Preparing humane ML experts for a better future. Experiments with design and engineering students

10.29073/jer.v2i1.21

Received: January 15, 2024.
Accepted: March 7, 2024.
Published: July 9, 2024.

Author: Martina Sciannamè, Politecnico di Milano, Italy, martina.scianname@polimi.it.

Abstract

Recognizing the rising demand for well-trained professionals in the responsible AI (RAI) landscape, the study explores which skills might characterize humane ML experts. A literature review outlines human centricity, ML-savvy, and value sensitivity as pivotal qualities for responsible practices, materializing an overarching multidisciplinary approach to the design of meaningful ML-infused solutions. For a preliminary and qualitative investigation, four experimental workshops were conducted in different European universities, targeting design and computer engineering students across different educational levels. They were intended to (i) translate the presented skills into educational experiences; (Q2) assess the effectiveness of the experimentations to foster these competences; and (Q3) evaluate their suitability and meaningfulness. Adapting the theoretical assumptions to the target audiences' backgrounds, positive results emerged. Both design and engineering students exhibited receptiveness and appreciation for the contents, methods, and tools presented in the workshops, emphasizing the transversal and essential nature of the proposed skills in diverse educational contexts. Despite the limits of the experimentations, the research argues that the depicted skills might orient designerly and technical ML experts toward meaningful outcomes, especially if they build effective collaborations leveraging their complementary strengths. Hopefully the contribution offers insights to advance the discourse about future professional figures in RAI.

Keywords: Humanity-Centered Design; ML Design Education; ML-Savvy; RAI Design Skills; Value Sensitivity.

Introduction

Technological innovations have always had a unique capability to shape the evolution of human history, in ways that go far beyond products, services and other applications. Electricity, personal computers, the Internet, just to name a few, have dramatically impacted people's lives and, indirectly, affected their entire ecosystems. Repercussions, in fact, can be measured in personal well-being, social dimension, natural environment, and even climate.

Recent advancements in technology have significantly boosted computational power and data storage and retrieval capabilities. This, together with substantial investments by public and private entities — China and the US being exemplar cases — has created an ideal environment for the flourishing and spreading of machine learning (ML), which has emerged as a crucial subset of artificial intelligence (AI). Undoubtedly, ML has the potential to improve people's lives in ways that are not fully foreseeable yet. However, this newfound power also entails new responsibilities, as malicious scopes could lead to harmful outcomes. From deepfakes and misinformation, to reinforcement of discrimination and even social manipulation, there are several risks to watch out for.

As technological development and profit prospects have fueled the current growth of ML, the design of products and services incorporating this technology is predominantly driven by marketing objectives or technological experimentations, motivated by the desire to unveil novel frontiers. Unfortunately, these approaches tend to overlook broader perspectives rooted in the socio-technical nature of these systems (Antonelli, 2018; van de Poel, 2020; Yang, Banovic, et al., 2018). This issue is acknowledged both by computer science and ethics fields, and has led to the emergence of several research, experimentations, and theoretical assumptions aimed at mitigating the risks associated with the unchallenged advancement of AI and ML. Human-related disciplines are particularly keen at questioning, rising concerns, and attempting at reframing AI-related principles, as evidenced by the escalation of publications of ethical guidelines (Algorithmic Watch, 2020).

Two main barriers hinder the trustworthiness of this technology and contribute to undesired outcomes, such as confusion, uncertainty, frustration, and ultimately, mistrust, which the public interprets as a sense of *creepiness* (Fruchter & Liccardi, 2018). First, AI remains overly opaque, especially — but not only — for those who are not experts in the field. Starting from the ambiguous allusion to human intelligence, which is a difficult point for even computer scientists to agree on (Russell & Norvig, 2020), the narrative around AI is often misleading. The inner functioning and capabilities are rarely properly communicated, which often leads to the perception of AI as overly technical and arcane. Several authors have noted how it can be conceived as a mysterious or even monstrous entity, placing users in a context of unfamiliarity and lack of control that not only generates concern but also instills fear (Antonelli, 2018; Dove & Fayard, 2020; Johnson & Verdicchio, 2017; Kulesz, 2018). A further layer of unclarity is related to the concept of responsibility which, for AI and ML systems, may be contested among different human parties and the technology itself. In particular, this brings us back to the second barrier: human factors are often not considered or incorporated. As noted by Johnson and Verdicchio (2017), on the one hand, the prevalent narrative tends to associate the autonomy of ML systems to the idea that machines can operate beyond human control. On the other, the authors highlight a *sociotechnical blindness*, meaning that "the essential role played by humans at every stage of the design and deployment of an AI system" is often neglected. Indeed, Ibo van de Poel (2020), professor in ethics and technology at TU Delft, sustains that AI systems should be regarded as *socio-technical systems*, implying that they should be considered in combination with human behavior, social arrangements, and meaning at a larger scale. What sets AI systems apart is their peculiarity as *artificial agents*, enabling them to actively influence other components within the larger systems they are part of.

Failing at acknowledging the socio-technical dimension of products and services integrating AI and ML systems means that users' expectations, needs, and mental models are not addressed in their development. As a consequence, at best, ML-infused artifacts turn out to be gadgets and toys (Levinson, 1977) that crowd people's homes without actually enriching their lives. Yet, incautious design policies might lead to more worrisome situations like misinformation, power imbalance, manipulation, or intendedly harmful outcomes.

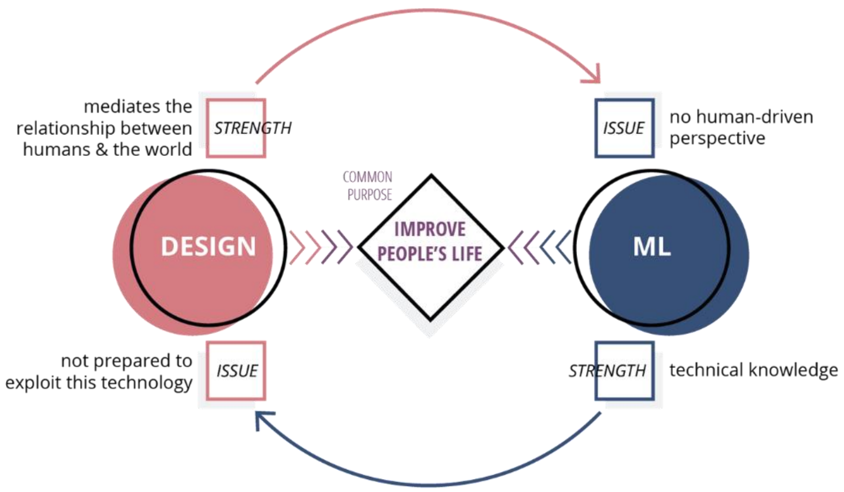
Therefore, although the work of AI and ML experts is essential to implement and further this technology, a more holistic perspective is imperative. In this rapidly evolving field, additional competences and professionals from other domains might complement current practices. New, essential figures should possess the ability to comprehend and navigate a context determined by the intricate web of relationships within it. This aligns with the principles of the Actor-Network-Theory (Latour, 1996), for which a clear distinction between humans and non-humans, people and technology, science and nature (Law, 2015) cannot be traced, as they are all actors in an interconnected reality. Thus, multi-competencies teamwork, as advocated by several researchers and practitioners (Davis, n.d.; Frascara, 2020; K. Friedman et al., 2019; M. W. Meyer & Norman, 2020; Rismani & Moon, 2023; Voûte et al., 2020; Wang et al., 2023), might play a central role in the future of digital transformation.

Based on these premises and considering the swift broadening of AI-related job positions, especially toward AI ethics expertise (Rismani & Moon, 2023), the research presented in this article aims at identifying the basic skills that future practitioners dealing with ML — especially from a design and development perspective — need to acquire for an adequate preparation to conceive and implement responsible ML-based solutions. Specifically, a theoretical inquiry allowed the identification of three main competencies for humane ML experts (human-centricity, ML-savvy, and value sensitivity) with an overarching need for multidisciplinarity. Their effectiveness and meaningfulness for the context under investigation has been tested in workshops addressing ML and design students, envisioned as primary actors in the future of designing ML-infused artifacts.

Indeed, incorporating diversity and multidisciplinary competences in development teams is widely advocated, as it is essential for the responsible exploration of AI and ML. This is warmly recommended by multiple sources such as Cutler et al. (n.d.), the High-Level Expert Group on Artificial Intelligence (2019b), and the World Economic Forum Global Future Council on Human & Rights 2016–18 (2018). In particular, while the fundamental importance of the role of ML engineers and computer scientists is obvious, that of designers is starting to be recognized (sometimes implicitly) and encouraged but not yet effectively fostered in practice. To go beyond technology and market-driven solutions, designers are equipped with relevant skills to handle the uncertainties that ML presents today (Antonelli, 2018; Yang, 2020), especially leveraging multidisciplinary approaches (M. W. Meyer & Norman, 2020). Designers are accustomed to navigating through ill-defined, fluid, and constantly evolving contexts (Auernhammer & Ford, 2022; M. W. Meyer & Norman, 2020). Thanks to their long-standing "experience with technology, transformative influence, cross-disciplinary predisposition, system-level thinking, and empathy" (Sciannamè, 2023), they best respond to several requirements that technology companies are looking for but in different disciplinary fields.

Therefore, both ML experts and designers possess pivotal capabilities to affect the development of ML-infused solutions and they have the potential to steer it towards meaningful applications for a flourishing future (Figure 1). However, some skills to overcome existing gaps need to be identified and nurtured. In the following, the results of this investigation are presented to contribute to the conversation about the future of professional figures embodying digital maturity and responsibility.

Figure 1: Framework for Design and ML cooperation.



Source: Developed by the author.

2. Identifying Strategic Skills for Humane ML Experts. A Literature Review

Over the last five years, substantial research has surfaced across academia, public institutions, and private sector, aiming to instill greater ethical considerations into AI and ML. This has led to the generation of research strands such as Explainable AI (XAI), Interpretable ML, Responsible AI (RAI), and Fair ML. Yet, coherently with the premises and the scope of this study, two main weaknesses can be recognized in the plethora of publications about this subject. On the one hand, most of the explorations are computational in nature, solely relying on computer science and engineering principles and practices (Varanasi & Goyal, 2023). Focusing on solutions exclusively related to the creation of datasets, models performances and similar issues, they perpetuate the *sociotechnical blindness* (Johnson & Verdicchio, 2017) discussed above. On the other, a lot of ethics frameworks and guidelines emerged but they predominantly remain at a theoretical level, delineating principles and values without often delving into practical implementation strategies or emphasizing the essential skills needed to effectively advocate for responsible practices. Only very recent studies are recognizing the necessity to broaden the purview and focus on practical skills and applications (Rismani & Moon, 2023; Varanasi & Goyal, 2023; Wang et al., 2023).

In this variegated context, a thematic analysis was employed to identify relevant elements for inferring key skills to make the professionals who deal with the design of ML-infused solutions more responsible and humane. As the specific research for necessary competences or recommendations for responsible AI or ML produced scarce pertinent results from Scopus and ACM Library, the ethics guidelines for AI collected in the online repository Algorithmic Watch (2020) composed the main reference set. Only documents in English were incorporated, specifically those directly touching on the design and development of AI and ML systems that did not present an exclusively technical and specialized perspective. Particularly noteworthy ethics toolkits, which offered principles that otherwise were not well-represented, were identified through snowball sampling and subsequently included.

Given the frequent absence of explicitly mentioned desirable skills or competencies, the analysis focused on extracting the emphasized values, the impediments posing risks to responsible ML outcomes, and potential measures and solutions. Despite the granularity of the results, the author identified four overarching areas of interest, with three of them acknowledged as transferable skills for humane ML experts. These areas aim to highlight the most notable competences that are currently missing and would be beneficial in the development of meaningful ML-infused artifacts. Namely, human-centricity, ML-savvy and value sensitivity, under the encompassing need for multidisciplinarity, resulted the most relevant for achieving responsible practices (Norman, 2023; Umbrello & van de Poel, 2021).

2.1. Human Centricity

Probably a direct consequence to the acknowledgement of ML solutions lacking the inclusion of human factors, human centricity is the most pervasive competence encountered. Considering the possible threats that ML systems might pose, *misuse* was the most recurrent (with 17 occurrences). Major concerns are the harmfulness, weaponization, and abusability of the systems, mischievous tracking purposes, as well as possible discrimination and oppression. Dependency on (biased) training data (14 occurrences) and non-neutrality of the technology (9), because it embeds and amplifies beliefs and behaviors, are the following retrieved risks. All of them imply the poor consideration of people's influence, interests, and mental models throughout the design process. This is also reflected in the most common possible solutions to current flaws of ML-infused solutions, which encompass: including different perspectives and people (28), facilitating the understanding of ML systems behavior (26), monitoring and evaluating with not only technical parameters (22), using foresight methodologies (13), and applying a design-driven approach for problem framing (11). How literature leans toward human-centered approaches and methods as beneficial means to achieve meaningful solutions is confirmed by some studies aiming at identifying ways and skills to design responsible AI. Specifically, Wang and colleagues (2023) report the underrated but essential work that some UX practitioners and RAI experts at a big tech company are doing to sensitize AI teams toward possible responsible practices by exploring and introducing AI users' mental models, feasibility and user acceptance, UX methods and tools to conduct user tests early in the process, and participatory approaches to AI design. Furthermore, while investigating responsible AI practitioners' roles and skills from a computer science-centric perspective, Rismani and Moon (2023) manifest the need for "*creative problem solvers* who can work in a *fast-changing environment*" for "translating research into design, technology development, and policy." Interestingly, their depiction of the role of a more human-centered researcher does not explicitly recognize designers as suitable figures, while outlining most of their characterizing qualities. Instead, they suggest looking for individuals with backgrounds in, for instance, human-computer interaction, cognitive psychology, experimental psychology, digital anthropology, and social sciences. Still, this is further confirmation of the pivotal role a human-centered perspective can play in addressing the multifaceted challenges posed by ML systems.

2.2. ML-Savvy

In-depth technical expertise is undisputedly a fundamental requirement for the realization of ML systems. However, current products and services prove that it is not sufficient to achieve meaningful results. In order to find a purpose for ML applications in the real world, a different level of knowledge is needed. It should be more oriented toward practice and making sound judgments, as reported especially in the literature outside of computer science. Several sources note how just high-level conceptualizations allow UX practitioners to envision ML solutions (Wang et al., 2023; Zdanowska & Taylor, 2022; Yang, Scuito, et al., 2018).

Indeed, lacking specific disciplinary knowledge, methods, and tools, some UX designers developed a functional understanding of this technology, utilizing abstractions to comprehend the capabilities of ML systems in relation to users' needs (Yang, Scuito, et al., 2018). They found ways to adapt traditional UX ways of prototyping and testing to better grasp ML models capabilities and communicate them to users, to set their mental models in the right direction, and anticipate potential harms (Wang et al., 2023). Additionally, they managed to navigate ML complexity, sometimes with PoCs that helped them comprehend the constraints of models, datasets, and technical feasibility at large. They were also enabled to clarify the scope, potential issues, and acceptability of a solution with respect to the problem to solve (Zdanowska & Taylor, 2022).

Therefore, multiple layers for acknowledging ML systems can be uncovered and refined to cope with very practical issues. Interestingly, these approaches diverge significantly from conventional ML educational contents and express a different modality to make sense of this technology, possibly favoring the identification of meaningful applications.

2.3. Value Sensitivity

The massive efforts in outlining principles and values to steer the development of AI and ML systems has limited value in terms of practical applications. Yet, it provides a useful indication about what is still missing: a value sensitive approach to design. Even human-centered methods and tools have no guarantee of ethically acceptable outcomes.

Recently, ethics guidelines, Responsible Research and Innovation (RRI), and vertical studies on responsible AI practices have spotlighted the necessity to think about the values a ML-infused solution might threaten, promote, or preserve. Value Sensitive Design (VSD), a methodology that Batya Friedman depicted in the 1990s to advocate human principles in the planning of technology (2019), has gained renewed significance. Umbrello and Van de Poel (2021) proposed a value-sensitive design process tailored for AI, encompassing the entire life cycle of AI systems to monitor algorithm evolution and proactively detect potential unintended consequences. Their adaptation of VSD encompasses four phases. The first, *context analysis*, considers societal challenges, existing technology, and stakeholder values. The second, *values identification*, recommends conceptual, empirical, and technical investigations, with a differentiation between promoted and respected values. The authors particularly insist on an explicit orientation toward positive impacts, like the Sustainable Development Goals of the United Nations. The third phase involves the formulation of design requirements based on the previous steps. Finally, the prototyping phase tests whether the system meets design requirements and incorporates the identified values.

The challenge of translating theoretical concepts into tangible strategies and measures persists. Even regulatory efforts, like the AI Act, primarily operate at a high level. Nevertheless, there is a growing acknowledgment of the imperative to integrate value sensitivity into everyday practices. (Rismani & Moon, 2023; Varanasi & Goyal, 2023; Wang et al., 2023).

2.4. Multidisciplinarity

Finally, it is worth mentioning how being able to deal with and mediate across multiple disciplines is a competence that was encountered in almost all the resources during the analysis. Having multidisciplinary teams, multidisciplinary ML practitioners able to bridge disciplinary boundaries, work across different functionalities and disciplines, and act as facilitators or translators (High-Level Expert Group on Artificial Intelligence, 2019b; Rismani & Moon, 2023; Varanasi & Goyal, 2023; Wang et al., 2023) are only some examples of how the discourse articulate, and they build on a larger debate about necessary skills for tackling complex challenges (K. Friedman et al., 2014; M. Meyer, 2010).

However, this was not included in the set of transferrable competences for humane ML because, according to the researcher it corresponds to a higher level that is difficult to observe and measure. Yet, it can be considered an overarching requirement that is inherently intertwined and can be effectively conveyed through the combination of the previously identified skills. In fact, these are representative of three different disciplinary components, respectively, design, ML, and ethics.

3. Methodology

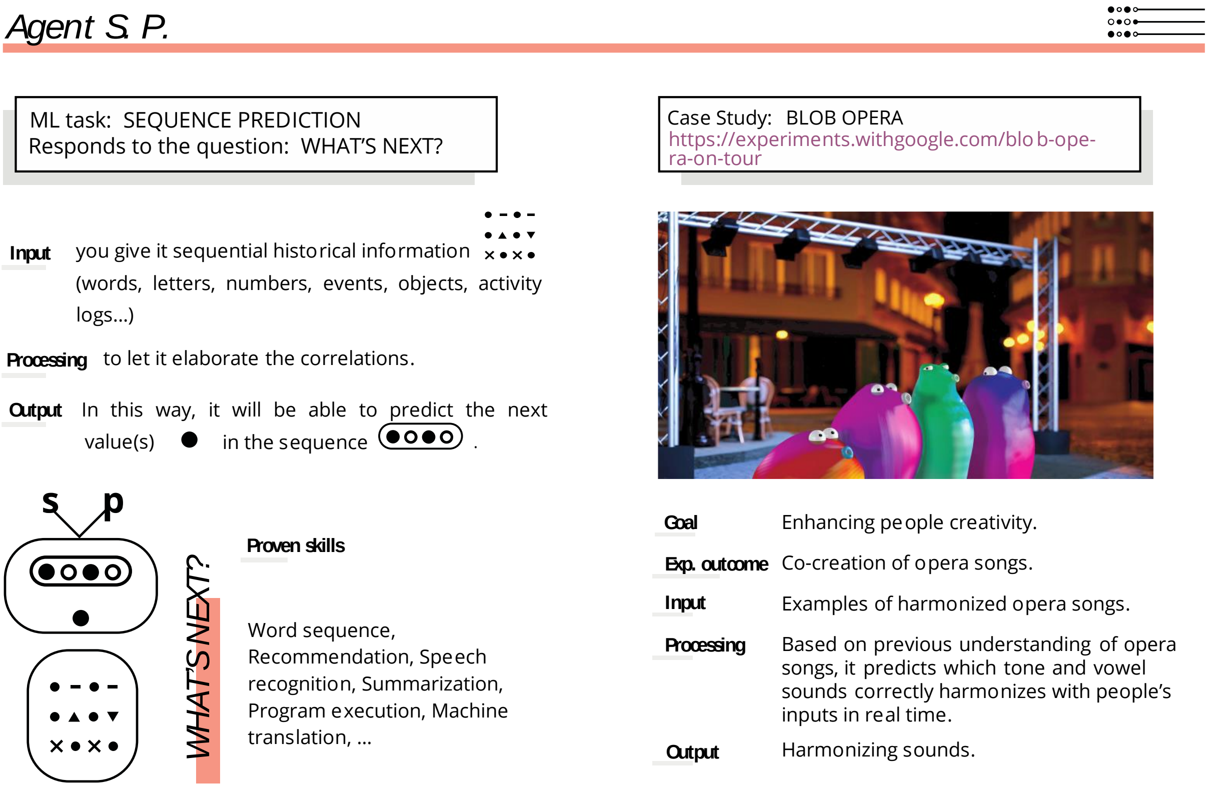
Due to the relatively recent attention AI and ML are attracting, especially from the empirical perspective of the design discipline, an action research approach has been selected to test the theoretical assumptions emerging from the literature review through a practical and systematic investigation (Archer, 1995). Specifically, following a long-lasting tradition (Bresler, 2021; Robson & McCartan, 2015), the action research enquiry applies to the educational context as a natural sandbox to nurture new skills and experiment with formative modalities, and it acquires a mainly qualitative character, building on *a reflection-in-action* process (Schön, 1983).

Therefore, four experimental workshops, addressing design and engineering students as key figures for the future development of meaningful ML-infused products and services, were organized, and are presented here as case studies. Their main purposes were to (Q1) understand how to translate the presented skills for humane ML experts into educational experiences; (Q2) assess the effectiveness of the workshops to elicit such competences (considering aspects like the contents, the format, the principles, the methods, and the tools provided); and (Q3) understand the suitability and meaningfulness of the promoted skills.

To grant diversity, the workshops took place in different European universities and involved international students at different stages of their formative paths. More precisely, W1 and W2 targeted design students. The first was held in FH Joanneum University of Applied Science, in Graz (Austria). It developed over three days, for a total of 18 hours, and engaged seven students (5 females and 2 males) attending their first year of master’s degree in interaction design (6) and in media design (1). They worked in groups of two or three people. The second 16-hours workshop also spanned over three days and involved 15 third-year bachelor design students (13 females and 4 males) from the Universidade da Madeira, where the program is not differentiated in any design specialization. They were organized into 3 groups of 5 people. The other two workshops (W3 and W4) were open to both design and engineering students and took place in Politecnico di Milano as part of an ATHENS network initiative. Eventually, though, they resulted in workshops for people with a technical background and quite familiar with ML. Indeed, W3 involved 17 students (11 males, 6 females), of which only two had an industrial design background and one was enrolled in a computing architecture program — all at Istanbul Technical University. The others were studying computer science (6), computer or software engineering (7), and technical physics (1). W4 counted 24 students (17 males, 7 females) from diverse technical backgrounds — industrial engineering and management (6), computer science (5), computer engineering (5), economy and finance (3), data science (1), telecommunication (1), geomatics (1), aerospace (1), physics (1) — and none from a design background. The participants to both the latter workshops differed by year of enrollment (first through sixth) and by university of origin, namely Instituto Superior Técnico (Lisbon, Portugal), Istanbul Technical University (Turkey), KU Leuven (Belgium), Paris Tech (France), Technische Universität München (Germany), TU Wien (Austria), TU Delft (Netherlands), University Politehnica of Bucharest (Romania), and Warsaw University of Technology (Poland).

All the workshops shared the theoretical assumptions previously presented as constituents of the formative experience as well as the main objectives of (i) putting the essential skills for humane ML experts in practice, and (ii) enabling the participants to envision meaningful solutions that integrate ML capabilities for improving the quality of life (at the level preferred by the participants) to ensure a better future. Additionally, the same tools were used to facilitate the comprehension and application of ML systems towards responsible applications. For the first purpose, the *ML Agents* (Figure 2) aimed at transferring basic ML knowledge, focusing on the main capabilities the technology offers and synthesizing its description according to an input-processing-output structure and complementing it with a concise question, a graphic representation and an example.

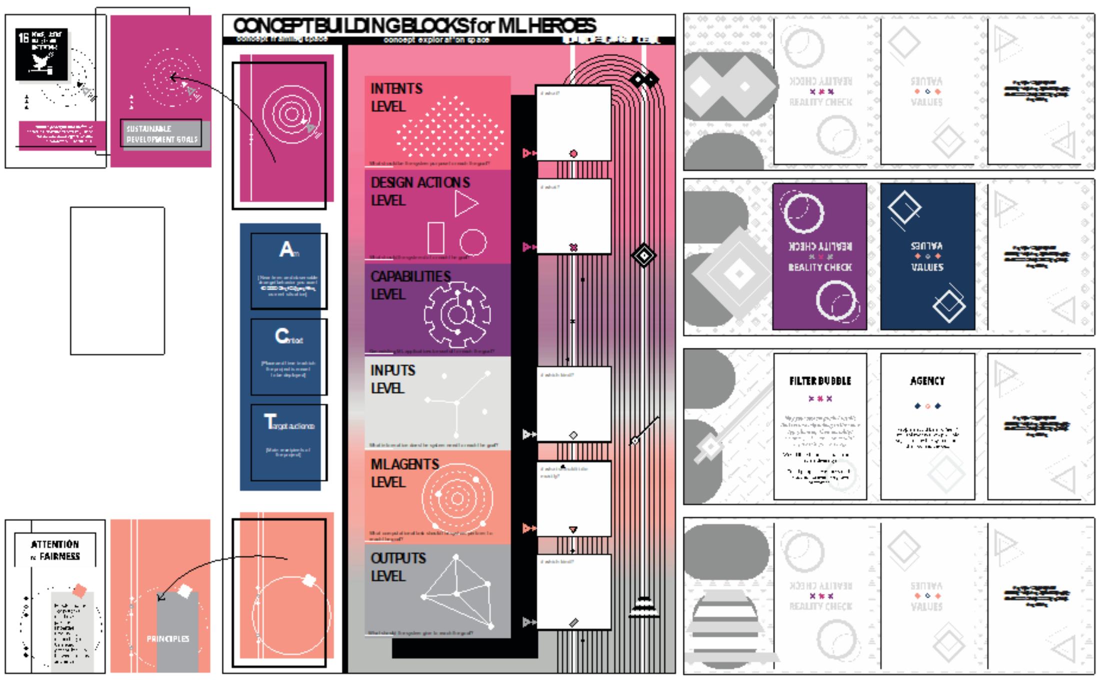
Figure 2: ML Agent example: Sequence Prediction Agent.



Source: Developed by the author.

For the latter, the *Concept Building Blocks* (CBB) and *VALUable by Design Expansion* (VDE) (Figure 3) were meant to give procedural support and inspiration for envisioning meaningful ML-infused artifacts, respectively focusing on the materialization of the technical possibilities into design requirements and actions and on the ethical implications. Further details can be found in the author's doctoral thesis (Sciannamè, 2023).

Figure 3: Concept Building Blocks and VALUable by Design Expansion.



Source: Developed by the author.

For the scope described in this article, qualitative methods were applied to collect data, specifically observation, analysis of the workshops' outcomes, oral and written feedback (Robson & McCartan, 2015), and rigorous procedures were followed to ensure the validity and reliability of the data. To support the researcher's observation, an annotation sheet was structured to report on student's reactions and responses to teaching-learning activities, contents, and tools. The delivered concepts were evaluated by both the researcher and the class according to their relevance, consistency with the technological capabilities, ethical acceptability, (personal, social, environmental) sustainability, and overall desirability. The feedback collection, instead, was encouraged both in an anonymous, paper-based form, via the digital platform used throughout the workshops (Miro), and in a final semi-structured focus group. It revolved around assessing the effectiveness of educational experience in transferring the identified skills for humane ML experts, as well as their usefulness. A thematic analysis was employed to identify recurring themes and patterns serving the purposes of the investigation.

Of course, the study presents some limitations. Because of its qualitative nature the researcher's perspective and the contexts in which the workshops took place may affect the non-deterministic results. Additionally, the workshops produced *"local"* understanding (Koskinen et al., 2011), which might hinder the generalization of the findings. For this reason, some mitigating measures have been adopted, such as involving a diverse target audience — in terms of cultural and educational background — and collecting rich and detailed feedback for analysis.

4. Q1 Results: an Educational Framework for Humane ML Experts

The identified skills for humane ML experts are assimilable to the soft skills domain. In fact, even if they have a specific objective, they can be transversally applied to different fields and, more importantly, they do not rely on any predetermined formula, but rather on personal understanding and elaboration. For this reason, a *constructivist approach* was adopted as the foundation for the design of the educational experience. It is based on the premise that knowledge arises through the construction of meaning, individually produced through the internal synthesis of emotions, prior knowledge, value systems, and beliefs. Accordingly, and in line with the practical goal of envisioning meaningful ML applications, a *project-centered* pedagogical framework was proposed to engage students in activities grounded in experiences, abstractions, inferences, problem-solving, information recombination, and collaboration with peers (Sancassani et al., 2019). Indeed, Kirschener and Norman's (2021) perspective underlines how a project-centered approach resonates with the holistic character of the skills to be elicited, as it entails a wider scope than just solving a problem. It includes *"social, societal, economic, ethical, ecological aspects and so further of that solution"* (Kirschener & Norman, 2021). Therefore, finally, the *studio format*, dear to the design education tradition, was implemented in all the workshops.

Additional references for the construction of the educational framework are the key principles for RRI (von Schomberg, 2013), Gagné's *events of instruction* (Gagné et al., 1992), and the prior experimentations unfolded throughout the author's doctoral research (Sciannamè, 2023). After a few preliminary workshops to test and determine the preferable forms, language, essential contents and tools for synthesizing ML technical knowledge and related ethics principles and make them operational materials for designing meaningful ML solutions, an educational framework was outlined by adjusting Gagné's *events of instruction* (Gagné et al., 1992) to include prior findings and RRI focal points.

Specifically, three main parts characterized the workshops that had to lead students to the definition of a meaningful ML-infused concept: (i) an initial *context introduction* (including steps 1-3 described below), (ii) *exploration & framing* (steps 4–5), and (iii) a conclusive *critical reflection and evaluation* stage (steps 6–8) (Figure 4). These articulated as follows:

1. **Engagement and responsibility.** The first step focuses on capturing students' attention and arousing their interest in the topic. In light of Gagné's work, a learner-centered and holistic approach is emphasized since the beginning, trying to leverage their sense of responsibility and understanding of their role as designers of ML solutions. To achieve this, the workshops problematized the current

development of ML-infused artifacts to elicit practical and ethical commitment, encouraging the participants to make a difference by experimenting with non-technology-driven approaches.

2. **Expectations and impact.** Beyond explaining the objectives and intended learning outcomes of the educational activity, as Gagné envisioned this step, an infusion of RRI principles would be beneficial. To set the initial stages of the experimental design process on a positive note, the expectations should be broadened to include relevant challenges to address and strong motivations behind them. Indeed, "*designing for the right impact*" (von Schomberg, 2013) clarifies the purpose of the activity also from a higher level, and it can make students feel involved and aware of the relevance of the solution they should envision.

3. **Interaction with design materials.** Given that design and engineering are practical fields of study, this point underlines the significance of an interactive approach to knowledge transfer, which has demonstrated considerable success in workshops, and it emphasizes that every nurtured notion or skill becomes an integral part of a designer's toolkit for their work. This encompasses theoretical content (such as machine learning and ethics), tools (like ML Agents, CBB and VDE), values, examples, and case studies. The activation of metaphors, following Schön's perspective (1983), is also a component of this stage, contrasting with prior knowledge that may not be readily available to the target audience. These elements can be presented to students or left for independent discovery through research, observation, experiences, or prompts. The crucial aspect is active involvement in the learning process.

4. **Practical exploration.** Subsequently, the newly acquired materials need to be applied in practice to acquaint individuals with and comprehend the potential ways to utilize and benefit from them.

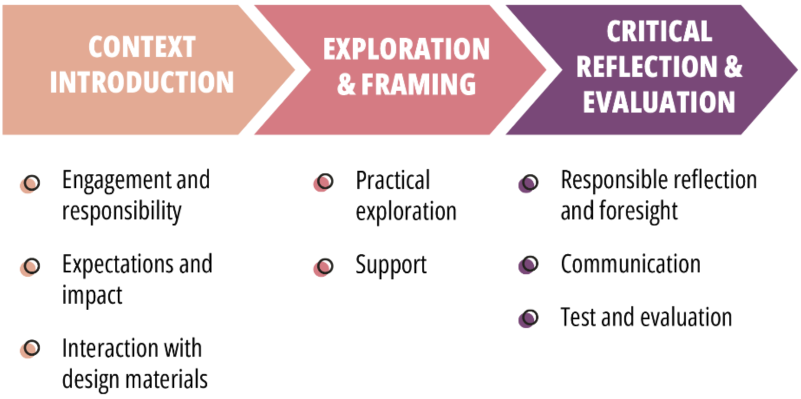
5. **Support.** Like in any studio format, support throughout all the practical activities has a pivotal importance and, as Gagné did, it needs to be marked as a specific point. Indeed, the figure of a facilitator reassuring, giving feedback, and orienting students in their explorations with new materials proved to be indispensable.

6. **Responsible reflection and foresight.** Necessarily, the generative phase should be complemented by explicit questioning and evaluating the emerging ideas. This involves contemplating UX and ethical aspects, guiding the concepts toward positive impacts with a value-driven approach, and foreseeing potential positive and negative outcomes to proactively address or mitigate certain risks. The VDE can serve as a valuable tool for this purpose.

7. **Communication.** This skill is overarching those identified for humane ML experts, which is why it should be expressly cultivated as part of the educational method. Indeed, for a successful design and development of meaningful ML solutions, multiple professional figures should be able to effectively collaborate. Thus, having the right vocabulary and means to properly communicate one's idea is essential for engaging with colleagues, users, or experts with different specializations, even if they share the same skills for humane ML experts.

8. **Test and evaluation.** Finally, testing and evaluation activities complete the learning process and contribute to determine how a responsible approach can unfold. While the workshops could not result in the practical development of the envisioned solutions, limited to low-fidelity prototyping, this stage holds importance even at the concept generation level. In fact, early foresight, along with peer and user evaluations, can help prevent unnecessary costs and efforts that might otherwise arise only after the deployment.

Figure 4: Workshops educational framework.



Source: Developed by the author.

From a content perspective, a few key elements were identified to stimulate the development of the skills for humane ML experts. Human centricity is pivotal in designerly processes and methods, which acquired a primary and pervasive importance, as already evident in the educational framework. To achieve ML-savvy, a significant way to synthesize ML knowledge as a tool to achieve meaningful ends was outlined to be useful for both designers and engineers. It bridged ML and design disciplines by merging the technical capabilities — ML tasks in (Russell & Norvig, 2020) — and the potential value this technology can bring to people through concrete design actions. For the comprehension of the basic functioning principle beneath the technology at hand, the definition of ML systems as agents (High-Level Expert Group on Artificial Intelligence, 2019a; Russell & Norvig, 2020), with a core input-processing-output structure, was retained as the founding element for the synthesis, resulting in the *ML Agents* knowledge-transferring tool. Eventually, value sensitivity, was encouraged by highlighting two main concepts: (i) ML systems are special kinds of sociotechnical systems in which the technology plays an active role in affecting people responses and behaviors (van de Poel, 2020), and (ii) embedding values in artifacts early in the design process should be an essential and explicit activity (van den Hoven, 2013).

Inevitably, different methodological approaches were adopted to structure the educational activities based on the different background and mindset of the target audience. Specifically, the workshops intended for design students insisted on the transfer of technical knowledge, leaving them more freedom in the development of their ideas, as they were following a process they were familiar with. On the contrary, engineering students needed more guidance for designing humanity-centered ML applications, as the conveyed design and value-driven process took them out of their comfort zone. However, in both cases, an introduction of ethics principles to be openly and intentionally introduced in the design process was necessary. Though, while ethics opened the theoretical argumentation of the workshops involving ML-related students, it was introduced to design students only after they intuitively started to figure out possible applications of ML capabilities, as a way to assess and iterate their concepts for improvements. Both strategies proved quite effective for their audiences, even if improvements might be implemented, as the following section depicts.

5. Findings for Q2: Assessing Skills Transfer for Humane ML Experts

As discussed in the previous sections, the transfer of skills to cultivate humane ML expertise involves a multifaceted strategy. This is explicitly outlined based on the inherent components of educational experiences, namely: the theoretical contents discussed and the fostered practical knowledge; the methods characterizing the educational activities and adopted to convey the skills; and the tools provided in support. Consequently, the results derived from the thematic analysis of delivered concepts, observation and feedback notes are synthesized and presented highlighting these dimensions and differentiating on the three skills under investigation — human centricity, tech-savvy, and value sensitivity.

5.1. Human Centricity

Contents. As prior knowledge and experience with this skill was expected from the design students involved in the workshops, their educational experience fostered human centricity completely through the practical definition of the design problem and the envisioning of a possible solution. The unfolding of the workshop and the final outputs confirmed that the future designers were already equipped with a people-oriented vision from an individual level to more comprehensive social and environmental scales. Having the implementation of (at least) one ML capability as a project constraint and giving design students the freedom to approach the development of the idea following a *problem-based, data-driven, technology-driven, artifact-driven, value-based,* or *human-driven* perspective could have compromised the results. However, the workshops demonstrated that no matter the starting point, finding a relevant solution for people was always central in the participants minds and they confirmed that they felt to have always followed a human-centered process.

On the contrary, in workshops with a predominance of attendees from an engineering background it was essential to clearly depict the theoretical principles characterizing the design process and to highlight, step-by-step, how the practical activities related to the procedural structure. In general, during the classes, it was difficult to assess whether these abstract contents were effectively interesting and comprehensible for the audience, as the researcher observed a moderate attention level. Nonetheless, some of the students, in both W3 and W4, followed up with questions comparing their technical design experience with what was being explained. They expressed appreciation for the simple examples describing the different approaches to the same design problem (technology, data, human, and value-driven). However, further practical examples or case studies would have been desirable for a better understanding.

Methods. Promoting a studio environment, with the researchers' constant support and feedback provision throughout all the stages of the concept development and following a design-thinking process prioritizing people's perspectives (users and stakeholders alike) was expectedly very functional to the purpose at hand. Again, whether it did not represent a novelty for design students, it was useful for them to consistently define their problem setting (aim, context, and target audience) in a limited timeframe and with no possibility to conduct proper preliminary research.

Very different was the perception of engineering students, who were sufficiently disoriented and positively impressed by the common designerly way of approaching the development of a project. Two main aspects emerged as fundamental to shift their mindset toward a human-centered one: (i) the initial problem framing and (ii) the encouragement to iterate the idea. (i) All the groups presented a solution-oriented approach to the definition of the challenge to tackle, which means that instead of identifying possible issues or needs for people to fulfill they directly and regardless determined the solution they wanted to design. Only by providing several case-by-case examples and possibilities to start the design process from people's problems and perspectives (reframing the solution they were exploring in terms of a human-centered problem definition), the participants started to modify the narrative depicting their starting point in a way that focused on a problem to solve and opened the space to multiple possibilities. Ultimately, most of the groups were able to integrate a human-centered problem framing, even if after several attempts and with some effort. (ii) Analogously, once they set the outline of their solution, it was difficult to make them consider alternatives that could as well or better address their design problem, even if they encountered obstacles that hindered the feasibility or effectiveness of the idea itself. Here, the researcher's intervention, challenging them to explore further options, was essential to steer the solution toward a more meaningful result for the target audience. Although radical iterations have faced much resistance, realizing that they had the opportunity to pursue different paths, even after a first idea had been portrayed, was game changing in leaning toward human centricity. An example of this troubled process is represented by a group in W3. They aimed to detect employees' burnout through the collection of several physiological data, causing relevant privacy and accuracy issues. Being at an *impasse*, the researcher suggested they look at how people suffering from burnout are diagnosed and treated by professionals, hinting at the use of conversational data instead of physiological ones. However, while they seemed to get the point and the importance of looking at the problem from a lateral perspective, they finally could not strongly detach from their initial idea. They finally introduced a compromise: they kept the physiological parameters, which were transparently available to the users, and their ML solution would point at possible burnout indicators for them to raise their awareness (self-defining whether substantial changes were caused by physical activity or anxiety states) and communicate with their psychologist, who had further information to help them diagnose their patient's condition and provide personalized suggestions.

Despite the final outcomes, in the end, even the most technology-based thinkers attested that they have understood the value of framing the problem and iterate the solution trying to put themselves in the shoes of the people that should benefit from their projects. Some even affirmed that they wanted to implement these methods in their own university projects, despite the common optimization-oriented approach promoted by their professors.

An additional help to foster a human-centric mindset consisted in assigning to each group member a different role and personality to embody during the design and peer evaluation processes, similarly to de Bono's thinking hats (De Bono, 2016). Nonetheless, this strategy has led to limited outcomes, and the students involved in W4 conclusively recognized the potential of the role-playing but would have preferred even more explicit activities in this sense.

Tools. The CBB tool was specifically intended to facilitate and support the definition of the structural features of a meaningful ML-infused idea by combining a technical perspective (input data, ML capability represented by a ML agent, and expected output) with a human-centered one, articulating in three main phases. First, the *concept framing* phase required the definition of the problem in terms of aim, context, and target audience. Then, the *concept exploration* space equipped by a deck of multi-level cards (providing guidance and inspiration) allowed the connection of technical possibilities to a general intent and a more specific design action aimed at helping people achieve their goals. Finally, the *concept definition* space encouraged the participants to better explain how the elements previously identified could actually benefit their intended users. This tool was used for all workshops indistinctly, introduced by a collective example of use to show the underlying mechanics and purposes. All design students pointed out the effectiveness of the tool in helping them envision a meaningful experience. However, very few of them recognized "creating value for people" as its main strength. Instead, they mostly appreciated it for its guidance, procedural, and inspirational value. In their opinion, the VDE — meant to promote ethical reflections — played a crucial role in encouraging a human-centered approach. In fact, they openly said that it fostered the envisioning of problems otherwise missed, and that it was eye opening in exploring different perspectives and making decisions.

None of the ML-related students shared explicit comments about the tools effectiveness in conveying a human centered approach. However, as anticipated, some inferences can be drawn by the activities the CBB encouraged. Throughout the three phases presented, it clearly pushed the participants out of the familiarity of the processes they were used to and toward considering other people's points of view more than their own. Both problem framing and defining user-oriented objectives for their ML-infused artifacts were challenging exercises that required some attempts before being consistently aligned with the requests. Eventually, most of the participants seemed to get the principles behind these activities, and even if some groups also succeeded in accomplish them, it was patent that a five-day workshop was not enough for the students to fully interiorize them. Further iterations would be needed. At least, though, it managed to raise awareness and appreciation for human centricity.

5.2. ML-Savvy

Contents. The weight and significance of the synthesis of ML theoretical knowledge for meaningful applications are evidently different for the two workshop typologies. Anyways, the same contents were provided with different tones and scopes. As previously depicted, ML systems were explained by characterizing them as agents and focusing on their practical capabilities instead of more technical functioning specifications. Additionally, basic information such as definitions, differences with traditional programming and demystification of the most common myths complemented the knowledge transfer.

Design students had no or little prior knowledge on the subject. Therefore, the offered introduction represented the first approach to the technical discipline for many of them. It was essential that the language was clear and that the contents could relate to something familiar for them. For this reason, different examples were provided in support of the structured theoretical explanations and elicited also from the class. From the researchers' perspective and the final outcomes, design students seemed to have grasped the foundational principles of what ML is, what it is not, and how it can be exploited for meaningful causes. However, they suggested that even more examples and case studies were presented for better comprehension. Nonetheless, in both the bachelor and master level workshops, students showed a great interest toward the topic and the graduated ones, who intentionally decided to participate to the educational activity, were also more active in making consistent interventions and asking questions. Overall, the covered arguments proved to be sufficient for making ML a new design material. Still, on the last day some basic ML concepts were weaker in students' minds than on the first day, when they had been introduced. This might be the consequence of the shift of attention toward further design aspects there were more familiar to them, but it might also imply that more practice and exercise requiring the application of ML knowledge could be useful.

Engineering students, on the other hand, were required to have prior knowledge and experience with ML systems. Therefore, the approach was much more interactive. The students were prompted to anticipate the contents that were then explained from the researcher's perspective, and this includes the definition of ML capabilities based on a couple of exemplar applications that the participants could interact with before providing their interpretation. This modality favored the conditions for opening debates about the topic and allowed the author to get in touch with the audience's point of view. As she could only know from literature reviews and reported testimonies, the engineering students involved in the workshops seemed quite unaware of the practical implications and possibilities offered by ML, being more focused on the technical details underneath the algorithms. Also, none of them could provide a complete definition of ML, mostly outlining its statistical nature and capability to improve over time. Indeed, they clearly stated that they were not used to reason about ML systems as presented by the author, yet they were able to draw interesting insights from it. Surprisingly, the fuzzy conception of AI as a technology able to imitate human reasoning was also frequent. In both workshops some fruitful discussions emerged to better specify the capabilities of ML systems or related examples starting from what the researcher proposed, even though the class of W4 appeared less interested in paying attention to what they thought they already knew. Eventually, the researcher's synthesis of ML knowledge proved more necessary than expected to guide engineering students toward meaningful applications of the technology, but they were particularly quick and proficient in applying the lessons learned.

Finally, in both workshop typologies, the researcher and the participants agreed that more practical experimentations with the technology would have been utterly beneficial to have a deeper understanding of the design material and to acquire a more complete ML-savvy.

Methods. To help sediment the basic understanding of ML, not only the theoretical presentations were very interactive, but they were complemented by simple formative tests to reiterate the foundational concepts by both eliciting a response and explaining which answers were correct and why. They gained much more success than expected among design students, who attributed a great part of their comprehension of the subject to these questions and, especially, to making mistakes. The researcher equally assumed that the formative tests could be too easy for her more technical audience, but it really was not the case. While the majority of the class correctly answered, a good number of respondents also indicated wrong options, giving the chance to fellow students to clarify the underlying motivations.

Anyway, the focal method for assessing and crystallizing the participants' ML-savvy was its practical application into a meaningful design challenge. Ultimately, all the ML-infused concepts were intendedly consistent with the ML Agent selected. Though, some design students required further explanations for a couple of capabilities (namely, action selection and regression), which were also rarely considered for implementation. Instead, engineering students also proved capable of suggesting coherent alternatives to achieve the goals depicted by their colleagues during the peer review sessions.

Tools. Two main ingredients were considered to convey ML-savvy: basic ML knowledge and practical experimentation, although only at a conceptual level. For these, two different tools were designed to support educational activities. The former was fostered by ML Agents, the latter by the CBB.

Both design and engineering students did not consult the ML Agents booklet during the design process, primarily relying on their previous introduction by the researcher. The questions synthesizing the principle beneath each capability seemed useful to facilitate the comprehension of the related ML Agent. Yet, in the end, the name of the capability was the most referred to in the next stages, therefore it was important for them to be directly intelligible. The visual representations of the ML Agents had a limited impact on students' understanding. Indeed, a design student in W1 suggested the implementation of animated ones. The examples provided to complement the definition of ML capabilities worked fine, especially with engineering students who were asked to use them to deduce the ML task. In both classes, the recognition of a capability was not immediate, but this reversed process gave the occasion to discuss and depict differences. However, even more examples would have been appreciated by both target audiences.

Regarding the CBB, all design students agreed on the effectiveness of the tool to help envision a meaningful ML-infused artifact. The input — ML capability (ML Agent) — output structure was straightforward regardless of familiarity with the subject. Many affirmed that the guidance provided by the tool gave them the confidence to experiment in the ideation phase, like ML systems were common design materials. Indeed, all the groups delivered concepts that consistently integrated ML. However, besides the coherent application of this technology to serve a predefined purpose, a further step to handover their ideas to their technical counterparts, in a slightly more detailed and effective way was missing.

This also emerged in W4, as some students noted that the CBB tool was missing a more technical part they could better relate to. Indeed, an additional level of more operative knowledge was outlined by the researcher in the theoretical framework beneath the construction of the CBB tool, the ML Designerly Taxonomy (Sciannamè, 2023), but it was not yet included in the tool. Overall, the participants with a technical background appreciated the tool and, more specifically, the modality of thinking to the purpose of the technological solution that it fostered, because it was something they were not used to. Additionally, all the groups in the W3 were grateful for the physical nature of the tool, as it gave them the possibility to design without their laptops, incentivizing discussion and divergent thinking. Although following the tool as a guide in their creative process instead of jumping to a possible solution right away was more of a challenge for engineering students, the CBB ultimately was a helpful support for questioning and modifying initial ideas, allowing them to look at the project from different angles. Thus, despite some difficulties in approaching the design process differently and considering new variables (like the intent and design actions beyond the technical core of the solution), ML-savvy resulted a complementary way to think of ML even for those who already possessed technical knowledge.

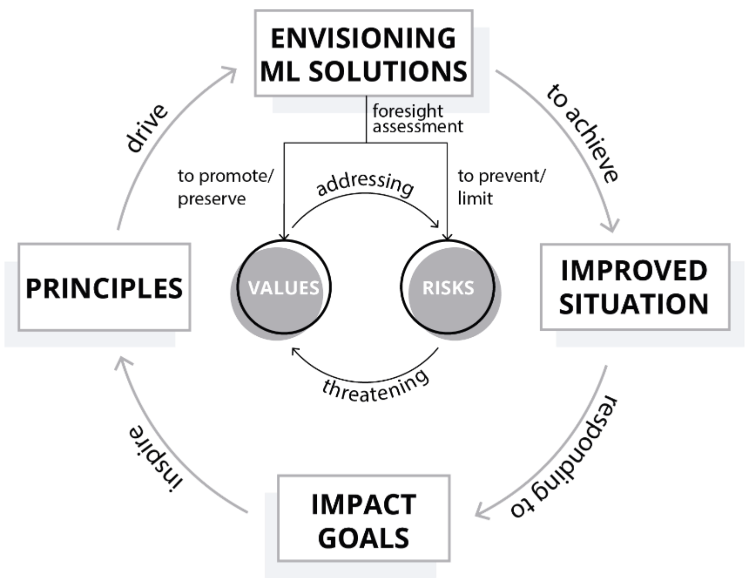
5.3. Value Sensitivity

Contents. As depicted above, two main concepts were at the base of transferring value sensitivity: (i) ML systems as special kinds of sociotechnical systems (van de Poel, 2020), and (ii) embedding values in the design process (van den Hoven, 2013). To convey them, all the workshops presented basic principles of RRI, the definition of sociotechnical systems in relation to ML, and a focus on value-sensitive design. To underline the ethical and social implications of the design of ML systems, and because W3 and W4 were longer, they also introduced the concepts of nudging and the non-neutrality of design and technology, case studies representing questionable and beneficial applications of ML capabilities, and an extended explanation of the value sensitive design process in relation to common designerly ones. Having an entire morning dedicated to this topic resulted in sessions planned to be very much participative, where students were necessarily required to critically reflect on the arising issues, providing examples and personal sense-making. Despite being the same, the lectures for the technical workshops encountered very different responses. In W3, the students intervened with many pertinent and thoughtful comments and questions, demonstrating interest and critical reasoning; in W4, active participation was poor. Once again, the students indicated that augmenting the number of practical examples might engage them more, limiting theoretical concepts to the bare minimum.

For design students, who practically demonstrated their implicit orientation toward responsible solutions but lacked formal preparation on ethical issues, the theoretical presentation was much more limited, and the transferring of value sensitivity was deferred to the design activity. Indeed, because of their background, they did not need to understand the importance of ethics principles as many of them overlap with common design practices. However, the discussion of debatable and positive case studies would have enriched their educational experience.

Methods. Based on a systematic analysis of the main AI ethical guidelines collected in the online repository Algorithmic Watch (2020), a *Responsible Cycle for ML Design* (Figure 5) was outlined by the author (Sciannamè, 2023) to encourage a responsible approach toward the design of ML-infused solutions. The focal implied assumption, included and tested in all the workshop, established that, in order to educate humane ML experts, the translation should frame essential ethical concerns and incentivize reflection-in-action, without providing predefined solutions. Indeed, the available guidelines mainly portray principles and values to overcome or limit unexpected or undesirable implications of AI and ML systems. They rarely suggest possible remedies to prevent or limit negative impacts. And when they do, these highly resonate with common human-centered design approaches. For this reason, while ethics can be a fruitful source for unfolding critical insights throughout the design process, the development of valuable solutions should be a designers' responsibility. As this was implemented through the VDE, the considerations about this approach are depicted in the following point.

Figure 5: Concept Building Blocks and VALUable by Design Expansion.



Source: Developed by the author.

A characterizing difference between the workshops involving design and ML-related students regards the actual introduction and operationalization of value sensitivity. As mentioned, the ethical aspects of designing with and for ML capabilities are the opening argumentation of the educational path for engineering students. For this, the starting point of their design process is the identification of a great challenge, one of the Sustainable Development Goals (SDGs), to help them frame a more specific problem to tackle (aim, context, and target audience) with a principle in mind from the very beginning. Design students, instead, were required to frame and define their idea starting from personal intuition, so that the researcher could assess the human centricity of their approach, only based on the design brief and no predefined processes. Only after the first draft, they were asked to reframe their problem, aiming at an SDG and including an ethical principle as main reference for the development, and iterating their idea accordingly. Both methods were adequate for their intended targets. In the first case, as framing the design problem in a human-centered way was already a novelty for engineering students, setting high-level ethical boundaries helped them intuitively orient the concept generation toward relevant challenges. Indeed, they appreciated having a significant context to start with, even if identifying a more specific problem within their topic required a little effort. For design students, identifying a valuable problem to address was more natural. At the beginning they needed to focus on the consistency with the technological possibilities they were dealing with for the first time, and adding too many variables would have been counterproductive. The subsequent inclusion of ethics principles was quite seamless, as most of the ideas would perfectly fit or were implicitly inspired by SDGs. This method reflected the processes that they, more or less implicitly, usually follow.

Tools. The VDE tool was developed to explicitly guide the participants to intentionally and explicitly introduce a value-driven approach to the design process. It encompasses the already mentioned SDGs and ethics principles to inform problem framing, and risks and values (all in card form) to consider in impact anticipation and foresight activities at a concept generation level. Once an idea was firstly drafted, the focus shifted on identifying the possible issues that might emerge and finding value-inspired ways to prevent or limit them from happening. In general, not having predetermined solutions to draw from did not seem a problem. Indeed, all the groups, regardless of their background, could find mitigating measures to address the *reality check* cards (depicting possible risks) quite easily. Even if no particularly innovative solutions emerged, the participants were free and managed to look at multiple aspects for addressing their issues.

This facilitated anticipatory exercise was certainly unusual for engineering students, especially because it required them to find solutions beyond the technical system and embracing aspects like communication, user experience, or social behavior. Although at first it was difficult for them not to concentrate only on datasets and algorithms, after the researchers' feedback, they understood how to holistically reason about the sociotechnical system they were envisioning. In these workshops, the researcher observed how the value cards played a crucial role in inspiring implementable solutions to prevent or limit possible risks. Moreover, having a role / personality to impersonate brought to light very interesting insights about the values of the different stakeholders, especially during peer review sessions. In the end, engineering students admitted it was a fun experience and, even if they do not know how much of this value sensitive approach can be directly included in their day-to-day practice as it is framed today, they surely recognized its importance.

The design students attending the workshops found the VDE an intuitive and enjoyable tool, suitable as reflection starter, able to provide a general overview and "perfect to start from scratch," but also a synthetic tool "that really put everything together." Its capability to give direction to the design process was particularly valued. As well, "introducing more perspectives," "pushing ML Heroes [how their concepts were defined in the design brief] further" and "making [them] sustainable and acceptable products" are among the main qualities that the students reported. Overall, the researcher's assumptions for triggering value sensitive solutions were largely confirmed by the comments and final outputs. Especially, the explicitness and intentionality of a value sensitive approach proved useful also in designerly contexts. In fact, although design students are more inclined to look for valuable solutions for people, this does not automatically correspond to actual responsible outcomes, as the lack of deliberate reasoning can easily lead to some key features being overlooked. This was underlined by a student who appreciated the VDE "because it forces you to be critical on your work," translating abstract concepts into practice.

6. Q3: Understanding the Suitability and Meaningfulness of the Promoted Skills

No objective measure can determine the suitability and meaningfulness of the proposed skills for humane ML experts only based on the presented educational activities. Nonetheless, to understand whether these skills were appropriately tuned for their target audiences, it is possible to draw some qualitative inferences considering three main factors: the consistency and quality of students' interventions throughout the workshops, their ability to accomplish the different design tasks, and the final outputs. Regarding the meaningfulness of these skills, students' feedback is the only available source for the argumentation. Certainly, this implies that the reported results are context-specific, and their generalization cannot be assured.

The experimentations in design institutions (different from the researcher's one) confirmed that a human-centered approach is at the core of design education and can be considered a fundamental prerequisite of designers. Thus, a project-based activity, with the broad brief to envision meaningful ML solutions and no related theoretical contents, was sufficient to convey human centricity to design students. Differently, the participants with an engineering background needed some theoretical introduction to justify the process and activities required. Abstract theories and frameworks were not enough, while examples were the most significant means to favor the comprehension of the peculiarity of human-centered perspective. During the workshops, the students could recognize that engineers and designers have different mindsets, distinguishing their strengths and pain points. Some curiosity was also raised to better understand the differences, overlaps or possible integrations between the agile method and the design thinking process. Additionally, after initial difficulties and a few attempts, the groups of engineering students managed to grasp how to frame a human-centered problem and develop a concept accordingly. Eventually, all their ML-infused solutions showed some empathy toward the users or people otherwise impacted (also because of the value-driven specifications), demonstrating a marked change of attitude and, in some cases, even deeper reflections than in the projects of their design counterparts. Finally, it is complicated to assess how meaningful human centricity is for design students, as it can be considered an identity mark for them, who almost take it for granted. Greater importance was recognized by engineering students, who reported that they benefited from the exposure to a different way of approaching a project.

In current times, the significance of having basic knowledge about ML is undisputed. How this could be framed to be practice-oriented was the focal point of the exploration. The workshops highlighted that the way in which ML knowledge was synthesized and communicated to foster tech-savvy for the envisioning of meaningful ML applications suited both the participants from a designerly and a ML-related background. Indeed, it represented the first formal approach to the topic for the former, but also offered a complementary perspective to the latter. As depicted above, the contents, their form and language were appropriate for both novices and more expert audiences, specifically with the aim of finding relevant ways to apply this technology to real-world problems. In all cases examples provided by both the researcher and the class were essential and they should have a central role when designing educational activities for all kinds of humane ML experts. All the participants could structure technologically consistent ML-infused solutions, even if restricting them to portray one prominent capability (for the sake of simplicity and time) can be perceived as limiting. Probably, giving space for the complexity of these systems to be acknowledged and implemented would enrich their understanding and awareness.

The incorporation of ethical principles into the design process through a value-driven approach represented a novel experience for all participants. Echoing earlier points, the use of examples to provoke questioning and reflection proved to be an effective means. Nonetheless, the hands-on activity was pivotal. It revealed the transversal nature of this skill, making it easily attainable for diverse target audiences. Indeed, universal values, easily relatable to everyone, required minimal explanations but necessitated the right guidance to ensure their conscious and intentional integration. Still, a couple of issues need attention. On the one hand, the workshops primarily focused on concept generation. This is why the envisioned solutions provided broad value-based indications rather than detailed ones, thus possibly overlooking viability and feasibility. On the one hand, participants across all workshops acknowledged how this approach could enhance design perspectives, fostering increased thoughtfulness. Yet, an engineering student raised a noteworthy concern, highlighting the potential conflict between the undeniable importance of embedding values in ML-infused solution design and the practical constraints faced by employees in real-world scenarios.

7. Discussion

Before delving into identifying the implications that could be inferred by the results of the presented experimentations, it is crucial to emphasize that they were set in safe educational environments. Therefore, the effectiveness of the activities might not have an immediate transposition in the professional field, as they offer a partial perspective on the actual development of ML-infused solutions. For instance, the most pragmatic constraints are not included, such as the actual communication possibilities among different stakeholders and organizational structures (Rakova et al., 2021). Nevertheless, preliminary insights to inform some essential traits for future professionals involved in the development of ML-infused artifacts can be drawn. Hopefully, they might inspire further experiments in business contexts, for more comprehensive analyses.

7.1. The Converging Pathways of Design and Engineering Students

How design and engineering students coped with the workshop activities does not differ significantly. Indeed, the influence of the promoted skills can be detected in the final outcomes. All the concepts showcase meaningful use of ML systems, fairly aligned with their capabilities. They address problems that affect people at personal, societal, or environmental level and include measures to enhance their ethical acceptability. Of course, for optimal results, the learning path should extend over a single learning experience, including more iterations, practical experimentations, or even contaminating current academic programs.

The diverse backgrounds of the participants did not compromise skill transfer but influenced the educational strategy (contents and approach to the design activity), and their performance in specific tasks. Expectedly, engineering students initially struggled to focus on people-centered problems, being used to a more straightforward declination of solutions. Additionally, they did not expect to explore multiple ideas instead of optimizing their first option, while this is the common way designers operate. On the contrary, the groups with a technical background could find suitable applications for ML systems, demonstrating a deeper understanding of their requirements and complexity. Instead, their designerly counterparts limited the reflections about the technology to the essential elements elicited. In the end, the participants from both disciplinary areas managed to reach the goal of envisioning meaningful ML solutions and understood the value of the contents and approach proposed — those having a broader expertise with ML more than the others. Still, the demonstrated disciplinary strengths would perfectly complement the weaknesses of the other and emphasizing them in a collaborative perspective could benefit everyone.

7.2. Future Humane ML Experts Bridging Disciplinary Gaps

The synthesis of needed competences, based on the deficiencies of current ML applications, highlights the importance of the conjunction of different disciplinary perspectives for preparing future ML experts. Moreover, they represent pivotal features that anyone engaged in the design of ML-infused solutions should and can be aware of, regardless of their educational background. Their disciplinary expertise, however, should be valorized. Therefore, in the domain of the design of products and services integrating ML, humane ML experts should maintain the specificity of the roles of programmers and designers, but in a context of close collaboration. Thus, they should be able to effectively communicate and work together, sharing enough knowledge and skills to comprehend the challenges, the potential, and the limitations of their counterparts. Then, while the technical ML experts would be in charge of materializing ideas and making things work, the designerly ML experts should focus on delivering value and meaning to people. But only together they could make sense of what ML can and should do, leveraging human centricity, ML-savvy, and value sensitivity.

8. Conclusion

In the evolving industry of AI, the research explores fruitful skills for humane ML experts. Human centricity, ML-savvy, and value sensitivity emerged from the thematic analysis of recent studies, guidelines, and toolkits. These were implemented and tested in four experimental workshops, targeting design and engineering students as pivotal contributors to the future creation of meaningful ML-infused artifacts. The workshops aimed at (Q1) understanding how to translate the presented skills into educational experiences; (Q2) assessing the effectiveness of the experimentations to foster such competences; and (Q3) evaluating their suitability and meaningfulness.

Despite the preliminary and qualitative nature of the investigation, positive results could be observed. The experimented skills for humane ML experts demonstrated to be transversal and useful for the two educational contexts. In particular, the explicit orientation towards meaningful and responsible solutions, with an attentive initial problem framing, an encouraged value-driven approach, and a non-technical starting point (focusing instead on ML capabilities) proved effective with both target audiences. Designers were introduced a new material to design with. Engineers were exposed to a new mindset.

However, being set in the educational context, the study does not encompass the complexities of business organizations, including the optimization of time and resources. Therefore, while educational activities have value for formative purposes and to foster the skills for humane ML experts, they do not reflect the working reality. Instead, they suggest a viable possibility for the two professional figures involved to effectively communicate and understand each other, valuing their complementary competencies and sharing the same scopes.

Finally, the research intends to bring a qualitative contribution to the ongoing conversation about the prospective competences of professionals in the emerging RAI field, paving the way to further inquiries. Potential areas for expansion include evaluating the long-term influence of the workshops on students' academic and professional careers. Additionally, the scope can be broadened to finding innovative approaches to translate the suggested skills in both educational and professional settings, or to critically discuss and build upon them. In any case, furthering the exploration of the essential characteristics of future AI-related professionals is indeed a first step toward ensuring a more responsible and meaningful future.

References

Algorithmic Watch. (2020, April). *AI Ethics Guidelines Global Inventory*. https://inventory.algorithmwatch.org/

Antonelli, P. (2018, August 2). AI Is Design's Latest Material. *Google Design Library*. https://design.google/library/ai-designs-latest-material/

Archer, B. (1995). The Nature of Research. *Co-Design: Interdisciplinary Journal of Design*, 6–13.

Auernhammer, J., & Ford, C. (2022). What is design? The semantic core and periphery of design definition. In G. Goldschmidt & E. Tarazi (Eds.), *Expanding the frontiers of design: A blessing or a curse?* (pp. 24–37). DesignTech, Faculty of Architecture & Town Planning TechnionIsrael Institute of Technology.

Bresler, L. (2021). Ethnography, Phenotllenology And Action Research In Music Education. *Visions of Research in Music Education, 16(17)*. https://opencommons.uconn.edu/cgi/viewcontent.cgi?article=1808&context=vrme

Cutler, A., Pribić, M., & Humphrey, L. (n.d.). Everyday Ethics for Artificial Intelligence. *IBM*. ibm.biz/everydayethics

Davis, M. (n.d.). Introduction to Design Futures. *AIGA | the Professional Association for Design*. https://www.aiga.org/aiga-design-futures/introduction-to-design-futures/

De Bono, E. (2016). *Six thinking hats* (Revised and updated). Penguin Life, an imprint of Penguin books.

Dove, G., & Fayard, A.-L. (2020). Monsters, Metaphors, and Machine Learning. https://www.semanticscholar.org/paper/Monsters%2C-Metaphors%2C-and-Machine-Learning-Dove/4b1ea05c83d44b984db8b7d5764ef306d602dc35

Dove, G., Halskov, K., Forlizzi, J., & Zimmerman, J. (2017). UX Design Innovation: Challenges for Working with Machine Learning as a Design Material. *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, 278–288. https://doi.org/10.1145/3025453.3025739

European Commission. (2018). *Artificial Intelligence for Europe*.

Frascara, J. (2020). Design Education, Training, and the Broad Picture: Eight Experts Respond to a Few Questions. *She Ji: The Journal of Design, Economics, and Innovation, 6(1)*, 106–117. https://doi.org/10.1016/j.sheji.2019.12.003

Friedman, B., & Hendry, D. (2019). *Value sensitive design: Shaping technology with moral imagination*. The MIT Press.

Friedman, K., Lou, Y., Norman, D., Stappers, P. J., Voûte, E., & Whitney, P. (2014, December 6). Why DesignX? Designers and Complex Systems. *Core77*. https://www.core77.com/posts/27986/Why-DesignX-Designers-and-Complex-Systems

Friedman, K., Lou, Y., Norman, D., Stappers, P. J., Voûte, E., & Whitney, P. (2019, August 1). *DesignX: A Future Path for Design*. *Jnd.Org*. https://jnd.org/designx_a_future_path_for_design/

Fruchter, N., & Liccardi, I. (2018). Consumer Attitudes Towards Privacy and Security in Home Assistants. *Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems*, 1–6. https://doi.org/10.1145/3170427.3188448

Gagné, R. M., Briggs, L. J., & Wager, W. W. (1992). *Principles of instructional design* (4th ed.). Harcourt Brace Jovanovich College Publ.

High-Level Expert Group on Artificial Intelligence. (2019a). *A definition of AI: Main capabilities and scientific disciplines*. European Commission. https://ec.europa.eu/digital-single-market/en/news/definition-artificial-intelligence-main-capabilities-and-scientific-disciplines

High-Level Expert Group on Artificial Intelligence. (2019b). *Ethics Guidelines for Trustworthy AI*. European Commission. https://ec.europa.eu/futurium/en/ai-alliance-consultation/guidelines

Johnson, D. G., & Verdicchio, M. (2017). Reframing AI Discourse. *Minds and Machines*. https://doi.org/10.1007/s11023-017-9417-6

Kirschener, P., & Norman, D. A. (2021). The Teaching and Learning of Design. *Future of Design Education*. https://assets.website-files.com/6010c2ec764764887cfefee7/60515ff38234e90fac4c64cb_Kirschner%20and%20Norman.%20Teaching%20and%20Learning%20of%20Design%2021.01.21.pdf

Koskinen, I., Zimmerman, J., Binder, T., Redström, J., & Wensveen, S. (2011). *Design Research Through Practice: From the Lab, Field, and Showroom*. Morgan Kaufmann.

Kulesz, O. (2018). *Culture, platforms and machines: The impact of Artificial Intelligence on the diversity of cultural expressions* (DCE/18/12.IGC/INF.4) [Information Document]. UNESCO. https://en.unesco.org/creativity/sites/creativity/files/12igc_inf4_en.pdf

Latour, B. (1996). On actor-network theory: A few clarifications. *Soziale Welt, 47(4)*, 369–381.

Law, J. (2015). *STS as Method*. http://heterogeneities.net/publications/Law2015STSAsMethod.pdf

Levinson, P. (1977). Toy, Mirror, and Art: The Metamorphosis of Technological Culture. *Learning Cyberspace, 34(2)*, 151–167.

Meyer, M. (2010). The Rise of the Knowledge Broker. *Science Communication, 32(1)*, 118–127. https://doi.org/10.1177/1075547009359797

Meyer, M. W., & Norman, D. (2020). Changing Design Education for the 21st Century. *She Ji: The Journal of Design, Economics, and Innovation, 6(1)*, 13–49. https://doi.org/10.1016/j.sheji.2019.12.002

Norman, D. (2023). Humanity-Centered versus Human-Centered Design. In *Design for a Better World: How to create a meaningful, sustainable, and humanity-centered future*. MIT Press. https://jnd.org/humanity-centered-versus-human-centered-design/

Rakova, B., Yang, J., Cramer, H., & Chowdhury, R. (2021). Where Responsible AI meets Reality: Practitioner Perspectives on Enablers for Shifting Organizational Practices. *Proceedings of the ACM on Human-Computer Interaction, 5(CSCW1)*, 7:1-7:23. https://doi.org/10.1145/3449081

Rismani, S., & Moon, Aj. (2023). What does it mean to be a responsible AI practitioner: An ontology of roles and skills. *Proceedings of the 2023 AAAI/ACM Conference on AI, Ethics, and Society*, 584–595. https://doi.org/10.1145/3600211.3604702

Robson, C., & McCartan, K. (2015). *Real World Research. A Resource for Users of Social Research Methods in Applied Settings* (4th ed.). Wiley Blackwell. https://www.wiley.com/en-ie/Real+World+Research%2C+4th+Edition-p-9781118745236

Russell, S., & Norvig, P. (2020). *Artificial Intelligence: A Modern Approach* (4th ed.). Pearson.

Sancassani, S., Brambilla, F., Casiraghi, D., & Marenghi, P. (2019). *Progettare l'innovazione didattica*. Pearson.

Schön, D. A. (1983). *The reflective practitioner: How professionals think in action* (Reprinted). Ashgate.

Sciannamè, M. (2023). *Machine Learning (for) Design. Towards designerly ways to translate ML for design education* [Ph.D. Thesis, Politecnico di Milano, Department of Design]. http://hdl.handle.net/10589/204613

Umbrello, S., & van de Poel, I. (2021). Mapping value sensitive design onto AI for social good principles. *AI and Ethics, 1(3)*, 283–296. https://doi.org/10.1007/s43681-021-00038-3

van de Poel, I. (2020). Embedding Values in Artificial Intelligence (AI) Systems. *Minds and Machines, 30(3)*, 385–409. https://doi.org/10.1007/s11023-020-09537-4

van den Hoven, J. (2013). Value Sensitive Design and Responsible Innovation. In R. Owen, J. Bessant, & M. Heintz (Eds.), *Responsible Innovation* (pp. 75–83). John Wiley & Sons, Ltd. https://doi.org/10.1002/9781118551424.ch4

Varanasi, R. A., & Goyal, N. (2023). "It is currently hodgepodge": Examining AI/ML Practitioners' Challenges during Co-production of Responsible AI Values. *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*, 1–17. https://doi.org/10.1145/3544548.3580903

von Schomberg, R. (2013). A Vision of Responsible Research and Innovation. In R. Owen, J. Bessant, & M. Heintz (Eds.), *Responsible Innovation* (pp. 51–74). John Wiley & Sons, Ltd. https://doi.org/10.1002/9781118551424.ch3

Voûte, E., Stappers, P. J., Giaccardi, E., Mooij, S., & van Boeijen, A. (2020). Innovating a Large Design Education Program at a University of Technology. *She Ji: The Journal of Design, Economics, and Innovation, 6(1)*, 50–66. https://doi.org/10.1016/j.sheji.2019.12.001

Wang, Q., Madaio, M., Kane, S., Kapania, S., Terry, M., & Wilcox, L. (2023). Designing Responsible AI: Adaptations of UX Practice to Meet Responsible AI Challenges. *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*, 1–16. https://doi.org/10.1145/3544548.3581278

World Economic Forum Global Future Council on Human & Rights 2016–18. (2018). *How to Prevent Discriminatory Outcomes in Machine Learning*. https://www3.weforum.org/docs/WEF_40065_White_Paper_How_to_Prevent_Discriminatory_Outcomes_in_Machine_Learning.pdf

Yang, Q. (2020). *Profiling Artificial Intelligence as a Material for User Experience Design* [Carnegie Mellon University]. http://reports-archive.adm.cs.cmu.edu/anon/hcii/abstracts/20-100.html

Yang, Q., Banovic, N., & Zimmerman, J. (2018). Mapping Machine Learning Advances from HCI Research to Reveal Starting Places for Design Innovation. *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems — CHI '18*, 1–11. https://doi.org/10.1145/3173574.3173704

Yang, Q., Scuito, A., Zimmerman, J., Forlizzi, J., & Steinfeld, A. (2018). Investigating How Experienced UX Designers Effectively Work with Machine Learning. *Proceedings of the 2018 Designing Interactive Systems Conference*, 585–596. https://doi.org/10.1145/3196709.3196730

Zdanowska, S., & Taylor, A. S. (2022). A study of UX practitioners roles in designing real-world, enterprise ML systems. *CHI Conference on Human Factors in Computing Systems*, 1–15. https://doi.org/10.1145/3491102.3517607

Ethical Statement

Conflict of Interest: Nothing to declare. **Funding:** Nothing to declare. **Peer Review:** Double-blind peer review.