companion)*. Association for Computing Machinery, New York, NY, USA, 397–400. https://doi.org/10.1145/3393914.3395912

[205] An T. Nguyen, Aditya Kharosekar, Saumyaa Krishnan, Siddhesh Krishnan, Elizabeth Tate, Byron C. Wallace, and Matthew Lease. 2018. Believe it or not: Designing a Human-AI Partnership for Mixed-Initiative Fact-Checking. In *Proceedings of the 31st Annual ACM Symposium on User Interface Software and Technology (UIST '18)*. Association for Computing Machinery, New York, NY, USA, 189–199. https://doi.org/10.1145/3242587.3242666~~

[206] Renee Noortman, Peter Lovei, Mathias Funk, Eva Deckers, Stephan Wensveen, and Berry Eggen. 2022. Breaking up data-enabled design: expanding and scaling up for the clinical context. *Artificial Intelligence for Engineering Design, Analysis and Manufacturing* 36, 1 (May 2022), 1-13. https://doi.org/10.1017/S0890060421000433

[207] Mahsan Nourani, Chiradeep Roy, Jeremy E. Block, Donald R. Honeycutt, Tahrima Rahman, Eric D. Ragan, and Vibhav Gogate. 2022. On the Importance of User Backgrounds and Impressions: Lessons Learned from Interactive AI Applications. *ACM Trans. Interact. Intell. Syst.* 12, 4, Article 28 (December 2022), 29 pages. https://doi.org/10.1145/3531066

[208] Eshed Ohn-Bar and Mohan Manubhai Trivedi. 2017. Are all objects equal? Deep spatio-temporal importance prediction in driving videos. *Pattern Recogn.* 64, C (April 2017), 425–436. https://doi.org/10.1016/j.patcog.2016.08.029

[209] Andreea M. Oprescu, Gloria Miró-Amarante, Lutgardo García-Díaz, Victoria E. Rey, A. Chimenea-Toscano, Ricard Martínez-Martínez, and M. C. Romero-Ternero. 2022. Towards a data collection methodology for Responsible Artificial Intelligence in health: A prospective and qualitative study in pregnancy. *Information Fusion* 83 (July 2022), 53-78. https://doi-org/10.1016/j.inffus.2022.03.011~

[210] Kirsten Ostherr. 2020. Artificial intelligence and medical humanities. *Journal of Medical Humanities* 43, 2 (July 2020), 211-232. https://doi-org/10.1007/s10912-020-09636-4

[211] Giovanni Ottoboni, Fabio La Porta, Roberto Piperno, Rabih Chattat, Annalisa Bosco, Patrizia Fattori, and Alessia Tessari. 2022. A Multifunctional Adaptive and Interactive AI system to support people living with stroke, acquired brain or spinal cord injuries: A study protocol. *PloS one* 17, 4, Article e0266702 (April 2022), 10 pages. https://doi.org/10.1371/journal.pone.0266702

[212] Sharon Oviatt. 2021. Technology as Infrastructure for Dehumanization: Three Hundred Million People with the Same Face. In *Proceedings of the 2021 International Conference on Multimodal Interaction (ICMI '21)*. Association for Computing Machinery, New York, NY, USA, 278–287. https://doi.org/10.1145/3462244.3482855~~

[213] S. Oviatt, K. Hang, J. Zhou, K. Yu, and F. Chen. 2018. Dynamic Handwriting Signal Features Predict Domain Expertise. *ACM Trans. Interact. Intell. Syst.* 8, 3, Article 18 (September 2018), 21 pages. https://doi.org/10.1145/3213309~~

[214] Andrea Papenmeier, Dagmar Kern, Daniel Hienert, Yvonne Kammerer, and Christin Seifert. 2022. How Accurate Does It Feel? – Human Perception of Different Types of Classification Mistakes. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems (CHI '22)*. Association for Computing Machinery, New York, NY, USA, Article 180, 1–13. https://doi.org/10.1145/3491102.3501915~~

[215] Pradeep Paraman and Sanmugam Anamalah. 2022. Ethical artificial intelligence framework for a good AI society: principles, opportunities and perils. *AI & Society* (June 2022), 1-17. https://doi-org/10.1007/s00146-022-01458-3

[216] Hyanghee Park, Daehwan Ahn, Kartik Hosanagar, and Joonhwan Lee. 2022. Designing Fair AI in Human Resource Management: Understanding Tensions Surrounding Algorithmic Evaluation and Envisioning Stakeholder-Centered Solutions. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems (CHI '22)*. Association for Computing Machinery, New York, NY, USA, Article 51, 1–22. https://doi.org/10.1145/3491102.3517672~~

[217] Gaia Pavoni, Massimiliano Corsini, Federico Ponchio, Alessandro Muntoni, Clinton Edwards, Nicole Pedersen, Stuart Sandin, and Paolo Cignoni. 2022. TagLab: AI-assisted annotation for the fast and accurate semantic segmentation of coral reef orthoimages. *Journal of Field Robotics* 39, 3 (May 2022), 246-262. https://doi-org/10.1002/rob.22049

[218] Marieke MM Peeters, Jurriaan van Diggelen, Karel Van Den Bosch, Adelbert Bronkhorst, Mark A. Neerincx, Jan Maarten Schraagen, and Stephan Raaijmakers. 2021. Hybrid collective intelligence in a human–AI society. *AI & Society* 36, 1 (June 2020), 217-238. https://doi.org/10.1007/s00146-020-01005-y

[219] Debra P.C. Peters, N. Dylan Burruss, Luis L. Rodriguez, D. Scott McVey, Emile H. Elias, Angela M. Pelzel-McCluskey, Justin D. Derner, T. Scott Schrader, Jin Yao, Steven J. Pauszek, Jason Lombard, Steven R. Archer, Brandon T Bestelmeyer, Dawn M. Browning, Colby W. Brungard, Jerry L. Hatfield, Niall P. Hanan, Jeffrey E. Herrick, Gregory S. Okin, Osvaldo E. Sala, Heather Savoy and Enrique R. Vivoni. 2018. An integrated view of complex landscapes: a big data-model integration approach to transdisciplinary science. *BioScience* 68, 9 (Sept. 2018), 653-669. https://doi.org/10.1093/biosci/biy069~

[220] Evangelos Pournaras. 2020. Collective learning: A 10-year odyssey to human-centered distributed intelligence. In *2020 IEEE International Conference on Autonomic Computing and Self-Organizing Systems (ACSOS 20202)*. IEEE, 205-214. http://doi.org/10.1109/ACSOS49614.2020.00043

[221] Forough Poursabzi-Sangdeh, Daniel G Goldstein, Jake M Hofman, Jennifer Wortman Wortman Vaughan, and Hanna Wallach. 2021. Manipulating and Measuring Model Interpretability. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems (CHI '21)*. Association for Computing Machinery, New York, NY, USA, Article 237, 1–52. https://doi.org/10.1145/3411764.3445315~~

[222] Mahima Pushkarna, Andrew Zaldivar, and Oddur Kjartansson. 2022. Data Cards: Purposeful and Transparent Dataset Documentation for Responsible AI. In *2022 ACM Conference on Fairness, Accountability, and Transparency (FAccT '22)*. Association for Computing Machinery, New York, NY, USA, 1776–1826. https://doi.org/10.1145/3531146.3533231
[223] Ming Qian and Davis Qian. 2020. Defining a Human-Machine Teaming Model for AI-Powered Human-Centered Machine Translation Agent by Learning from Human-Human Group Discussion: Dialog Categories and Dialog Moves. In *Artificial Intelligence in HCI: First International Conference, AI-HCI 2020, Held as Part of the 22nd HCI International Conference, HCII 2020,* Copenhagen, Denmark, July 19–24, 2020, Proceedings. Springer-Verlag, Berlin, Heidelberg, 70–81. https://doi.org/10.1007/978-3-030-50334-5_5
[224] Emilee Rader, Kelley Cotter, and Janghee Cho. 2018. Explanations as Mechanisms for Supporting Algorithmic Transparency. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (CHI '18)*. Association for Computing Machinery, New York, NY, USA, Paper 103, 1–13. https://doi.org/10.1145/3173574.3173677~~
[225] Gonzalo Ramos, Jina Suh, Soroush Ghorashi, Christopher Meek, Richard Banks, Saleema Amershi, Rebecca Fiebrink, Alison Smith-Renner, and Gagan Bansal. 2019. Emerging Perspectives in Human-Centered Machine Learning. In Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems (CHI EA '19). Association for Computing Machinery, New York, NY, USA, Paper W11, 1–8. https://doi.org/10.1145/3290607.3299014
[226] Daniel Reichman, Joshua Peterson, Kiran Tomlinson, Annie Liang, and Tom Griffiths. 2021. Workshop on Human and Machine Decisions. Retrieved from https://nips.cc/Conferences/2021/Schedule?showEvent$=$21835.
[227] Jennifer Renoux, Jasmin Grosinger, Andrew Howes, Antti Oulasvirta, and Mohamed Chetouani. 2022. Communication in Human-AI Interaction Workshop. Retrieved from https://chai-workshop.github.io/.
[228] André Renz and Gergana Vladova. 2021. Reinvigorating the Discourse on Human-Centered Artificial Intelligence in Educational Technologies. *Technology Innovation Management Review* 11, 5 (May 2021), 5-16. https://doi.org/10.22215/timreview/1438
[229] Jimin Rhim, Ji-Hyun Lee, Mo Chen and Angelica Lim. 2021. A deeper look at autonomous vehicle ethics: an integrative ethical decision-making framework to explain moral pluralism. *Frontiers in Robotics and AI* 8, Article 632394 (May 2021) 18 pages. https://doi.org/10.3389/frobt.2021.632394~
[230] Dalai Dos Santos Ribeiro, Gabriel Diniz Junqueira Barbosa, Marisa Do Carmo Silva, Hélio Lopes, and Simone Diniz Junqueira Barbosa. 2021. Exploring the impact of classification probabilities on users' trust in ambiguous instances. In *2021 IEEE Symposium on Visual Languages and Human-Centric Computing (VL/HCC 2021)*. IEEE, 1-9. https://doi.org/10.1109/VL/HCC51201.2021.9576291~
[231] Maria Riveiro and Serge Thill. 2021. "That's (not) the output I expected!" On the role of end user expectations in creating explanations of AI systems. *Artif. Intell.* 298, C (Sep 2021). https://doi.org/10.1016/j.artint.2021.103507
[232] Samantha Robertson, Tonya Nguyen, and Niloufar Salehi. 2021. Modeling Assumptions Clash with the Real World: Transparency, Equity, and Community Challenges for Student Assignment Algorithms. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems (CHI '21)*. Association for Computing Machinery, New York, NY, USA, Article 589, 1–14. https://doi.org/10.1145/3411764.3445748~~
[233] Christina Rödel, Susanne Stadler, Alexander Meschtscherjakov, and Manfred Tscheligi. 2014. Towards Autonomous Cars: The Effect of Autonomy Levels on Acceptance and User Experience. In *Proceedings of the 6th International Conference on Automotive User Interfaces and Interactive Vehicular Applications (AutomotiveUI '14)*. Association for Computing Machinery, New York, NY, USA, 1–8. https://doi.org/10.1145/2667317.2667330
[234] Kat Roemmich and Nazanin Andalibi. 2021. Data Subjects' Conceptualizations of and Attitudes Toward Automatic Emotion Recognition-Enabled Wellbeing Interventions on Social Media. *Proc. ACM Hum.-Comput. Interact.* 5, CSCW2, Article 308 (October 2021), 34 pages. https://doi.org/10.1145/3476049~~
[235] Avi Rosenfeld and Ariella Richardson. 2019. Explainability in human–agent systems. *Autonomous Agents and Multi-Agent Systems* 33, 6 (Nov 2019), 673–705. https://doi.org/10.1007/s10458-019-09408-y
[236] Jože Rožanec, Elena Trajkova, Inna Novalija, Patrik Zajec, Klemen Kenda, Blaž Fortuna, and Dunja Mladenić. 2022. Enriching Artificial Intelligence Explanations with Knowledge Fragments. *Future Internet* 14, 5, Article 134 (April 2022), 14 pages. https://doi.org/10.3390/fi14050134~
[237] Stuart J. Russell and Peter Norvig. 2021. *Artificial Intelligence: A Modern Approach* (4th. ed.). Pearson Education.
[238] Dominik Sacha, Michael Sedlmair, Leishi Zhang, John A. Lee, Jaakko Peltonen, Daniel Weiskopf, Stephen C. North, and Daniel A. Keim. 2017. What you see is what you can change: Human-centered machine learning by interactive visualization. *Neurocomputing* 268 (Dec. 2017), 164-175. https://doi-org.ezproxy.library.uq.edu.au/10.1016/j.neucom.2017.01.105~
[239] Koustuv Saha, Ted Grover, Stephen M. Mattingly, Vedant Das swain, Pranshu Gupta, Gonzalo J. Martinez, Pablo Robles-Granda, Gloria Mark, Aaron Striegel, and Munmun De Choudhury. 2021. Person-Centered Predictions of Psychological Constructs with Social Media Contextualized by Multimodal Sensing. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 5, 1, Article 32 (March 2021), 32 pages. https://doi.org/10.1145/3448117~~
[240] Nithya Sambasivan, Shivani Kapania, Hannah Highfill, Diana Akrong, Praveen Paritosh, and Lora M Aroyo. 2021. "Everyone wants to do the model work, not the data work": Data Cascades in High-Stakes AI. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems (CHI '21)*. Association for Computing Machinery, New York, NY, USA, Article 39, 1–15. https://doi.org/10.1145/3411764.3445518~~
[241] Supraja Sankaran and Panos Markopoulos. 2021. "It's like a puppet master": User Perceptions of Personal Autonomy when Interacting with Intelligent Technologies. In *Proceedings of the 29th ACM Conference on User Modeling, Adaptation and Personalization (UMAP '21)*. Association for Computing Machinery, New York, NY, USA, 108–118. https://doi.org/10.1145/3450613.3456820
[242] Supraja Sankaran, Chao Zhang, Henk Aarts and Panos Markopoulos. 2021. Exploring Peoples' Perception of Autonomy and Reactance in Everyday AI Interactions. *Front Psychol* 12, Article 713074 (Sept. 2021), 14 pages. https://doi.org/10.3389/fpsyg.2021.713074
[243] Supraja Sankaran, Chao Zhang, Marisela Gutierrez Lopez, and Kaisa Väänänen. 2020. Respecting Human Autonomy through Human-Centered AI. In *Proceedings of the 11th Nordic Conference on Human-Computer Interaction: Shaping Experiences, Shaping Society (NordiCHI '20)*. Association for Computing Machinery, New York, NY, USA, Article 134, 1–3. https://doi.org/10.1145/3419249.3420098
[244] Sebastin Santy, Kalika Bali, Monojit Choudhury, Sandipan Dandapat, Tanuja Ganu, Anurag Shukla, Jahanvi Shah, and Vivek Seshadri. 2021. Language Translation as a Socio-Technical System:Case-Studies of Mixed-Initiative Interactions. In *ACM SIGCAS Conference on Computing and Sustainable Societies (COMPASS '21)*. Association for Computing Machinery, New York, NY, USA, 156–172. https://doi.org/10.1145/3460112.3471954
[245] Laura Sartori and Andreas Theodorou. 2022. A sociotechnical perspective for the future of AI: narratives, inequalities, and human control. *Ethics and Inf. Technol.* 24, 1 (Mar 2022). https://doi.org/10.1007/s10676-022-09624-3
[246] Sanjay Sekar Samuel, Nik Nailah Binti Abdullah and Anil Raj. 2020. Interpretation of SVM using data mining technique to extract syllogistic rules. In *International Cross-Domain Conference for Machine Learning and Knowledge Extraction (CD-MAKE 2020)*. Springer, Chalm, 249-266. https://doi.org/10.1007/978-3-030-57321-8_14
[247] Téo Sanchez, Baptiste Caramiaux, Jules Françoise, Frédéric Bevilacqua, and Wendy E. Mackay. 2021. How do People Train a Machine? Strategies and (Mis)Understandings. *Proc. ACM Hum.-Comput. Interact.* 5, CSCW1, Article 162 (April 2021), 26 pages. https://doi.org/10.1145/3449236~~
[248] Téo Sanchez, Baptiste Caramiaux, Pierre Thiel, and Wendy E. Mackay. 2022. Deep Learning Uncertainty in Machine Teaching. In *27th International Conference on Intelligent User Interfaces (IUI '22)*. Association for Computing Machinery, New York, NY, USA, 173–190. https://doi.org/10.1145/3490099.3511117
[249] Liz Sanders. 2008. ON MODELING: An evolving map of design practice and design research. *interactions 1*5, 6 (November + December 2008), 13–17. https://doi.org/10.1145/1409040.1409043
[250] Morgan Klaus Scheuerman, Alex Hanna, and Emily Denton. 2021. Do Datasets Have Politics? Disciplinary Values in Computer Vision Dataset Development. *Proc. ACM Hum.-Comput. Interact.* 5, CSCW2, Article 317 (October 2021), 37 pages. https://doi.org/10.1145/3476058~~
[251] Johannes Schleith, Milda Norkute, Mary Mikhail, and Daniella Tsar. 2022. Cognitive Strategy Prompts: Creativity Triggers for Human Centered AI Opportunity Detection. In *Creativity and Cognition (C&C '22)*. Association for Computing Machinery, New York, NY, USA, 29–37. https://doi.org/10.1145/3527927.3532808~~
[252] Stefan Schmager and Sonia Sousa. 2021. A Toolkit to Enable the Design of Trustworthy AI. In *HCI International 2021 - Late Breaking Papers: Multimodality, eXtended Reality, and Artificial Intelligence: 23rd HCI International Conference, HCII 2021,* Virtual Event, July 24–29, 2021, Proceedings. Springer-Verlag, Berlin, Heidelberg, 536–555. https://doi.org/10.1007/978-3-030-90963-5_41
[253] Ralf Schmälzle and Shelby Wilcox. 2022. Harnessing Artificial Intelligence for Health Message Generation: The Folic Acid Message Engine. *J Med Internet Res* 24, 1, Article e28858 (Jan. 2022), 14 pages. https://doi.org/10.2196/28858~
[254] Damaris Schmid, Dario Staehelin, Andreas Bucher, Mateusz Dolata, and Gerhard Schwabe. 2022. Does Social Presence Increase Perceived Competence? Evaluating Conversational Agents in Advice Giving Through a Video-Based Survey. *Proc. ACM Hum.-Comput. Interact.* 6, GROUP, Article 26 (January 2022), 22 pages. https://doi.org/10.1145/3492845~~
[255] Bruno Schneider, Dominik Jäckle, Florian Stoffel, Alexandra Diehl, Johannes Fuchs, and Daniel Keim. 2018. Integrating data and model space in ensemble learning by visual analytics. *IEEE Transactions on Big Data* 7, 3 (Oct. 2018), 483-496. https://doi.org/10.1109/TBDATA.2018.2877350
[256] Anna Marie Schröder and Maliheh Ghajargar. 2021. Unboxing the Algorithm: Designing an Understandable Algorithmic Experience in Music Recommender Systems. In *Proceedings of the Perspectives on the Evaluation of Recommender Systems Workshop 2021. co-located with the 15th ACM Conference on Recommender Systems (RecSys 2021)*. Amsterdam, The Netherlands, Article 12, 1-11.

[257] Sophia Schulze-Weddige and Thorsten Zylowski. 2021. User Study on the Effects Explainable AI Visualizations on Non-experts. In *International Conference on ArtsIT, Interactivity and Game Creation (ArtsIT 2021)*. Springer, Cham, 457-467. https://doi.org/10.1007/978-3-030-95531-1_31
[258] Candice Schumann, Zhi Lang, Nicholas Mattei, and John P. Dickerson. 2022. Group Fairness in Bandits with Biased Feedback. In *Proceedings of the 21st International Conference on Autonomous Agents and Multiagent Systems (AAMAS '22)*. International Foundation for Autonomous Agents and Multiagent Systems, Richland, SC, 1155–1163.
[259] Hugo Scurto, Baptiste Caramiaux, and Frederic Bevilacqua. 2021. Prototyping Machine Learning Through Diffractive Art Practice. In *Designing Interactive Systems Conference 2021 (DIS '21)*. Association for Computing Machinery, New York, NY, USA, 2013–2025. https://doi.org/10.1145/3461778.3462163~~
[260] Hugo Scurto, Bavo Van Kerrebroeck, Baptiste Caramiaux, and Frédéric Bevilacqua. 2021. Designing Deep Reinforcement Learning for Human Parameter Exploration. *ACM Trans. Comput.-Hum. Interact.* 28, 1, Article 1 (February 2021), 35 pages. https://doi.org/10.1145/3414472~~
[261] Jessica Zeitz Self, Michelle Dowling, John Wenskovitch, Ian Crandell, Ming Wang, Leanna House, Scotland Leman, and Chris North. 2018. Observation-Level and Parametric Interaction for High-Dimensional Data Analysis. *ACM Trans. Interact. Intell. Syst.* 8, 2, Article 15 (June 2018), 36 pages. https://doi.org/10.1145/3158230~~
[262] Hong Shen, Haojian Jin, Ángel Alexander Cabrera, Adam Perer, Haiyi Zhu, and Jason I. Hong. 2020. Designing Alternative Representations of Confusion Matrices to Support Non-Expert Public Understanding of Algorithm Performance. *Proc. ACM Hum.-Comput. Interact.* 4, CSCW2, Article 153 (October 2020), 22 pages. https://doi.org/10.1145/3415224~~
[263] Murtuza N. Shergadwala and Magy Seif El-Nasr. 2021. Human-centric design requirements and challenges for enabling human-AI Interaction in engineering design: an interview study. In *Proceedings of the ASME 2021 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference (IDETC-CIE 2021)*. American Society of Mechanical Engineers, New York, NY, USA, Article V006T06A054. https://doi.org/10.1115/DETC2021-69809
[264] Donghee Shin. 2021. The effects of explainability and causability on perception, trust, and acceptance: Implications for explainable AI. *International Journal of Human-Computer Studies* 146, Article 102551 (Feb. 2021), 10 pages. https://doi-org/10.1016/j.ijhcs.2020.102551
[265] Donghee Shin, Kerk F. Kee, and Emily Y. Shin. 2022. Algorithm awareness: Why user awareness is critical for personal privacy in the adoption of algorithmic platforms? *Int. J. Inf. Manag.* 65, C (Aug 2022). https://doi.org/10.1016/j.ijinfomgt.2022.102494
[266] Ben Shneiderman. 2020. Bridging the Gap Between Ethics and Practice: Guidelines for Reliable, Safe, and Trustworthy Human-centered AI Systems. *ACM Trans. Interact. Intell. Syst.* 10, 4, Article 26 (December 2020), 31 pages. https://doi.org/10.1145/3419764~~
[267] Ben Shneiderman. 2020. Human-centered artificial intelligence: Three fresh ideas. *AIS Transactions on Human-Computer Interaction* 12, 3 (Sept. 2020), 109-124. https://doi.org/10.17705/1thci.00131~
[268] Ben Shneiderman. 2020. Design lessons from AI's two grand goals: Human emulation and useful applications. *IEEE Transactions on Technology and Society* 1, 2 (June 2020), 73-82. https://doi.org/10.1109/TTS.2020.2992669~
[269] Ben Shneiderman. 2020. Human-centered artificial intelligence: Reliable, safe & trustworthy. *International Journal of Human–Computer Interaction* 36, 6 (March 2020), 495-504. https://doi-org/10.1080/10447318.2020.1741118
[270] Akshat Shrivastava and Jeffrey Heer. 2020. ISeqL: interactive sequence learning. In *Proceedings of the 25th International Conference on Intelligent User Interfaces (IUI '20)*. Association for Computing Machinery, New York, NY, USA, 43–54. https://doi.org/10.1145/3377325.3377503~~
[271] Jesper Simonson and Toni Robertson, eds. 2013. *Routledge international handbook of participatory design*. Routledge, New York, NY.
[272] Jim Smith, Phil Legg, Milos Matovic, and Kristofer Kinsey. 2018. Predicting User Confidence During Visual Decision Making. *ACM Trans. Interact. Intell. Syst.* 8, 2, Article 10 (June 2018), 30 pages. https://doi.org/10.1145/3185524~~
[273] Chaehan So. 2020. Human-in-the-Loop Design Cycles – A Process Framework that Integrates Design Sprints, Agile Processes, and Machine Learning with Humans. In *Artificial Intelligence in HCI: First International Conference, AI-HCI 2020, Held as Part of the 22nd HCI International Conference, HCII 2020,* Copenhagen, Denmark, July 19–24, 2020, Proceedings. Springer-Verlag, Berlin, Heidelberg, 136–145. https://doi.org/10.1007/978-3-030-50334-5_9~
[274] Kacper Sokol and Peter Flach. 2020. One explanation does not fit all. *KI-Künstliche Intelligenz* 34, 2 (Feb. 2020), 235-250. https://doi-org/10.1007/s13218-020-00637-y~
[275] Serge Sonfack Sounchio, Laurent Geneste, and Bernard Kamsu Foguem. 2022. Combining expert-based beliefs and answer sets. *Applied Intelligence* 53, 3 (Feb 2023), 2694–2705. https://doi.org/10.1007/s10489-022-03669-z
[276] Fabian Sperrle, Astrik Jeitler, Jürgen Bernard, Daniel Keim, and Mennatallah El-Assady. 2021. Co-adaptive visual data analysis and guidance processes. *Comput. Graph.* 100, C (Nov 2021), 93–105. https://doi.org/10.1016/j.cag.2021.06.016
[277] Logan Stapleton, Min Hun Lee, Diana Qing, Marya Wright, Alexandra Chouldechova, Ken Holstein, Zhiwei Steven Wu, and Haiyi Zhu. 2022. Imagining new futures beyond predictive systems in child welfare: A qualitative study with impacted stakeholders. In *2022 ACM Conference on Fairness, Accountability, and Transparency (FAccT '22)*. Association for Computing Machinery, New York, NY, USA, 1162–1177. https://doi.org/10.1145/3531146.3533177~~
[278] Hariharan Subramonyam, Colleen Seifert, and Eytan Adar. 2021. ProtoAI: Model-Informed Prototyping for AI-Powered Interfaces. In *26th International Conference on Intelligent User Interfaces (IUI '21)*. Association for Computing Machinery, New York, NY, USA, 48–58. https://doi.org/10.1145/3397481.3450640~~
[279] Hariharan Subramonyam, Colleen Seifert, and Eytan Adar. 2021. Towards A Process Model for Co-Creating AI Experiences. In *Designing Interactive Systems Conference 2021 (DIS '21)*. Association for Computing Machinery, New York, NY, USA, 1529–1543. https://doi.org/10.1145/3461778.3462012~~
[280] Jakob Suchan, Mehul Bhatt, and Srikrishna Varadarajan. 2021. Commonsense visual sensemaking for autonomous driving–On generalised neurosymbolic online abduction integrating vision and semantics. *Artificial Intelligence* 299, Article 103522 (Oct. 2021), 33 pages. https://doi-org/10.1016/j.artint.2021.103522
[281] Lucy A. Suchman. 1987. *Plans and situated actions: the problem of human-machine communication*. Cambridge University Press, USA.
[282] Jiao Sun, Q. Vera Liao, Michael Muller, Mayank Agarwal, Stephanie Houde, Kartik Talamadupula, and Justin D. Weisz. 2022. Investigating Explainability of Generative AI for Code through Scenario-based Design. In *27th International Conference on Intelligent User Interfaces (IUI '22)*. Association for Computing Machinery, New York, NY, USA, 212–228. https://doi.org/10.1145/3490099.3511119~~
[283] Harini Suresh, Steven R. Gomez, Kevin K. Nam, and Arvind Satyanarayan. 2021. Beyond Expertise and Roles: A Framework to Characterize the Stakeholders of Interpretable Machine Learning and their Needs. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems (CHI '21)*. Association for Computing Machinery, New York, NY, USA, Article 74, 1–16. https://doi.org/10.1145/3411764.3445088~~
[284] Gary K. L. Tam, Vivek Kothari, and Min Chen. 2017. An Analysis of Machine- and Human-Analytics in Classification. *IEEE Transactions on Visualization and Computer Graphics* 23, 1 (January 2017), 71–80. https://doi.org/10.1109/TVCG.2016.2598829
[285] Anissa Tanweer, Cecilia R Aragon, Michael Muller, Shion Guha, Samir Passi, Gina Neff, and Marina Kogan. 2022. Interrogating Human-centered Data Science: Taking Stock of Opportunities and Limitations. In *Extended Abstracts of the 2022 CHI Conference on Human Factors in Computing Systems (CHI EA '22)*. Association for Computing Machinery, New York, NY, USA, Article 99, 1–6. https://doi.org/10.1145/3491101.3503740
[286] Andreas Theissler, Mark Thomas, Michael Burch, and Felix Gerschner. 2022. ConfusionVis: Comparative evaluation and selection of multi-class classifiers based on confusion matrices. *Know.-Based Syst.* 247, C (Jul 2022). https://doi.org/10.1016/j.knosys.2022.108651
[287] Andreas Theissler, Simon Vollert, Patrick Benz, Laurentius A. Meerhoff, and Marc Fernandes. 2020. ML-ModelExplorer: An explorative model-agnostic approach to evaluate and compare multi-class classifiers. In *International Cross-Domain Conference for Machine Learning and Knowledge Extraction (CD-MAKE 2020)*. Springer, Cham, 281-300. https://doi.org/10.1007/978-3-030-57321-8_16
[288] Jana Thompson. 2021. Mental Models and Interpretability in AI Fairness Tools and Code Environments. In *HCI International 2021 - Late Breaking Papers: Multimodality, eXtended Reality, and Artificial Intelligence: 23rd HCI International Conference, HCII 2021,* Virtual Event, July 24–29, 2021, Proceedings. Springer-Verlag, Berlin, Heidelberg, 574–585. https://doi.org/10.1007/978-3-030-90963-5_43
[289] Filareti Tsalakanidou, Symeon Papadopoulos, Vasileios Mezaris, Ioannis Kompatsiaris, Birgit Gray, Danae Tsabouraki, Maritini Kalogerini Fulvio Negro, Maurizio Montagnuolo, Jesse de Vos, Philo van Kemenade, Daniele Gravina, Rémi Mignot, Alexey Ozerov, Francois Schnitzler, Artur Garcia-Saez, Georgios N. Yannakakis, Antonios Liapis and Georgi Kostadinov. 2021. The AI4Media Project: Use of Next-Generation Artificial Intelligence Technologies for Media Sector Applications. In *IFIP International Conference on Artificial Intelligence Applications and Innovations (AIAI 2021)*. Springer, Cham, 81-93. https://doi-org/10.1007/978-3-030-79150-6_7~
[290] Kristen Vaccaro, Karrie Karahalios, Deirdre K. Mulligan, Daniel Kluttz, and Tad Hirsch. 2019. Contestability in Algorithmic Systems. In *Conference Companion Publication of the 2019 on Computer Supported Cooperative Work and Social Computing (CSCW '19)*. Association for Computing Machinery, New York, NY, USA, 523–527. https://doi.org/10.1145/3311957.3359435
[291] Niels van Berkel, Omer F. Ahmad, Danail Stoyanov, Laurence Lovat, and Ann Blandford. 2021. Designing Visual Markers for Continuous Artificial Intelligence Support: A Colonoscopy Case Study. *ACM Trans. Comput. Healthcare* 2, 1, Article 7 (January 2021), 24 pages. https://doi.org/10.1145/3422156~~
[292] Niels van Berkel, Benjamin Tag, Jorge Goncalves, and Simo Hosio. 2022. Human-centred artificial intelligence: a contextual morality perspective. *Behaviour & Information Technology* 41, 3 (Sept. 2020), 502-518. https://doi.org/10.1080/0144929X.2020.1818828~

[293] Craig Vear. 2021. Creative AI and Musicking Robots. *Front. Robot. AI* 8, Article 631752 (Nov. 2021), 10 pages. https://doi.org/10.3389/frobt.2021.631752
[294] Erik Veitch and Ole Andreas Alsos. 2021. Human-Centered Explainable Artificial Intelligence for Marine Autonomous Surface Vehicles. *J. Mar. Sci. Eng.* 9, 11, Article 1227 (Nov. 2021), 24 pages. https://doi.org/10.3390/jmse9111227
[295] Q. Vera Liao and S. Shyam Sundar. 2022. Designing for Responsible Trust in AI Systems: A Communication Perspective. In *2022 ACM Conference on Fairness, Accountability, and Transparency (FAccT '22)*. Association for Computing Machinery, New York, NY, USA, 1257–1268. https://doi.org/10.1145/3531146.3533182~~
[296] Himanshu Verma, Guillaume Pythoud, Grace Eden, Denis Lalanne, and Florian Evéquoz. 2019. Pedestrians and Visual Signs of Intent: Towards Expressive Autonomous Passenger Shuttles. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 3, 3, Article 107 (September 2019), 31 pages. https://doi.org/10.1145/3351265~~
[297] Sruthi Viswanathan, Cecile Boulard, Adrien Bruyat, and Antonietta Maria Grasso. 2022. Situational Recommender: Are You On the Spot, Refining Plans, or Just Bored? In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems (CHI '22)*. Association for Computing Machinery, New York, NY, USA, Article 473, 1–19. https://doi.org/10.1145/3491102.3501909~~
[298] Dieter P. Wallach, Lukas A. Flohr, and Annika Kaltenhauser. 2020. Beyond the Buzzwords: On the Perspective of AI in UX and Vice Versa. In *Artificial Intelligence in HCI: First International Conference, AI-HCI 2020, Held as Part of the 22nd HCI International Conference, HCII 2020*, Copenhagen, Denmark, July 19–24, 2020, Proceedings. Springer-Verlag, Berlin, Heidelberg, 146–166. https://doi.org/10.1007/978-3-030-50334-5_10
[299] April Yi Wang, Dakuo Wang, Jaimie Drozdal, Michael Muller, Soya Park, Justin D. Weisz, Xuye Liu, Lingfei Wu, and Casey Dugan. 2022. Documentation Matters: Human-Centered AI System to Assist Data Science Code Documentation in Computational Notebooks. *ACM Trans. Comput.-Hum. Interact.* 29, 2, Article 17 (April 2022), 33 pages. https://doi.org/10.1145/3489465~~
[300] Dakuo Wang, Josh Andres, Justin D. Weisz, Erick Oduor, and Casey Dugan. 2021. AutoDS: Towards Human-Centered Automation of Data Science. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems (CHI '21)*. Association for Computing Machinery, New York, NY, USA, Article 79, 1–12. https://doi.org/10.1145/3411764.3445526~~
[301] Dakuo Wang, Justin D. Weisz, Michael Muller, Parikshit Ram, Werner Geyer, Casey Dugan, Yla Tausczik, Horst Samulowitz, and Alexander Gray. 2019. Human-AI Collaboration in Data Science: Exploring Data Scientists' Perceptions of Automated AI. *Proc. ACM Hum.-Comput. Interact.* 3, CSCW, Article 211 (November 2019), 24 pages. https://doi.org/10.1145/3359313
[302] Liuping Wang, Dakuo Wang, Feng Tian, Zhenhui Peng, Xiangmin Fan, Zhan Zhang, Mo Yu, Xiaojuan Ma, and Hongan Wang. 2021. CASS: Towards Building a Social-Support Chatbot for Online Health Community. *Proc. ACM Hum.-Comput. Interact.* 5, CSCW1, Article 9 (April 2021), 31 pages. https://doi.org/10.1145/3449083~~
[303] Qiaosi Wang, Ida Camacho, Shan Jing, and Ashok K. Goel. 2022. Understanding the Design Space of AI-Mediated Social Interaction in Online Learning: Challenges and Opportunities. *Proc. ACM Hum.-Comput. Interact.* 6, CSCW1, Article 130 (April 2022), 26 pages. https://doi.org/10.1145/3512977~~
[304] Qiaosi Wang, Koustuv Saha, Eric Gregori, David Joyner, and Ashok Goel. 2021. Towards Mutual Theory of Mind in Human-AI Interaction: How Language Reflects What Students Perceive About a Virtual Teaching Assistant. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems (CHI '21)*. Association for Computing Machinery, New York, NY, USA, Article 384, 1–14. https://doi.org/10.1145/3411764.3445645~~
[305] Pontus Wärnestål. 2022. Multi-disciplinary Learning and Innovation for Professional Design of AI-Powered Services. In *International Conference on Design, Learning, and Innovation (DLI 2021)*. Springer, Cham, 21-36. https://doi.org/10.1007/978-3-031-06675-7_2
[306] Anna Christina Weigand and Martin Christof Kindsmüller. 2021. HCD3A: An HCD Model to Design Data-Driven Apps. In *Artificial Intelligence in HCI: Second International Conference, AI-HCI 2021, Held as Part of the 23rd HCI International Conference, HCII 2021*, Virtual Event, July 24–29, 2021, Proceedings. Springer-Verlag, Berlin, Heidelberg, 285–297. https://doi.org/10.1007/978-3-030-77772-2_19
[307] Justin D. Weisz, Mary Lou Maher, Hendrik Strobelt, Lydia B. Chilton, David Bau, and Werner Geyer. 2022. HAI-GEN 2022: 3rd Workshop on Human-AI Co-Creation with Generative Models. In *27th International Conference on Intelligent User Interfaces (IUI '22 Companion)*. Association for Computing Machinery, New York, NY, USA, 4–6. https://doi.org/10.1145/3490100.3511166
[308] Justin D. Weisz, Michael Muller, Steven I. Ross, Fernando Martinez, Stephanie Houde, Mayank Agarwal, Kartik Talamadupula, and John T. Richards. 2022. Better Together? An Evaluation of AI-Supported Code Translation. In *27th International Conference on Intelligent User Interfaces (IUI '22)*. Association for Computing Machinery, New York, NY, USA, 369–391. https://doi.org/10.1145/3490099.3511157~~
[309] John Wenskovitch, Michelle Dowling, and Chris North. 2020. With respect to what? simultaneous interaction with dimension reduction and clustering projections. In *Proceedings of the 25th International Conference on Intelligent User Interfaces (IUI '20)*. Association for Computing Machinery, New York, NY, USA, 177–188. https://doi.org/10.1145/3377325.3377516~~
[310] Maximiliane Windl, Sebastian S. Feger, Lara Zijlstra, Albrecht Schmidt, and Pawel W. Wozniak. 2022. 'It Is Not Always Discovery Time': Four Pragmatic Approaches in Designing AI Systems. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems (CHI '22)*. Association for Computing Machinery, New York, NY, USA, Article 50, 1–12. https://doi.org/10.1145/3491102.3501943
[311] Marcus Winter and Phil Jackson. 2020. Flatpack ML: How to Support Designers in Creating a New Generation of Customizable Machine Learning Applications. In *Design, User Experience, and Usability. Design for Contemporary Interactive Environments: 9th International Conference, DUXU 2020, Held as Part of the 22nd HCI International Conference, HCII 2020*, Copenhagen, Denmark, July 19–24, 2020, Proceedings, Part II. Springer-Verlag, Berlin, Heidelberg, 175–193. https://doi.org/10.1007/978-3-030-49760-6_12
[312] Doris Xin, Eva Yiwei Wu, Doris Jung-Lin Lee, Niloufar Salehi, and Aditya Parameswaran. 2021. Whither AutoML? Understanding the Role of Automation in Machine Learning Workflows. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems (CHI '21)*. Association for Computing Machinery, New York, NY, USA, Article 83, 1–16. https://doi.org/10.1145/3411764.3445306~~
[313] Wei Xu. 2019. Toward human-centered AI: a perspective from human-computer interaction. *interactions* 26, 4 (July-August 2019), 42–46. https://doi.org/10.1145/3328485~~
[314] Qian Yang, Alex Scuito, John Zimmerman, Jodi Forlizzi, and Aaron Steinfeld. 2018. Investigating How Experienced UX Designers Effectively Work with Machine Learning. In *Proceedings of the 2018 Designing Interactive Systems Conference (DIS '18)*. Association for Computing Machinery, New York, NY, USA, 585–596. https://doi.org/10.1145/3196709.3196730~~
[315] Qian Yang, Aaron Steinfeld, Carolyn Rosé, and John Zimmerman. 2020. Re-examining Whether, Why, and How Human-AI Interaction Is Uniquely Difficult to Design. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems (CHI '20)*. Association for Computing Machinery, New York, NY, USA, 1–13. https://doi.org/10.1145/3313831.3376301~~
[316] Qian Yang, Jina Suh, Nan-Chen Chen, and Gonzalo Ramos. 2018. Grounding Interactive Machine Learning Tool Design in How Non-Experts Actually Build Models. In *Proceedings of the 2018 Designing Interactive Systems Conference (DIS '18)*. Association for Computing Machinery, New York, NY, USA, 573–584. https://doi.org/10.1145/3196709.3196729~~
[317] Xinfen Yang, Hao Zhang, Renmei Chen, Shasha Li, Ni Zhang, Biao Wang, Xin Wang, and Deepak Kumar Jain. 2022. Research on Forecasting of Student Grade Based on Adaptive K-Means and Deep Neural Network. *Wirel. Commun. Mob. Comput.* 2022 (2022). https://doi.org/10.1155/2022/5454158
[318] Alexey Zagalsky, Dov Te'eni, Inbal Yahav, David G. Schwartz, Gahl Silverman, Daniel Cohen, Yossi Mann, and Dafna Lewinsky. 2021. The Design of Reciprocal Learning Between Human and Artificial Intelligence. *Proc. ACM Hum.-Comput. Interact.* 5, CSCW2, Article 443 (October 2021), 36 pages. https://doi.org/10.1145/3479587~~
[319] Michael Zbyszyński, Balandino Di Donato, Federico Ghelli Visi, and Atau Tanaka. 2019. Gesture-Timbre Space: Multidimensional Feature Mapping Using Machine Learning and Concatenative Synthesis. In *Perception, Representations, Image, Sound, Music: 14th International Symposium, CMMR 2019*, Marseille, France, October 14–18, 2019, Revised Selected Papers. Springer-Verlag, Berlin, Heidelberg, 600–622. https://doi.org/10.1007/978-3-030-70210-6_39
[320] Shanghang Zhang, Hao Dong, Wei Pan, Pradeep Ravikumar, Vittorio Ferrari, Fisher Yu, Xin Wang, and Zihan Ding. 2022. Workshop on Human in the Loop Learning. Retrieved from https://neurips-hill.github.io/.
[321] Yingwei Zhang, Yiqiang Chen, Weiwen Yang, Hanchao Yu, and Zeping Lv. 2022. Human-centered intelligent healthcare: explore how to apply AI to assess cognitive health. *CCF Trans. Pervasive Comp. Interact.* 4 (May 2022), 189-206. https://doi-org/10.1007/s42486-022-00102-9
[322] Chengbo Zheng, Dakuo Wang, April Yi Wang, and Xiaojuan Ma. 2022. Telling Stories from Computational Notebooks: AI-Assisted Presentation Slides Creation for Presenting Data Science Work. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems (CHI '22)*. Association for Computing Machinery, New York, NY, USA, Article 53, 1–20. https://doi.org/10.1145/3491102.3517615~~
[323] Qi Zhou, Wannapon Suraworachet, Stanislav Pozdniakov, Roberto Martinez-Maldonado, Tom Bartindale, Peter Chen, Dan Richardson, and Mutlu Cukurova. 2021. Investigating Students' Experiences with Collaboration Analytics for Remote Group Meetings. In *Artificial Intelligence in Education: 22nd International Conference, AIED 2021*, Utrecht, The Netherlands, June 14–18, 2021, Proceedings, Part I. Springer-Verlag, Berlin, Heidelberg, 472–485. https://doi.org/10.1007/978-3-030-78292-4_38
[324] Zhibin Zhou, Lingyun Sun, Yuyang Zhang, Xuanhui Liu and Qing Gong. 2020. ML lifecycle canvas: designing machine learning-empowered UX with material lifecycle thinking. *Human–Computer Interaction* 35, 5-6 (April 2020), 362-386.

https://doi-org/10.1080/07370024.2020.1736075~
[325] Tehseen Zia, Nauman Bashir, Mirza Ahsan Ullah, and Shakeeb Murtaza. 2022. SoFTNet: A concept-controlled deep learning architecture for interpretable image classification. *Know.-Based Syst.* 240, C (Mar 2022). https://doi.org/10.1016/j.knosys.2021.108066
[326] Alexandra Zytek, Dongyu Liu, Rhema Vaithianathan, and Kalyan Veeramachaneni. 2022. Sibyl: Understanding and Addressing the Usability Challenges of Machine Learning In High-Stakes Decision Making. *IEEE Transactions on Visualization and Computer Graphics* 28, 1 (Jan. 2022), 1161–1171. https://doi.org/10.1109/TVCG.2021.3114864