

Using human-in-the-loop and explainable AI to envisage new future work practices

KONSTANTINOS TSIAKAS and DAVE MURRAY-RUST, Industrial Design Engineering, TU Delft, Netherlands

In this paper, we discuss the trends and challenges of the integration of Artificial Intelligence (AI) methods in the workplace. An important aspect towards creating positive AI futures in the workplace is the design of fair, reliable and trustworthy AI systems which aim to augment human performance and perception, instead of replacing them by acting in an automatic and non-transparent way. Research in Human-AI Interaction has proposed frameworks and guidelines to design transparent and trustworthy human-AI interactions. Considering such frameworks, we discuss the potential benefits of applying human-in-the-loop (HITL) and explainable AI (xAI) methods to define a new design space for the future of work. We illustrate how such methods can create new interactions and dynamics human users and AI in future work practices.

Additional Key Words and Phrases: Human-AI interaction, Future of Work, Explainable AI, Human-in-the-Loop

ACM Reference Format:

Konstantinos Tsiakas and Dave Murray-Rust. 2022. Using human-in-the-loop and explainable AI to envisage new future work practices. In *PETRA '22: ACM Symposium on Neural Gaze Detection, June 03–05, 2018, Woodstock, NY*. ACM, New York, NY, USA, 10 pages. <https://doi.org/XXXXXXXX.XXXXXXX>

1 INTRODUCTION

The Future of Work (FoW) is being shaped by the growing adoption of Artificial Intelligence (AI) in the workplace. AI-based systems, methods and approaches have been deployed in several workplace contexts, including but not limited to, automation and industrial settings, human resources management, as well as remote work. The main goal of AI is to increase efficiency and productivity in the workplace. Despite the possible benefits of the digital transformation of the workplace and the transition to the FoW, there are many concerns related to ethical considerations, safety, and trust.

There is a growing interest in designing, developing, and evaluating methods to ensure that human users can safely interact with an AI system which is transparent and accountable, and makes fair decisions with respect to ethical considerations. As an example, Explainable AI (XAI) methods have been proposed to enhance trust in human-AI interaction. Moreover, fairness has been introduced as an evaluation metric for AI models, in order to mitigate bias, either due to pre-existing bias that is captured in the data, or due to technical bias introduced during data processing and modeling. Model cards and reports have been proposed to ensure transparency and intelligibility for developed AI models. Also, Human-in-the-Loop (HITL) methods aim to engage users during the interaction by enabling them to provide feedback to the system which can be used either as an evaluation metric for the system's performance, or as an additional feedback which can be utilized by the learning algorithm to facilitate learning. Most of these methods aim to "correct" possible problems that may arise from the integration of AI in real-world applications with human users, e.g., fairness evaluation of an (explainable) AI model before deployment with human users.

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 ACM 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.

© 2022 Association for Computing Machinery.

Manuscript submitted to ACM

53 In this paper, we discuss the possible benefits of defining a new design space for the future of work. From a post-
54 industrial design thinking perspective, our goal is to define design strategies for systems that may have fuzzy boundaries
55 and complex effects. More specifically, data-driven AI systems may use data from several sources and their design
56 may include a network of stakeholders and users. Moreover, socio-legal situations are likely to adapt to the impacts of
57 technology, so that machine operations drive human behaviour and vice versa.

58 Moreover, the paper highlights the aspects of *co-performance* and *response-ability* [8, 11]. Co-performance refers to
59 the synergistic interaction between a human and a machine as an opportunity for the different abilities of the two parts
60 to be combined in. Moreover, when the responsibility for a decision or an operation is diffused through a system rather
61 than centered on a person, how can we create a sensitivity to the needs and rights of the network of people involved.

62 From a human-centered design aspect, we focus on two emerging trends in human-centered AI (XAI and HITL),
63 arguing that the combination of these approaches can create new design possibilities in the future of work. We are
64 interested in the beneficial, ecosystemic possibilities given by the introduction of XAI and HITL methods in the particular
65 situation of future work practices. This is an area where use of AI systems becomes part of everyday life, mediating
66 between people and organisations with differing powers and agencies, but of deep importance to the quality of life of
67 many people and the performance of many organisations. In order to do this, we engage with the existing trends within
68 AI and the future of work, as well as guidelines for how human-AI interactions should be designed. From this, we look
69 at the new possibilities offered by HITL and XAI, and illustrate with short examples the kinds of new interactions and
70 configurations that are possible.

71 2 BACKGROUND AND RELATED WORK

72 2.1 AI and the Future of Work: Trends and Challenges

73 There is an increasing need to investigate the potential benefits and challenges of the digital transformation and
74 the integration of AI systems in the workplace. Applications of AI in the workplace may pose several implications
75 related to the AI capabilities, i.e., learning and predicting human behavior, as well as to establish new values, ethical
76 considerations and morals. A research study highlights the possible concerns of AI in the workplace, taking into
77 consideration the emerging challenges raised by COVID-19, as well as insights for the future of work [6]. The paper
78 suggests that explainability and accountability, as well as digital literacy among all stakeholders and users (decision
79 makers, employers, employees), is an important feature of a policy agenda and implementation strategy for AI in
80 the workplace. One example of an application of AI that may pose ethical, technical and societal implications in the
81 workplace is the management of organization employees, i.e., AI systems for organizations/companies to manage
82 their employees, including monitoring, coordinating, and making decisions (recruitment, promotion, etc.) about their
83 employees. A research study focuses on the applicant's perspective, conducting a user study with undergraduate
84 students, as the main future users of such AI-based recruitment systems [9]. Based on the results from the thematic
85 analysis, five dominant themes were identified: efficiency, impartiality, conformity, human interaction, and uncertainty.
86 Based on their analysis, the authors proposed a framework to integrate AI methods for recruiting, focusing on the
87 potential benefits of AI on different stages of the recruitment phase. Moreover, their results indicated that future users
88 of such systems are still hesitant towards a complete digitization of the recruitment process, since AI decisions may be
89 biased, unpredictable and invisible to the applicant who is being impacted by these decisions. AI prediction models
90 have been proposed to estimate the level of personnel competence in order to optimize job performance [4]. More
91 specifically, the proposed model predicts job performance, as a function of job competence, based on job knowledge,
92
93
94
95
96
97
98
99
100
101
102
103
104

105 self-motivation, self-concern, and role perception. The model is used to predict the competence of the applicants, and
106 ultimately make decisions (accept or reject application).

107 In order to address the issues of unfairness, privacy, and bias in such AI systems, a design agenda for AI systems for
108 employee management in organizations has been proposed [14]. In this paper, the authors present and discuss three
109 different types of fairness in the context of the organization (distributive, procedural, interactional), as well as ways
110 for justice to redress unfairness. Moreover, based on a literature review on AI design and fairness in organizations,
111 they propose a design framework as a set of primary components that need to be considered, including the aspect of
112 user affordances, i.e., *"the particular ways in which an individual perceives and interacts with a system"*. Considering
113 user affordances, a fair AI system should be transparent about its internal models and mechanics, explainable to
114 effectively communicate its models and decisions to human users, able to provide visualizations to represent the required
115 information. Finally, such systems should be able to satisfy the affordance of voice; to give users the opportunity to
116 provide feedback and communicate with the AI during the interaction. In the context of manufacturing, Artificial
117 Intelligence has been mainly applied towards the automation of human-like and repetitive tasks. Research works
118 focus on the interactions between human actors and automation, focusing on the capabilities of AI systems to support
119 and augment human performance. A research study investigates the impact of Human-in-the-Loop AI methods in a
120 manufacturing setup [7]. The paper illustrates how the integration of human-in-the-loop AI methods can create new
121 communication channels between human and non-human actors, highlighting the need to analyse and understand the
122 emergent outcomes of such synergistic interactions between human and AI actors, with both parts of the interaction
123 can augment and support each other. Focusing on the challenges that arise from the integration of such approaches
124 in manufacturing and automation, it is essential that organizations and companies provide sustainable training and
125 education for their workforce [10]

126 In the context of human-AI collaboration, a research article investigated the use of *cobots* in managerial professions
127 [17] arguing that *"the future of AI in knowledge work needs to focus not on full automation but rather on collaborative
128 approaches where humans and AI work closely together"*. In order to support human-human collaboration, AI can be
129 applied for real-time analysis of meetings, brainstorming sessions and digital collaboration. A research study presented
130 Meeting Mediator; an AI-supported real-time self-reflection tool for online meetings [13], which visualizes estimations
131 of key metrics, including group and individual performance, speaking time and influences between the participants.
132 They conclude that AI-based digital tools can trigger behavioral change and can offer immediate feedback to participants
133 which can help build and develop soft skills required to succeed in a new digital environment, e.g., increase perception
134 of dominance during virtual meetings.

144 2.2 Design Guidelines and Frameworks for Human-AI Interaction

145 The motivation of designing fair and transparent human-AI interactions is to involve users to the decision making,
146 learning, and adaptation process of an AI system for the following reasons. Such interactions can (a) augment the
147 users' perception about themselves, the system mechanics, and their (common) environment, (b) can enhance trust,
148 fairness, and reliability, and (c) can enable users to learn how to efficiently collaborate with AI systems towards hybrid
149 intelligence. A recent research article [19] defines Human-AI Interaction as "the completion of a user's task with the
150 help of AI support, which may manifest itself in non-intermittent scenarios". The authors present three main types of
151 Human-AI interaction: intermittent, continuous, and proactive, highlighting *"how differences in initiation and control
152 result in diverging user needs"*.

157 These three paradigms of human-AI interaction can exist in parallel. However, there is a need to design interaction
158 paradigms, focusing especially on the challenges of continuous and proactive interactions and support designers in
159 creating usable AI-driven systems. The main goal of integrating AI and ML methods to HCI systems is to improve the
160 interactions between the user and the system (trust, fairness, accountability, performance, etc.). However, there is a lack
161 of design innovation in envisioning how ML might improve user experience [21]. In this review paper, the authors
162 identified a lack of research integrating UX and ML methods. Based on their analysis, they identified value channels
163 through which the technical capabilities can provide value for users: self, context, optimal, and utility-capability. The
164 authors provide a schema of machine learning capabilities in terms of inference and actuation in order to increase a
165 user's perception of the aforementioned experiential value. Following the argument that Machine Learning as a design
166 material adds value to user experience, designers should be able to identify how existing AI and ML approaches and
167 methods can be integrated to the design process, e.g., which is an appropriate ML algorithm for a given design, or how
168 to design an AI system to support given design values? Considering the different ways that human-AI interactions can
169 be designed, a research study identified a set of design challenges for human-AI interactions [22]. More specifically,
170 they present five categories of challenges: (1) understanding AI capabilities, (2) envisioning novel and implementable AI
171 for a given UX problem, (3) iterative prototyping and testing human-AI interaction, (4) crafting thoughtful interactions,
172 and (5) collaborating with AI engineers throughout the design process. They present four levels of AI systems based on
173 the design complexity. In order to address different types of challenges, the authors present their suggestions towards
174 facilitating human-AI interaction design: "improving designers' technical literacy, facilitating design-oriented data
175 exploration, enabling designers to more easily "play with" AI in support of design ideation, to gain a felt sense of what
176 AI can do, aiding designers in evaluating AI outputs, and creating AI-specific design processes". Microsoft Research has
177 proposed 18 applicable design guidelines for human-AI interaction [2]. We identify a set of 8 categories which can relate
178 to system's fairness, explainability and adaptability, as well as autonomy and shared control. For example, providing
179 appropriate and accountable information to users requires fair and transparent ML approaches, while enabling the
180 user to intervene to the process (e.g., ignore or guide AI), requires the system design to enable the user to provide
181 granular feedback, to learn from user's input and behaviour. Focusing on the aspects of user-centric explainability, a
182 framework is proposed towards designing Theory-Driven User-Centric Explainable AI [20]. This paper discusses how
183 the proposed framework can bridge algorithm-generated explanations and human decision-making theories. Apart
184 from designing explanations that can be easily perceived by human users, a key module of this framework focuses
185 on how the application of XAI methods can be used to support reasoning and mitigate errors during the human-AI
186 interaction.

187 The Human-Centered Artificial Intelligence (HCAI) framework [15] describes how to (1) design for high levels of
188 human control and computer automation to increase human performance, (2) understand the situations in which full
189 human or computer control are necessary, and (3) avoid the dangers of excessive human or computer control. Reliability
190 requires appropriate technical practices, which support human responsibility, fairness, and explainability. Based on this
191 framework, the goal of Reliable, Safe and Trustworthy AI is achieved by a high level of human control and high level
192 of computer automation. The design decisions give human operators a clear understanding of the machine state and
193 their choices, guided by concerns such as the consequences and reversibility of mistakes. Well-designed automation
194 preserves human control where appropriate, thereby increasing performance and enabling creative improvements.

3 DESIGNING HUMAN-IN-THE-LOOP AND EXPLAINABLE AI INTERACTIONS

In this section, we focus on the design aspects of human-AI interaction, where human users and AI systems can communicate in a collaborative fashion through an exchange of feedback and explanations. We present a set of design aspects that need to be defined while designing XAI and HITL systems. Based on these, we briefly describe how to design aspects of AI systems, including types and roles of users, level of autonomy and human control, AI learning capabilities and personalization.

- (1) **Designing Explainable AI.** Explanations can be used for various reasons and purposes, e.g., to inform or persuade users, to help programmers debug complex models, to visualize the models to domain experts for knowledge extraction, etc. The purpose of the explanations is highly linked both to the sender and the the recipient of the explanations. According to the goal and the parts included in this interaction, an important aspect is to communicate the explanations to the user in an effective way. The effectiveness of the explanations depends on various factors, including: (a) form of explanations designed (text, visuals, speech, etc.,), (b) frequency or timing of explanations, (c) level of transparency and explainability, and (d) the type of explanations (e.g., contrastive, counterfactual, local vs. global explanations). Different users may need to have different access to the explanations or with different levels of transparency. Based on the system requirements and the purpose of explanations, designing an efficient explainable AI system requires a proper definition of when (or how often) explanations should be given, considering user's cognitive overload.
- (2) **Designing Human-in-the-Loop AI.** In order to design an HITL system, the first step is to define which user(s) can provide feedback to the system, as well as the motivation of including a human user in the AI loop. Human users can provide feedback to the system in different ways and for different purposes. For example, human users can provide evaluative feedback during the interaction (in the form of a numerical signal), which can be used as an evaluation metric. Moreover, user feedback can be used to intervene and control the system during the interaction, in the form of corrective action. Human feedback can be provided either implicitly or explicitly. Depending on the system requirements, human user must be able to provide the feedback to the system without burdening their capabilities during the interaction, e.g., mental workload. In order to ensure that providing feedback is not an additional effort to the user, it is important to answer the following questions: when does the user provide feedback?, does the system ask for feedback? is the user free to provide feedback at any point of the interaction?

Taking into consideration these aspects of XAI and HITL systems, we describe how they can be applied considering key design aspects of AI systems:

- (1) **AI System Design.** A definition of the AI input/output space is essential for a proper design of an XAI/HITL system, since they play a central role in the formulation of explanations. Moreover, model parameters and metrics (e.g., accuracy, uncertainty, etc.) can be used as design materials for the interaction. For example, if there is high model uncertainty for a given input, the system design can allow this information to be communicated to the user through the XAI communication channel. Additionally, the user can provide feedback back to the system through the HITL communication channel, which can be integrated to the learning mechanism of the algorithm. For example, hand gestures or facial expressions can be used as evaluative feedback to accept or reject the decision of an AI system.
- (2) **Types of Human-AI Interactions and Roles.** Apart from the categorization of an interaction as intermittent, continuous, or proactive [19], it is important to define the composition of the human-AI teams, as well as their

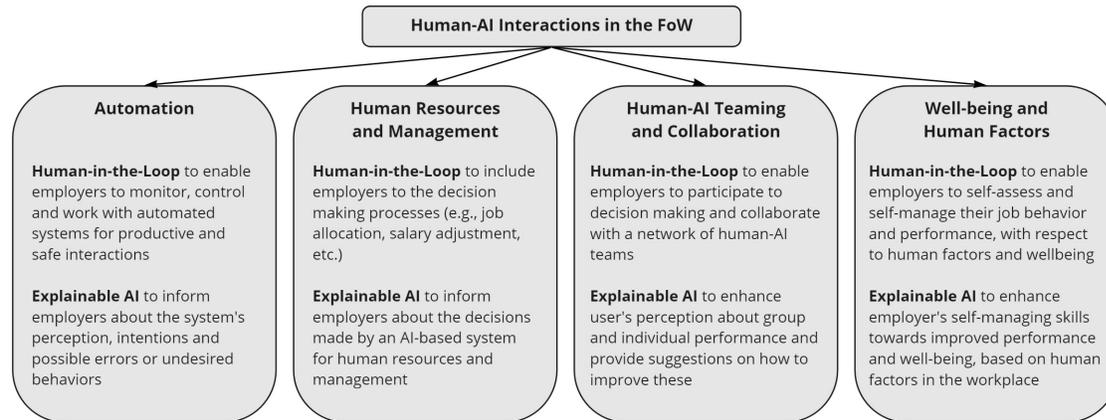
261 roles in the interaction, especially in a network of human and non-human (AI) actors. Considering the XAI
262 and HITL channels, the role of users can be characterized by the way they provide feedback to the system
263 or how the system provides explanations to the specific user. For example, in collaborative tasks, XAI can
264 be used to inform the members of the team about their group performance, but can also be used to provide
265 more personal information to each individual separately. In a similar manner, HITL methods can be used to
266 enable each team member evaluate their individual and group performance. Each human and AI role can create
267 different possibilities for interactions between users and an AI model and can change the way the model is used
268 [18].
269

270
271 **(3) Designing for Autonomy and Human Control** Integrating explanations and interactive HITL methods in the
272 interaction does not entirely address the issue of system autonomy and human control. An explainable system
273 can provide explanations to the user for an automated decision it made in order to ensure trust. However, such
274 interaction does not allow the user to have control over the decisions. When a human user provides feedback, the
275 system can use it as a command (human control), can negotiate it with the user for a shared-autonomy decision,
276 or it may not consider it for its decision but rather as an evaluative feedback at the end of the interaction. It is
277 important to define such aspects of autonomy for all possible types of users and interactions that may emerge.
278 Explanations are important for human control, since they can enhance user's perception and decision making
279 in order to make an informed decision which will be beneficial for the interaction. Moreover, HITL methods
280 can broaden the scope of responsibility. It can engage people to actively get involved in order to challenge the
281 decisions or even change the the behaviour of systems.
282

283
284 **(4) Designing for AI: Learning, Adaptation and Personalization** Learning and personalization are crucial
285 features of dynamic systems which enable them to adjust to environmental changes and unseen events, or
286 different and new users. While learning, adaptation, and personalization of AI models are technical challenges
287 of an AI-based system, it is important to define the AI capabilities of the system and how they can be realized
288 through design. For example, there are different types of learning based on the model's learning mechanism,
289 including online, offline, batch and active learning. Active learning approaches, i.e., when the learning algorithm
290 can query a user interactively for data annotation with the desired outputs, require the design of an interface
291 (button, gestures, speech, etc.) which will allow AI and user interact for the purpose of data annotation. If the
292 system makes learning updates during the interaction, a design feature could be used to let the user know
293 (e.g., a progress bar or inactive interface). In terms of personalization, it is important to consider (if and) how
294 the system personalizes its behavior to different users or contexts. In other words, Which are the control and
295 observed parameters should be observed and adjusted and based on which observed parameters?
296
297
298
299

300 301 302 303 304 **4 HUMAN-IN-THE-LOOP AND EXPLAINABLE AI IN THE FUTURE OF WORK**

305 In this section, we provide a summary on how XAI and HITL methods can be applied in the workplace context,
306 highlighting the potential benefits of such interactions in future work practices. Our goal is to identify the different
307 configurations of human-AI interactions that emerge when different types of human users interact with AI in a set of
308 different scenarios and application contexts.
309
310
311
312



328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364

Fig. 1. Human-in-the-Loop and Explainable AI in the Future of Work

Automation. In the domain of automation, AI plays an important role in achieving high performance (productivity) while ensuring safety and quality. While the main contribution of AI in manufacturing is automation, a central role of AI is to augment human's perception and performance while collaborating with AI systems. HITL methods have been proposed to enable collaboration between human workers and automated systems and investigate the potentials of human meta-learning capabilities in such sociotechnical systems, considering the possible physical, cognitive and mental demands for workers and how these can affect the overall job performance [7]. HITL AI agents enable human users to provide feedback to the system, guide it, or even take control of the operation if needed (e.g., AI errors and malfunctions). Achieving a high level of human agency and control while interacting with an system with high automation level is considered to ensure a safe, reliable and trustworthy interaction with human-centered AI systems [15]. In order to enhance user's perception about the system's capabilities or intentions, automated systems should be able to provide appropriate information about their models and decisions during the interaction in order to enhance the user's ability to provide useful and granular feedback back to the system. Considering existing taxonomies of Explainable AI methods for applications in Industry 4.0, Cyber-Physical systems for production lines and smart manufacturing [1, 16], XAI methods can enhance user's trust and reliability by informing the user about the system's understanding and capabilities. Moreover, XAI can be used to enhance the user's perception about their own decisions/actions, as well as shared understanding when interacting with multiple human users or AI agents (network of interactions).

Human Resources and Management. The goal of AI systems in human resources and management applications focuses on the automation of decision making tasks related to the organizational management of the employers, e.g., job performance and evaluation, recruitment, task allocation, etc. Such models are employed to make decisions based on predictive models, e.g., hiring of an applicant with a given profile/CV. In order to address challenges related to bias and unfair decisions, human supervisors must be able to have control over the decisions made by the AI and intervene when needed (human control). Human-in-the-Loop methods can enable employers to participate to the decision making process by providing their own feedback to the rest of the network (coworkers, supervisors and AI system), when needed. For example, if a company deploys an explainable CV mining system for recruitment or promotion applications, the same system can be provided and used by the candidates to evaluate their own CVs. This has several possible modes of interaction: Candidates can identify weaknesses in their CV, or missing skills that the algorithmic system has missed.

365 In this case, XAI and HITL can be utilized to enable the user to adjust their CV to ensure that the model can efficiently
366 model their actual skills. Moreover, candidates may also utilize explainability and transparency to identify possible
367 career development paths, e.g., identify what is required to obtain a skill or get a position. Candidates may also notice
368 skills or qualities which are missing from the model (job description) but may be important for the job. Additionally,
369 candidates may notice job skills and requirements that may be more demanding than the specific job should require.
370 Such feedback opens the potential to those offering a position to rethink the requirements.
371

372 **Human-AI Teaming and Collaboration.** Research in AI and collaborative systems aims to analyze and model the
373 dynamics of human-human collaboration towards developing methods to support the collaboration between human-
374 human and human-AI collaboration [5]. More specifically, AI methods have been proposed to identify teamwork skills
375 during collaborative problem solving. In terms of human-AI teaming, a research study investigated how people perceive
376 AI teammates and what they expect from AI teammates in human-AI teaming [23]. Based on their findings, the authors
377 suggest that AI should be considered as a subject in collaborative-activity design and they highlight the need to enhance
378 user's perception about the capabilities and intentions of AI (XAI) towards effective and safe human-AI teaming and
379 collaboration. For example, during a AI-mediated brainstorming session, the system can analyze the behavior dynamics
380 and patterns and visualize them to the users. Explainability gives the opportunity to the users to get insights about
381 the group dynamics (e.g., who are the dominant speakers), as well as the individual contribution to the activity. HITL
382 can enable all participants (team members, leader, moderator) to provide feedback to the system to either evaluate the
383 group performance or negotiate the visualized models.
384

385 **Well-being and Human Factors.** The integration of AI systems in the workplace can raise challenges considering
386 the relationship between emerging technologies and human workers. Research in Human Factors and Ergonomics
387 investigates the impact of digital technology on the mental health (workload and stress) of employees [?]. Human
388 factors, including job skills, job satisfaction, and job fatigue may have a significant effect on job performance and
389 should be considered during the design of human-AI interactions in the workplace. A recent research study investigated
390 the relationship between worker's emotions and reliance on automated systems [12]. Human employers should be
391 supported to self-manage the human factors that may affect their performance and well-being. Individual models of
392 workplace wellbeing can help managers to keep track of remote workers stress levels. However, they can also serve as
393 signals about what the organisation feels is important: a wellbeing model is an encoding of the ways in which the work
394 may be problematic for people, and as the models improve, they provide important documentation about what it is
395 to work in a place. Similarly, the possibility for workers to contest, ignore, reject or otherwise annotate the models
396 description of their mental wellbeing can serve as a site for an organisation to better understand the needs and practices
397 of its workers. HITL approaches offer moments to go beyond what the model is seeing, and develop context that is
398 particularly necessary when considering the health of remote and distributed workforces.
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416

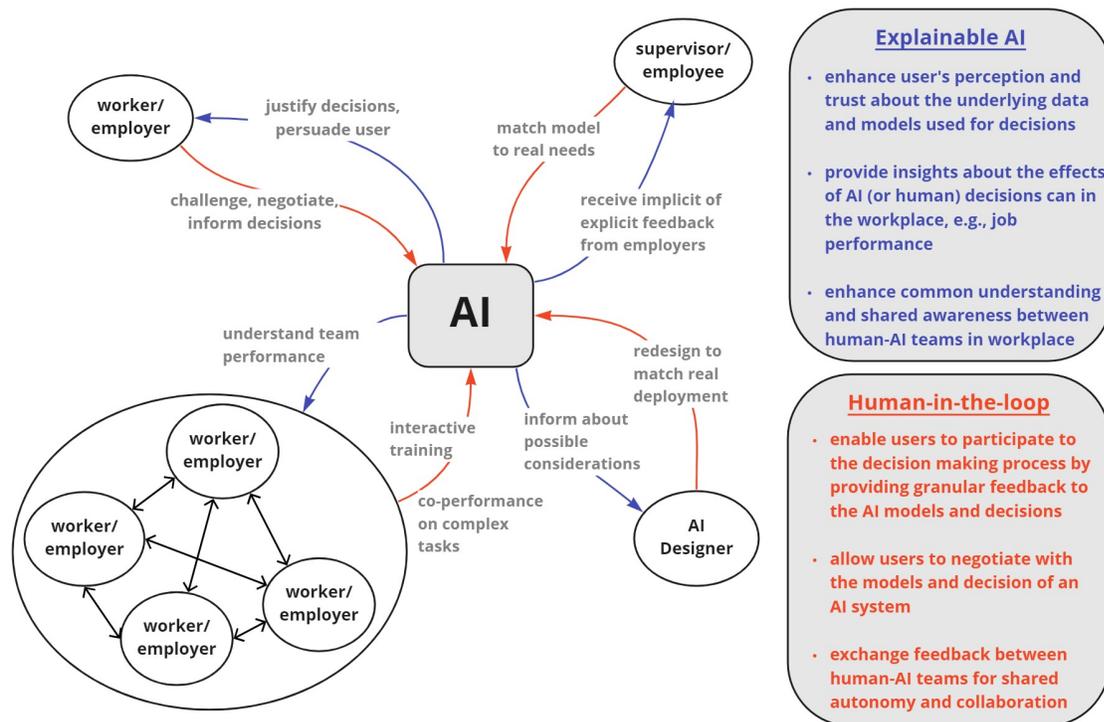


Fig. 2. Interactions with Human-in-the-Loop and Explainable AI in future work practices

5 CONCLUSION

In this paper, we discuss the potential benefits of XAI and HITL methods in future work practices. Taking onto consideration existing guidelines and framework fr the design of Human-AI interactions, as well as design practices for AI applications in the workplace, we present our discussion points towards defining a design space for explainable and interactive AI systems in the context of the Future of Work. We discuss the potential benefits of the integration of HITL and XAI methods in the FoW, as well as the possibilities for new design practices that can emerge in a synergistic AI framework.

REFERENCES

- [1] Imran Ahmed, Gwanggil Jeon, and Francesco Piccialli. 2022. From Artificial Intelligence to eXplainable Artificial Intelligence in Industry 4.0: A survey on What, How, and Where. *IEEE Transactions on Industrial Informatics* (2022).
- [2] Saleema Amershi, Dan Weld, Mihaela Vorvoreanu, Adam Fourney, Besmira Nushi, Penny Collisson, Jina Suh, Shamsi Iqbal, Paul N Bennett, Kori Inkpen, et al. 2019. Guidelines for human-AI interaction. In *Proceedings of the 2019 chi conference on human factors in computing systems*. 1–13.
- [3] Jcastillo2019study José Castillo, Edith Galy, Pierre Thérouanne, and Raoul Do Nascimento. [n. d.]. Study of the mental workload and stress generated using digital technology at the workplace. In *H-Workload 2019: 3rd International Symposium on Human Mental Workload: Models and Applications (Works in Progress)*. 105.
- [4] Chien-Chun Chen, Chiu-Chi Wei, Su-Hui Chen, Lun-Meng Sun, and Hsien-Hong Lin. 2022. AI Predicted Competency Model to Maximize Job Performance. *Cybernetics and Systems* 53, 3 (2022), 298–317.

- 469 [5] Pravin Chopade, David Edwards, Saad M Khan, Alejandro Andrade, and Scott Pu. 2019. CPSX: Using AI-Machine Learning for Mapping Human-
470 Human Interaction and Measurement of CPS Teamwork Skills. In *2019 IEEE International Symposium on Technologies for Homeland Security (HST)*.
471 IEEE, 1–6.
- 472 [6] Carina Dantas, Karolina Mackiewicz, Valentina Tageo, Giulio Jacucci, Diana Guardado, Sofia Ortet, Iraklis Varlamis, Michail Maniadakis, Eva de
473 Lera, João Quintas, et al. 2021. Benefits and Hurdles of AI In The Workplace–What Comes Next? *International Journal of Artificial Intelligence and*
474 *Expert Systems* (2021).
- 475 [7] Christos Emmanouilidis and Sabine Waschull. 2021. Human in the Loop of AI Systems in Manufacturing. In *Trusted Artificial Intelligence in*
476 *Manufacturing: A review of the emerging wave of ethical and human-centric AI technologies for smart production*. NOW PUBLISHERS INC, 158–172.
- 477 [8] Elisa Giaccardi and Johan Redström. 2020. Technology and more-than-human design. *Design Issues* 36, 4 (2020), 33–44.
- 478 [9] Sara Hekkala and Riitta Hekkala. 2021. Integration of Artificial Intelligence into Recruiting Young Undergraduates: the Perceptions of 20–23-Year-Old
479 Students. (2021).
- 480 [10] Jung-Sing Jwo, Ching-Sheng Lin, and Cheng-Hsiung Lee. 2021. Smart technology–driven aspects for human-in-the-loop smart manufacturing. *The*
481 *International Journal of Advanced Manufacturing Technology* 114, 5 (2021), 1741–1752.
- 482 [11] Lenneke Kuijer and Elisa Giaccardi. 2018. Co-performance: Conceptualizing the role of artificial agency in the design of everyday life. In *Proceedings*
483 *of the 2018 CHI Conference on Human Factors in Computing Systems*. 1–13.
- 484 [12] Stephanie M Merritt. 2011. Affective processes in human–automation interactions. *Human Factors* 53, 4 (2011), 356–370.
- 485 [13] Beth Porter and Francesca Grippa. 2020. A Platform for AI-Enabled Real-Time Feedback to Promote Digital Collaboration. *Sustainability* 12, 24
486 (2020), 10243.
- 487 [14] Lionel P Robert, Casey Pierce, Liz Marquis, Sangmi Kim, and Rasha Alahmad. 2020. Designing fair AI for managing employees in organizations: a
488 review, critique, and design agenda. *Human–Computer Interaction* 35, 5-6 (2020), 545–575.
- 489 [15] Ben Shneiderman. 2020. Human-centered artificial intelligence: Reliable, safe & trustworthy. *International Journal of Human–Computer Interaction*
490 36, 6 (2020), 495–504.
- 491 [16] Georgios Sofianidis, Jože M Rožanec, Dunja Mladenčić, and Dimosthenis Kyriazis. 2021. A Review of Explainable Artificial Intelligence in
492 Manufacturing. *arXiv preprint arXiv:2107.02295* (2021).
- 493 [17] Konrad Sowa, Aleksandra Przegalinska, and Leon Ciechanowski. 2021. Cobots in knowledge work: Human–AI collaboration in managerial
494 professions. *Journal of Business Research* 125 (2021), 135–142.
- 495 [18] Harini Suresh, Steven R Gomez, Kevin K Nam, and Arvind Satyanarayan. 2021. Beyond expertise and roles: A framework to characterize the
496 stakeholders of interpretable machine learning and their needs. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*.
497 1–16.
- 498 [19] Niels van Berkel, Mikael B Skov, and Jesper Kjeldskov. 2021. Human-AI interaction: intermittent, continuous, and proactive. *Interactions* 28, 6
499 (2021), 67–71.
- 500 [20] Danding Wang, Qian Yang, Ashraf Abdul, and Brian Y Lim. 2019. Designing theory-driven user-centric explainable AI. In *Proceedings of the 2019*
501 *CHI conference on human factors in computing systems*. 1–15.
- 502 [21] Qian Yang, Nikola Banovic, and John Zimmerman. 2018. Mapping machine learning advances from hci research to reveal starting places for design
503 innovation. In *Proceedings of the 2018 CHI conference on human factors in computing systems*. 1–11.
- 504 [22] Qian Yang, Aaron Steinfeld, Carolyn Rosé, and John Zimmerman. 2020. Re-examining whether, why, and how human-AI interaction is uniquely
505 difficult to design. In *Proceedings of the 2020 chi conference on human factors in computing systems*. 1–13.
- 506 [23] Rui Zhang, Nathan J McNeese, Guo Freeman, and Geoff Musick. 2021. "An Ideal Human" Expectations of AI Teammates in Human-AI Teaming.
507 *Proceedings of the ACM on Human-Computer Interaction* 4, CSCW3 (2021), 1–25.