MACHINE LEARNING (for)DESIGN

Towards designerly ways to translate ML for design education.

Martina Sciannamè

MACHINE LEARNING (FOR) DESIGN.

Towards designerly ways to translate ML for design education

Author
Martina Sciannamè

Supervisor
Davide Spallazzo

Co-supervisor
Mauro Ceconello

External examiners
Frédérique Krupa | L'école de design Nantes Atlantique | France
Seçil Uğur Yavuz | Free University of Bozen-Bolzano | Italy

PhD Programme in Design, XXXV cycle
Politecnico di Milano | Department of Design

Coordinator
Lucia Rosa Elena Rampino

Thesis deposit | *May 2023*

Defense | *June 2023*

POLITECNICO MILANO 1863
DIPARTIMENTO DI DESIGN

DOTTORATO DI RICERCA IN DESIGN

Table of contents

Abstract

Thanks to their empathy, system-level thinking, and transformative influence, designers have the potential to bring substantial changes in the way ML systems are designed, developed, and deployed in products and services. Yet, as they still lack ML-related knowledge, language, skills, and competencies, they are little or not involved in these processes.

Therefore, the doctoral research points to design education as a means for empowering designers to benefit both the ML field and society by conceptualizing meaningful solutions integrating ML systems in full respect of human values and needs as part of broader ecosystems.

Based on a designerly translation of ML knowledge within an ethical frame, theoretical systematizations, educational models, tools, and a method synthesizing them are proposed to enable designers to handle ML as a design asset to address different challenges and, ultimately, to participate in interdisciplinary teams to develop ML-enhanced solutions.

The argumentation retraces the exploratory and constructivist journey that, between theory, practice, and reflection – following an action research process and with a research-through design spirit – brought to the theoretical and practical contributions for the translation of ML (for) Design at the dawn of a field of burgeoning opportunities.

To all those who have supported and helped me on this challenging and troubled journey, a huge and heartfelt THANK YOU!

i. INTRODUCTION

Such productive thinking can be an everyday activity as well as employed to solve open and complex problems, which requires diverse domain-specific knowledge.

(Auernhammer & Ford, 2022)

I N T R O C H A P T E R

For an easier understanding and reading of the thesis, this Chapter frames the research by outlining the motivation and the identified **relevance** of the addressed topic (i.1). It introduces the pursued **objectives** and the **results** providing theoretical and practical **contributions** to scientific knowledge (i.2), as well as the **research nature and methodology** (i.3). Finally, it offers an **overview** of the chapters and the structure of the dissertation (i.4).

i

i.1 Rationale. Among disciplinary deficiencies and relevance

Artificial intelligence (AI) is not a discovery of our time. It dates back to the 1940s, and ever since, it has experienced exciting periods of disruptive visions coming to life, and winters in which innovative ideas could not be materialized, bringing the hype around this technology back to a silent state.

Recent technological developments have dramatically increased computational power and the possibility to retrieve and store data, creating the perfect conditions for machine learning (ML) to thrive, becoming the most important subset of AI and opening up new possibilities. This powerful technology is going to **affect numerous aspects of life** on Earth in ways that we cannot yet fully anticipate, and with a revolutionary force comparable to electricity (Kelly, 2016). However, with power comes responsibility. Currently, the development of products and services integrating ML is **led by technical experts**, often moved by the urge to unlock new technological capabilities or following the marketing objectives of their funders while missing the broader perspectives rooted in the socio-technical nature of these systems (Antonelli, 2018; van de Poel, 2020; Yang et al., 2018). The problem is acknowledged in both computer science and ethics fields, and many initiatives and guidelines emerged to counter the dangers of the reckless advancement of AI and ML, moved by experts from the public and private sectors (Algorithmic Watch, 2020). Among the most commonly suggested measures to responsibly manage such an influential technology that is still being studied, diversity in development teams is strongly encouraged (Cutler et al., n.d.; High-Level Expert Group on Artificial Intelligence, 2019; World Economic Forum Global Future Council on Human & Rights 2016-18, 2018).

However, in this domain, **designers are still marginally or not involved** (Dove et al., 2017), even though the recent history of the evolution of personal computers and the Internet shows that their empathy, system-level thinking, and transformative influence (Dalsgaard, 2017; Frascara, 2020) would allow them to bring beneficial contributions to the ML field and society at large. They have the potential to make sense of ML in innovative and more sensitive ways and to find new interpretations of what can be valuable to people (Norman & Verganti, 2014) through the conceptualization and materialization of meaningful solutions integrating ML systems. Though, designers are **not yet prepared** for this objective and have difficulties recognizing ML as a new material to include in their projects (Antonelli, 2018; Stoimenova & Price, 2020; Yang,

2020). Indeed, they lack ML-related knowledge, language, competencies, tools, and methods to deal with it (Dove et al., 2017; Meyer & Norman, 2020; Yang et al., 2020).
Therefore, **design education** appears to be a perfect context of intervention to start filling this gap and contribute to steering the development of ML towards beneficial impacts, a pressing matter that is still under-explored in the design discipline. This is also consistent with the strategic vision of the European Commission to disseminate AI knowledge and promote trustworthy results and with the principles that should drive design education for the 21st century. Specifically, interdisciplinary approaches that draw on technological knowledge are strongly encouraged, as this is a critical asset to cope with our changing environment (Findeli, 2001; Friedman et al., 2019; Meyer & Norman, 2020; Voûte et al., 2020).

i.2 Research objectives and scientific contribution

In order to enable (future) designers to exploit ML as a resource for their designerly interventions and to help them envision meaningful applications for this technology, the research aims to **identify and translate ML basic knowledge to raise design students' awareness** (main research question). On the one side, it focuses on making it accessible and operationalizable to them, understanding how this technology works, its possibilities, and limitations from a theoretical standpoint. On the other, it includes ethical concerns to enhance a responsible approach, reinforce and augment the **systemic and holistic** perspective of design students when addressing a problem, maintaining human life and its ecosystems at the center. Because of these premises, the doctoral research is positioned at the intersection of UX/interaction design, ML, HCI, and computer ethics.
Specifically, the investigation starts from the foundations of the ML discipline and related ethics to infer what are the essential concepts that can and must be translated to design in order to favor their comprehension and implementation (RQ1). Then, it concentrates on the modalities, forms, and language to actually transfer knowledge (RQ2), and it synthesizes the preliminary hypothesis into theoretical assumptions and constructs. To accomplish the translation, the research envisions, develops, and validates flexible and modular models and tools for different design educational contexts (RQ3). Finally, based on the experimental experiences and gained insights, it formulates a possible structure for an educational method to merge technical, ethical, and designerly knowledge and support the conceptualization of meaningful ML-infused solutions (RQ4).
In pursuit of these objectives, the investigation produced original contributions to knowledge at different levels. From a theoretical point of view, it resulted in the systematization of essential ML knowledge to bridge technical and human-centered perspectives on its implementation (*ML Designerly Taxonomy*) and in the identification and procedural framing of the foundational elements for a responsible ML design process *(Responsible Cycle for ML Design*). To apply these concepts in a hands-on educational environment, some tools were developed to transfer basic knowledge (*ML Agents*) and give procedural support and inspiration for envisioning meaningful ML-infused artifacts (*Concept Building Blocks* and *VALUable by Design Expansion*).

Educational models and an overarching method have been outlined to test and frame them.
Hopefully, the research outputs represent starting points to enhance interdisciplinary communication and collaboration toward addressing relevant challenges and impacts.

i.3 Research approach

Due to the uncharted territory under investigation (at least from a designerly perspective), the research is highly **exploratory**. It moves across ML and ethics to gain a better understanding of the subject matter and to identify the key ingredients to empower (future) designers to handle this technology as one of their design materials. With no solid foundations to lean on, the research follows an action-research process. It starts from wide-ranging qualitative inquiries and leverages field research methods (Koskinen et al., 2011) to test and investigate hypotheses from direct observation and participation. To infer actionable information, though, a significant role is played by reflective and interpretative activities. Indeed, it is highly iterative, and its development unfolds with a cyclic structure. The research questions are addressed in a **constructive progression**: each is characterized by four phases (planning, acting, observing, and reflecting) and creates the premises on which the next can be built. They are increasingly comprehensive in scope, moving from basic aspects of the translation to broader and more articulated constructs and, in the end, all the findings (tools, models, and theoretical assumptions) converge in the definition of an educational method merging design, ML, and ethics for design students to learn how to integrate and take advantage of ML capabilities to envision responsible and meaningful solutions.
Overall, it can be said that the investigation assumes a research-through-design spirit as this kind of research process actually corresponds to a design process (Swann, 2002).

i.4 Thesis overview & structure

The thesis follows the constructive progression of the investigation, subsequently portraying how the planning, acting, observing, and reflecting research activities led to answering each research question and established the foundations for further inquiry (addressed by the following research question until the time limits of the Ph.D. path).
As visible in Table 0.1, the argumentation offers five types of entry points that can guide readers in choosing the chapter that most interests them: (i) the chapter; (ii) the research question; (iii) the research activities portrayed; (iv) the tools developed and described; (v) and the main outcomes achieved.
Expectedly, the body of the research is represented and ordered according to the research questions. In addition, two introductory chapters and a conclusive one better frame the study.
To facilitate the navigation of the contents, each chapter has an introductory paragraph, highlighting how the topics develop throughout the sections, and a conclusive sum up of the key points of the argumentation. Both summaries help make the connections across chapters and between the research activities more explicit.

Research question	Chapter	Research action(s)	Main outcome(s)	Tool(s)
/	1. A designerly look to machine learning. Describing the research context and gap	Exploratory literature review	• Problem framing: foundations for the multidisciplinary work • Research questions	/
/	2. Methodology	/	• Identification of research approach and methods	/
RQ1: What can and has to be translated from ML to Design?	3. ML in translation. Substance and boundaries	• Literature review • Desktop research • Case studies analysis and auto-ethnography	• Core concepts to understand and translate ML	• *Encoder & Decoder* • *ML Suitability Matrix*
RQ2: How to frame ML knowledge for transfer?	4. Framing ML knowledge for transfer	• Theoretical synthesis • Workshop (ML Pills for Designers) • Content analysis of AI ethical guidelines • Expert interviews	• ML Designerly Taxonomy • *Responsible Cycle for ML Design* • Foundational assumptions for the educational experience	• *ML Agents*
RQ3: How can the theoretical constructs be operationalized into models and tools to be implemented and tested in educational contexts?	5. Concretizing hypotheses in models and tools to support design education	• Theoretical synthesis • Workshop and focus group (Introductory Game to ML Responsible Design) • Workshop (Superpowered Museums)	• Identification of the requirements for the educational models and method • Educational approaches and tools	• *Concept Building Blocks* • *VALUable by Design Expansion*
RQ4: Which design education method can support the conceptualization of ML-infused solutions?	6. Towards an educational method to frame models and tools	• Literature review • Workhops (ML Hero Agency, VALUable ML Heroes, VALUable ML Hero Agency) • Expert interviews	• Educational models (consistency, responsibility, integrated) • Educational method	/
/	7. Setting out a long journey ahead	/	Identification of: • contribution to knowledge • limitations • future steps • further research opportunities	/

Tab. i.1 | Thesis structure.

> Chapter 1 presents the research context from a designerly perspective and the mutual benefits that might emerge from bridging ML and design in an ethical frame. It also identifies the research gap, objectives and explains the research questions.

> Chapter 2 describes the characterizing nature of the research, illustrates the adopted methodology, and provides an overview of the entire process based on the planned actions and expected results.

> Chapter 3 addresses RQ1. Identifying which elements can and must be translated from ML to design, it includes an exploration of the technical knowledge, revealing inadequate narrative and contents for a direct transfer. This oriented the research towards a case studies analysis of ML outreach strategies that resulted in the identification of the most relevant elements to explain ML to non-experts, synthesized in the *Encoder* and *Decoder* tools and the *ML Suitability Matrix*.

> Chapter 4 has the purpose of finding ways to frame ML knowledge for transfer (RQ2). First, it is systematized in the *ML Designerly Taxonomy*, creating the basis for (interdisciplinary) communication and the implementation of ML systems to address real-world problems. To test the related assumptions and to explore the preferable forms and language for the translation, it reports the experience of the *ML Pills for Designers* workshop. Its findings and insights are synthesized in the *ML Agents* tool to transfer ML basic knowledge, and they inspired a content analysis of ethical guidelines for AI to complete the framing from an ethical perspective. The inquiry resulted in the definition of principles, values, and risks as basic components of a *Responsible Cycle for ML Design*.

> Chapter 5 focuses on elaborating models and tools to operationalize the previously obtained outcomes (RQ3). In particular, it investigates modalities to provide holistic educational experiences that can be flexible and modular to adapt to different contexts and necessities. Coherently, it portrays the development of the *Concept Building Blocks* tool and its *VALUable by Design Expansion* (aimed to support the concept development of consistent and responsible ML solutions) and two practical experimentations (the *Introductory Game to ML Responsible Design* and the *Superpowered Museums* workshop) to outline the requirements for an educational method.

> Chapter 6 finally outlines the research activities that led to the definition of the interdisciplinary educational method to support design students to envision meaningful ML solutions. In particular, it illustrates the construction of consistency, responsibility, and integrated educational models to synthesize and test all the previous assumptions, which were validated according to an evaluation research protocol of four workshops held in different European universities as case studies.

> **Chapter 7** gives a conclusive, informed overview of the research and its findings, including the contribution it can give to design education and beyond, its limitations, and the possible future steps and research opportunities.

1

1. A DESIGNERLY LOOK TO MACHINE LEARNING. DESCRIBING THE RESEARCH CONTEXT AND GAP

Without art and design and culture, any technology is bound to either fall flat or not really capture the wholeness of humanity.

(Antonelli, 2018)

C H A P T E R O N E

The argumentation opens with an **overview of ML**, aiming to shed light on aspects not usually evident in the computer science narrative. It portrays the main results of an exploratory literature review intended to investigate ML as a design issue, and to demonstrate these fields' mutual relevance.

After explaining the rationale behind the subject of interest (1.1) and the adopted method (1.2), the disciplinary affinities between ML and design are explored from a theoretical point of view (1.3) and in the practice of developing solutions integrating ML (1.4), highlighting different points of contact and shortcomings. Subsequently, the focus shifts to foundational design elements as keys to understanding ML, like interaction and experience (1.5), ethical concerns (1.6), and a systemic perspective (1.7). To conclude, a convergence is sought (1.8) investigating designers' potential to positively contribute to ML systems development and the educational context as a strategic means to spread ML knowledge and open possibilities for responsible innovation.
Overall, it is argued that envisioning solutions that integrate ML systems needs more humanity-oriented perspectives, and this challenge is consistent with designers' roles and capabilities. These premises finally set the groundwork for framing the research and indicate its direction (1.9), synthesized in the main question – ***How to translate ML basic knowledge for design students?*** – and the related sub-questions.

1

1.1 Why machine learning?

Even though the limitation to this specific disciplinary target came later in the research process, it is important to clarify, from the very beginning, why the investigation focuses on ML and does not consider AI in its entirety as design material. This explanation will give a general idea of the relevance of ML without going into the details of definitions (described in Chapter 3).
First of all, AI is a broad field, encompassing multiple approaches, applications, and sub-fields. While other branches of AI focus on logic, probabilistic reasoning, or knowledge representation (closely relating to statistics and data science), ML capabilities entail autonomous process definition to achieve a goal and adaptation to changing contexts. These are game-changing qualities bringing novelties to current services and products. Indeed, ML is the subset that has benefited most from recent technological developments, which have allowed its hitherto only theorized potential to flourish.
The fact that ML is the most common expression of AI at the moment was demonstrated already in 2019, at the beginning of this research, by the AI Index Report – a yearly publication edited by the Stanford Institute for Human-Centered

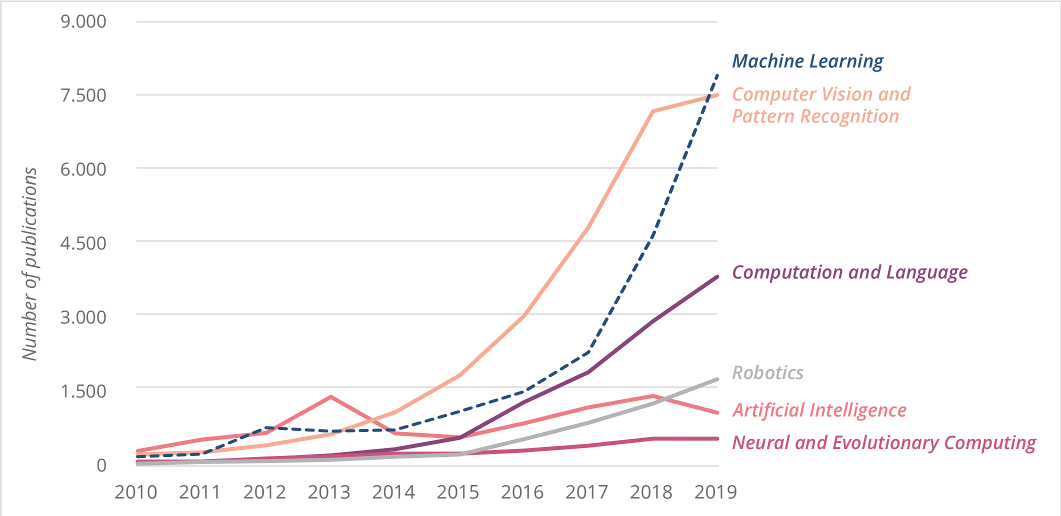


Fig. 1.1 | Number of AI-related publications on arXiv from 2010 to 2019, showing the relevance of ML publications. Source: AI Index Report (Perrault et al., 2019).

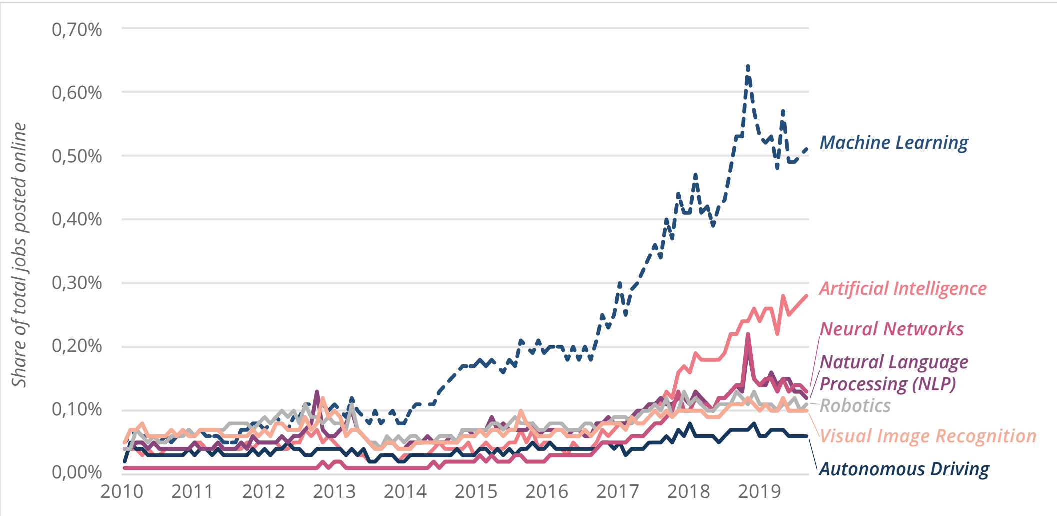


Fig. 1.2 | Share of total AI jobs posted online in the USA from 2010 to 2019. Source: AI Index Report (Perrault et al., 2019).

Artificial Intelligence (HAI) to provide unbiased, carefully examined, and worldwide sourced AI-related data. Perrault et al. (2019) show that both research (Fig. 1.1) and job offers (Fig. 1.2) have a great interest in ML among AI sub-fields. This remains mostly unvaried in the latest AI Index Report (Zhang et al., 2022), with the only difference being that pattern recognition counts more publications than ML. However, it is indeed a capability enabled by ML and – its subset – deep learning (DL) systems, more than an actual field.

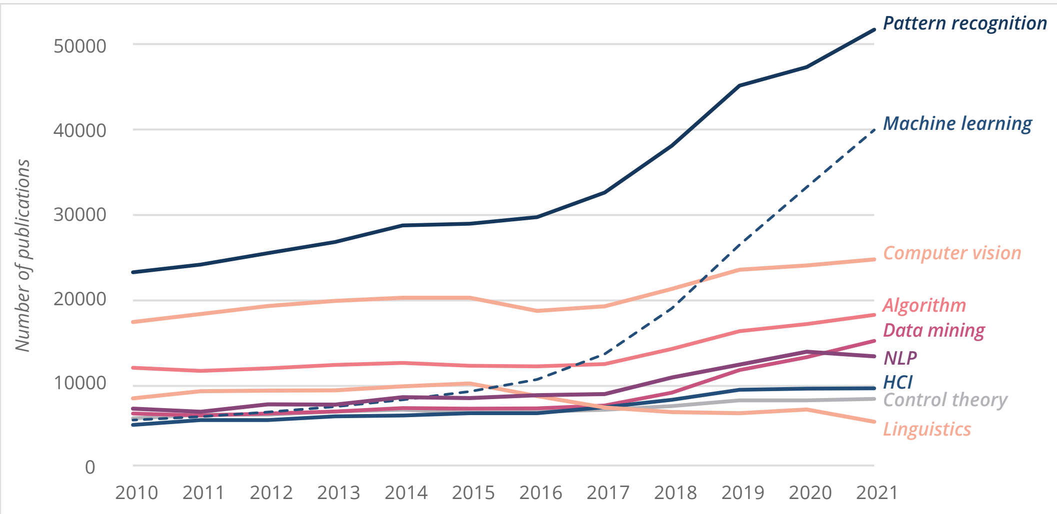


Fig. 1.3 | Number of AI-related publications from 2010 to 2021, where ML is second only to pattern recognition. Source: AI Index Report (Zhang et al., 2022).

Additionally, the worldwide reference textbook about AI, *Artificial Intelligence: A Modern Approach,* has recently released an update with important deepening and attention on ML as, since the previous edition (2010), the authors observed *"an increased availability of data, computing resources, and new algorithms"* (Russell & Norvig, 2020).
Also within the design field, literature specifically addresses ML as a *new design material* (Dove et al., 2017; Yang, 2018). In fact, ML systems directly impact human-computer interactions, changing the meanings that products and services can assume.

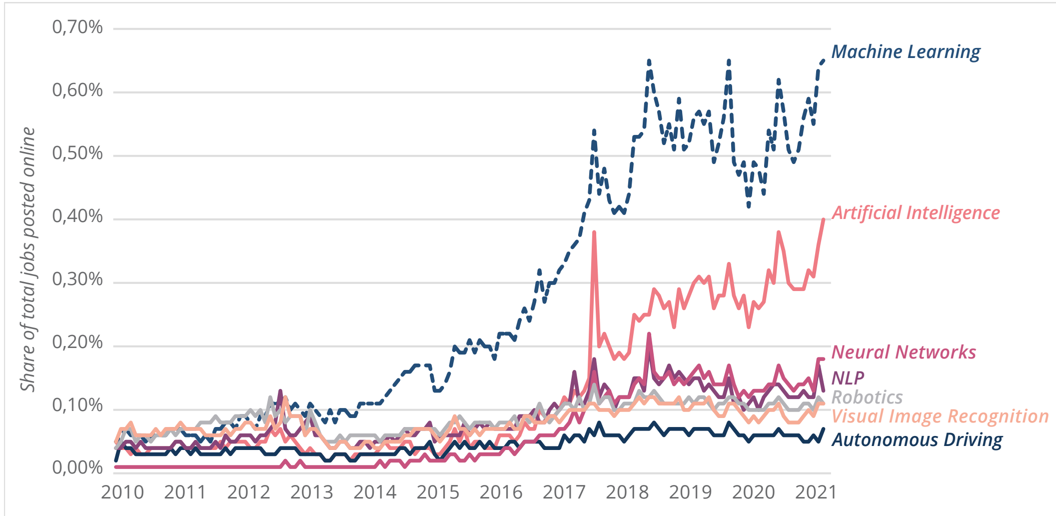


Fig. 1.4 | Share of total AI jobs posted online in the USA from 2010 to 2021. Source: AI Index Report (Zhang et al., 2022).

However, in the first exploratory part of the research, no restrictions were applied to the literature review, both because of the limited resources available and because the distinction could be operated after having defined whether the source applied to ML. Indeed, AI is often used to refer to ML capabilities, as if the terms were interchangeable.

1.2 Exploring ML from the design lens. Describing the research method

Aiming to investigate the relationship between design and AI (ML in particular), the first step was to conduct a literature review. For this purpose, a systematic approach based on the PRISMA model (Page et al., 2021) was attempted. The query was conducted on three databases – Scopus, Web of Science (WOS), and ACM Digital Library – to maximize the potential reach. It was initially limited to papers from 2010 to 2020 (when the research was conducted) to have up-to-date results. To begin with, they were searched from the databases according to their title, abstract, and keywords, using the entries: "design*" and "AI" or "artificial intelligence" or "ML" or "machine learning." Scopus returned 137 089 results, WOS 73 839, and ACM 73 527. This first exploratory attempt revealed that design is not a significant keyword and produced a high number of false positives. In fact, it was often included in the papers as a verb, largely interpreted as the act of designing AI and ML algorithms and thus spreading in all kinds of technical contributions. To reduce the sample, the search was filtered by language (English), type of document (conference proceedings and journal articles), and subject area, but this only uncovered further limitations. Design (or anything similar) is not listed among the possible subject areas, and the only relevant ones were computer science and engineering. The results decreased to 66 218 in Scopus, 21 562 in WOS, and 70 418 in ACM (since the only publication-type filter was available among those selected). Still, these were unmanageable amounts of documents, and it was not possible to filter out all the technical papers describing the development of algorithms.

Clearly, the term design needed better circumscribing to identify the correct reference context. Several attempts were made, including using UX and interaction as keywords. However, the outputs were still too generic as also these terms are easily interpreted in ways that do not concern the field of design.
The final strategy was to introduce a third element in the search to include possible objects of the investigation in addition to the disciplinary areas. "ML-infused" and "Human-ML interaction" concepts, with related alternatives, were added to the query, and satisfying results were obtained, with no need to limit the timeframe as the publications dated back maximum to 2018 on Scopus and WOS and to 2011 on ACM.
The search:

TITLE-ABS-KEY ("design material" OR "design matter" OR "UX Design" OR "User Experience Design" OR "Interaction Design") AND ("AI" OR "artificial intelligence" OR "ML" OR "machine learning") AND ("human-AI interaction" OR "human-ML interaction" OR "AI-infused" OR "ML-infused" OR "AI-enhanced" OR "ML-enhanced")

produced 17 results in Scopus, 7 in WOS, and 32 in ACM. The latter, though, needed a little refinement to exclude 9 full proceedings (resulting positive only because of the conference presentation), one tutorial, and one poster, for a total of 21 documents to process.
Excluding the duplicates, the initial set of results counted 38 papers . Of them, 15 were excluded based on their title and abstract, and 15 after reading the full papers because they were:

- vertical on niche applications of AI or ML systems somehow related to design concerns (e.g., conversational interactions, VR/AR applications, recommendation systems of social media platforms, or clinical decision support systems);
- still too much related to the technical development of algorithms, models, or interfaces to clarify their functioning to programmers;
- or because they were general calls for workshops (citing the relevant papers identified in this search when talking about the subject of interest).

Two could not be retrieved for full paper review, but the abstracts indicated they would have been of marginal interest to the research.
Ultimately, four relevant papers were identified as considerably relevant for the inquiry, and two more were kept in consideration for specific issues, as summarized in Tab. 1.1.
Therefore, not enough literature was retrieved to provide direction for the study, but some of the most influential authors working on the topic were identified in this way – namely (Dove et al., 2017; Yang, Sciuto, et al., 2018; Amershi et al., 2019; Dove & Fayard, 2020). The situation inevitably called for an exploratory approach, and the qualitative literature review was then based on a snowball sampling inquiry (Goodman, 1961), a technique that comes from sociology and statistics but can also be applied to systematic reviews. Starting from the identified set of relevant papers, the reference lists were carefully examined to find publications consistent with the research topic. Excluding already analyzed papers and those explicitly referring to other subjects, the remaining abstracts or full texts were reviewed for new insights.

	Title	Year	Authors	Affiliations	Source
Seminal references	***UX Design Innovation: Challenges for Working with Machine Learning as a Design Material***	2017	Graham Dove Kim Halskov Jodi Forlizzi John Zimmerman	• ***Aarhus University***, Denmark; • ***Carnegie Mellon University***, Pittsburgh, PA, USA	Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems
	Investigating how experienced UX designers effectively work with ML	2018	Qian Yang Alex Scuito John Zimmerman Jodi Forlizzi Aaron Steinfeld	• ***Carnegie Mellon University***, Pittsburgh, PA, USA	DIS 2018 - Proceedings of 2018 Designing Interactive Systems Conference, pp. 585-596
	Guidelines for human-AI interaction	2019	S. Amershi D. Weld M. Vorvoreanu A. Fourney B. Nushi P. Collisson J. Suh S. Iqbal P.N. Bennett K. Inkpen J. Teevan R. Kikin-Gil E. Horvitz	• ***University of Washington***, Seattle, WA, USA; • ***Microsoft***, Redmond, WA, USA	Conference on Human Factors in Computing Systems - Proceedings
	Monsters, Metaphors, and Machine Learning	2020	Dove G. Fayard A.-L.	• ***New York*** University, NY, USA	Conference on Human Factors in Computing Systems - Proceedings 3376275
References for specific issues	***Cognitive Mimetics for Designing Intelligent Technologies***	2018	Kujala T. Saariluoma P.	• ***University of Jyväskylä***, Finland	Advances in Human-Computer Interaction, 2018,9215863
	Replay enactments: Exploring possible futures through historical data	2020	Kenneth Holstein Erik Harpstead Rebecca Gulotta Jodi Forlizzi	• ***Carnegie Mellon University***, Pittsburgh, PA, USA	DIS 2020 - Proceedings of the 2020 ACM Designing Interactive Systems Conference, pp. 1509-1522

Tab. 1.1 | Set of references from a systematic literature review based on the PRISMA model.

The same process was applied to the newly included sources until no useful unread article could be found.
Overall, at the time of the study, the exploratory review confirmed that few contributions built theory or reported on previous research about the relations between AI and design. Rerunning the search at the end of the doctoral path demonstrated that the interest in the matter is rising, along with the number of resulting publications: 70 on Scopus, 19 on WOS, and 69 on ACM. Even though the construction of the theoretical premises presented in the following sections is based on the initial research, the insights from some relevant and more recent publications are also integrated.
Therefore, the inquiry tried to identify and focus on specific facets to develop hypotheses and research questions constructively.
In the following, the results of this process are depicted. The argumentation opens with an overview of ML, aiming to shed light on aspects not usually evident in the computer science narrative but that are key to framing it as a design issue and demonstrating the mutual relevance these fields have to each other. This finally led to identifying the research gap, objectives, and questions.

1.3 Approaching the world. Intertwined roots

Apparently, the discipline of design and AI could not be more distant. To most people's eyes, it might seem like the former is characterized by creative impulse and closeness to the humanities, while the latter by scientific rigor and mathematical logic. Therefore, parallels between the way designers and ML systems think and operate might come as a surprise while hinting at intriguing possibilities.
In this paragraph, the research explores the disciplinary affinities that design and ML have, highlighting different points of contact: from common theoretical origins to similar recognized skills and conditions in which to work.

1.3.1 Common theoretical background

One of the first theorists of the design process and design as a discipline was Herbert Simon in 1969, in his seminal book *The Sciences of the Artificial*. He is known for the famous definition of design as a devised action to change existing situations into preferred ones, and he claimed ownership of a *theory of design* and a *design curriculum* in the science of design (Simon, 1996). He felt the necessity to bring the design activity to the forefront as a practical application of natural sciences and a key competence for all professional (or, better, applied) disciplines like engineering, architecture, business, education, law, or medicine. In fact, he believed that while natural sciences try to explain *how* natural *things are*, design is about human-made artifacts (that constitute the greatest part of the world we live in) and *how they should be* to reach some goals. He also managed to make design an academically recognizable discipline and to have a Design Research Center founded at Carnegie Mellon in 1975. He needed to turn design's traditionally *soft*, *intuitive, informal, and cookbooky* character into a *tough*, *analytic, formalizable, and teachable* one (Simon, 1996, p. 112).
Even though his position might be arguable, he affirmed that design had been accepted in academia because of the need to include computers, and it had to be explicitly and precisely theorized (Simon, 1996). What he referred to when talking about computers was actually artificial intelligence that, along with human intelligence and the design process, represented a way to comprehend and deal with complex systems: the ultimate scope of the book. In his argumentation, then, he traces a clear connection between AI and design, making their roots inevitably intertwined.
Later on, the definition of design evolved, always struggling to find a place between sciences, arts, and technology. However, Simon's contribution was foundational in shaping its identity and marking its relevance. It left an important legacy in design research (Huppatz, 2015), which would not have existed without AI. Indeed, some parallels are still undeniable today and could bring interesting insights.

1.3.2 How designers and AI systems think and operate

At the beginning of his book, Simon (1996) recognizes the artificial world as a worthy subject of study and claims the necessity of a science of the artificial to make it more understandable and straightforward. Fascinated by the complexity of the environments in which humans, as simple *"behaving systems,"* live and are affected (Simon, 1996, p. 80), he identified the design process as the key to adapting one's means to the *outer* environment, in order to attain a goal. Essential for this activity is problem framing and solving. Indeed, Simon describes the steps to reach an objective in a complex, uncertain, and changing environment as understanding, learning, and finding new problem representations. Of course, this applies both to designers and computer programs (AI and ML systems).
Much of this argumentation resonates with pillar theories in the design discipline. The complex and ambiguous context of intervention is what calls for indeterminate, *wicked* problems (Rittel & Webber, 1973). According to Buchanan (1992), they characterize design problems because this discipline has no specific subject matter but could be universally applied to any area of human experience, the artificial world in Simon's words. Indeed, while there is no convergence in the definition of the design discipline, methods, or philosophy, design theorists shared an interest in *"the conception and planning of the artificial"* (Buchanan, 1992). An activity that could not rely on any scientific knowledge – not natural, nor social, nor humanistic – to address the wicked problems that design, as a liberal art in his perspective, deals with. Although they are not easy to describe, Rittel and Webber (1973) observe that these problems cannot be addressed following a linear path, a characteristic that also distinguishes ML from traditional programming. Peter Norvig stresses this aspect in the *Introduction to ML Crash Course* offered by Google (2018). He differentiates the logical and mathematical way of addressing a problem of a software engineer and the shift of focus that ML requires. In his view, this resembles natural science experiments that start from observing an uncertain world and continue with testing hypotheses and analyzing the results using statistics instead of logic.
The debate around the scientific character of design – embraced by Simon (1996) and confuted by many, as You and Hands (2019) well depicted – is still open and beyond the scope of this dissertation. However, another crucial analogy between design and AI is that both can be configured as a rational solving process, and the wicked problems they tackle always coincide with the definition of a solution. Also in Rittel and Webber's (1973) perspective, the design process divides into two phases:

an analytic problem definition and a synthetic problem solution. This *solution-focused* approach to dealing with ill-defined problems is what Cross (1982) defined as *designerly ways of knowing*, and it might be argued whether this also applies to AI and ML systems.
Then, a major feature they share probably lies in their purposes, as both tend to improve people's lives by making them easier, more pleasant, or better in some way.
Other affinities can be detected when detailing designers' ways of tackling problems. For instance, the second form of designers' abductive reasoning that Dorst (2011) describes involves defining a value to achieve with no specifications about what has to be built or which working principle should be applied. Similarly, ML systems are only given a goal to reach, and their task is determining how to do so. In this attempt, ML systems commonly work through examples and the identification of patterns, as well as designers are trained to learn from examples (Simon, 1996), to construct patterns – speaking the language of the artificial (Cross, 1982; Kun, 2020) – to reason with metaphors to make sense of familiar or unfamiliar situations (Schön, 1983), and to finally uncover possible solutions to a problem. In this respect, it is noteworthy how Simon's and Schön's approaches are usually considered to be at the antipodes but, in practice, share similar expedients (the results of which are then treated differently but tend to the same end).
Uncertainty also calls for flexibility and adaptability, attributes peculiar to ML systems throughout computer science and inherent to the design discipline that, as anticipated, could universally apply to a multitude of contexts and problems. Both are also prone to learning-by-doing in a trial-and-error modality and modifying their actions according to the environmental responses.
Ultimately, it should be remembered that the design curriculum envisioned by Simon (1996) has inspired the formalization of the design process and thinking, as Fig. 1.5, retrieved in (Dam & Siang, 2020), demonstrates. However, this simplification conceals the original AI-based and prescriptive tone that, at the same time, reinforces the AI-design relationship but also attracts much of the subsequent criticism. For instance, Simon (1996) refers to the design curriculum including utility theory and statistical decision theory as ways to frame the problem, algorithms for choosing optimal or, at least, satisfactory alternatives, or factorization and means-ends analysis for heuristic search.

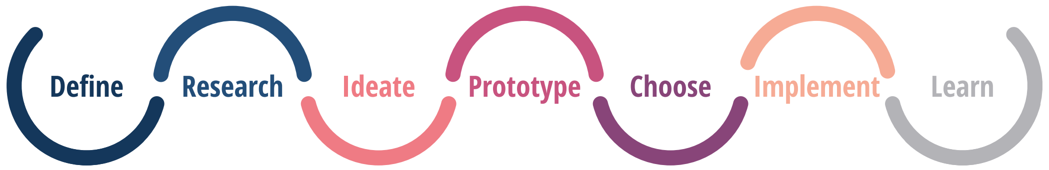


Fig. 1.5 | Design Process based on Simon's elaboration. Adapted from Daniel Skrok and the Interaction Design Foundation, CC BY-NC-SA 3.0.

Perhaps, going beyond Simon's theorization of the design process and using a framework that is built upon and synthesizes the lessons learned by the many design processes that have arisen over the years, it might be easier to portray the similarities that connect designers and ML systems' way of working.
For this purpose, (Fig. 1.6) illustrates an elaboration of the design process compared to other established ones. It makes explicit the often tacit *Problem space exploration*

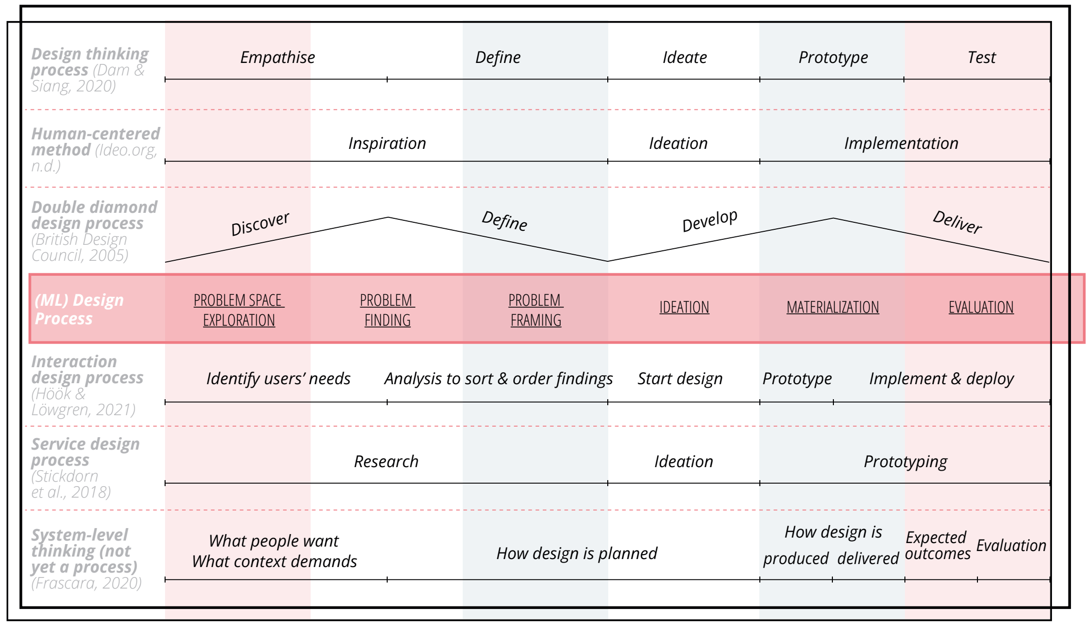


Fig. 1.6 | Reframing of the design process.

and *Evaluation* practices as they have great relevance in the definition of the "right" outcomes and impacts (von Schomberg, 2013); it uses the term *Materialization* to imply both prototyping and deployment as crucial phases to assess; and it distinguishes between *Problem finding* and *Problem framing*, to differentiate the understanding and representation of the problem.
Represented in a circular shape, to underline that the process may not have a clear beginning and is possibly never-ending, the commonly accepted design process can be easily assimilated with ML systems' process, as depicted in (Fig.1.7).
This paragraph attempts to bring the two disciplines closer by presenting the multiple traits that designers and ML systems have in common. Based on their affinities, it would be reasonable to think that the figure of the designer should intrinsically be well equipped to comprehend AI and ML systems. However, current reality shows a

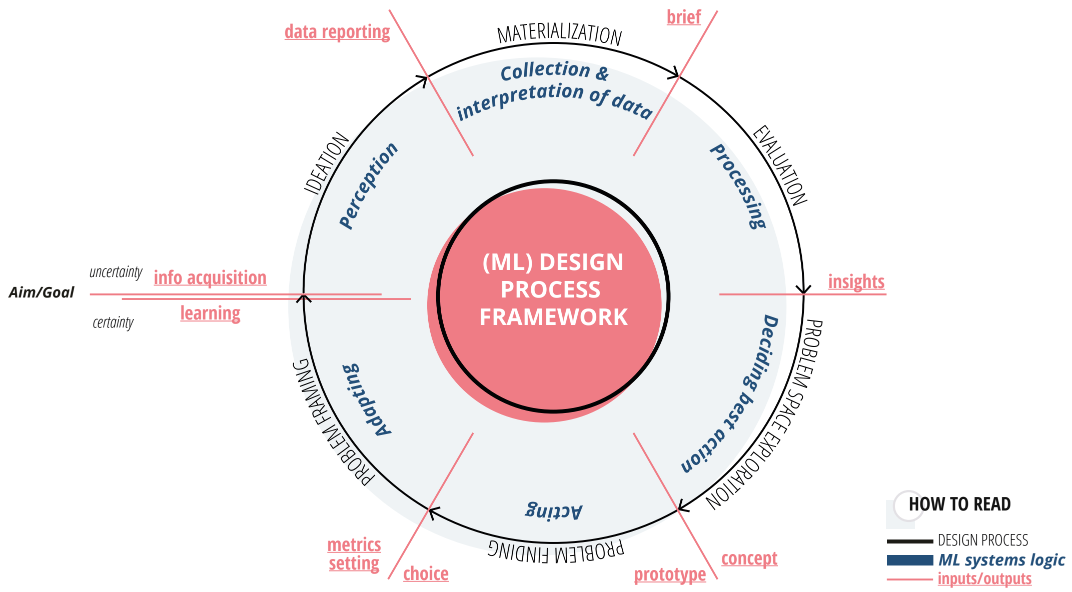


Fig. 1.7 | Combination of the design and ML system's processes.

very different picture, characterized by multi-level distances explored further in this chapter.

1.4 ML and design relations and challenges in the professional field

Moving from theoretical connections to practice, the real-world picture of designing AI or ML systems is still mainly technology driven. As Muratovski (2016) states, big tech corporations (such as IBM, Microsoft, Intel, etc.) do not look to the design field for researchers to investigate on future products and systems development. Instead, they hire PhDs from engineering, anthropology, or psychology backgrounds. If this is the case for economically powerful companies, smaller realities likely follow the same example. Similarly, we can imagine that this also applies to R&D of products and services integrating ML.

A possible indication of this may be the very limited literature investigating ML systems development and the role of design practitioners. To the best of the author's knowledge, at the beginning of this research, two were the prominent articles dealing with this topic (Dove et al., 2017; Yang, Sciuto, et al., 2018), and they have been recently complemented by one additional inquiry (Zdanowska & Taylor, 2022), which reaffirms the scarcity of resources.

Even though they offer useful insights for better framing the relationship between designers and ML, they present quite limited samples. So, the experiences they portray cannot be used to infer general tendencies about designers' roles in current AI and ML systems development. In fact, Dove et al. (2017) based their study on survey responses of 37 UX practitioners (from the US, UK, and Denmark) who worked with ML; Yang, Sciuto et al. (2018) on 13 interviews to design practitioners with at least four-year experience with ML-enhanced experiences; and Zdanowska and Taylor (2022) on semi-structured interviews to two interdisciplinary teams and seven UX practitioners.

It is relevant, in this sense, that Dove et al. (2017) identify ML as an *"under-explored opportunity for ideation and innovation led by UX design,"* despite 63% of the respondents to their survey confirmed to have worked with ML into commercial products or services. Of these 37 respondents, the majority (12) collaborated with ML experts to develop an idea, 8 were just called to give an interactive form to someone else's idea, and in just 7 cases, the concept generation integrating ML was left to the design team. Likewise, Yang, Sciuto, et al. (2018) recognize that no designer followed the project development in all of the stages (especially because ultimate decisions were based on data). Just the few participants who worked in large AI-focused organizations reported experiences in concept development, while only few projects even got to the release and refinement stage. The most likely scenario is that designers join the projects after the operational decisions have been made (Yang, 2018). Zdanowska and Taylor (2022), instead, depict a different picture, in which UX practitioners' importance is valued throughout the entire process (even if having the opportunity to do so is not a given), while ML experts are involved after the idea had been defined. The authors acknowledge a possible reason in the fact that UX practitioners are the target of their study, while the others address interaction and experience designers. To support this, also in light of Muratovski's (2016) assertion, only one of Zdanowska and Taylor's interviewees has a formal education in design (the others come from psychology, arts, humanities, or even natural science backgrounds). Being impossible to determine what is the current engagement of design practitioners in the development of AI/ML-infused artifacts, the common challenges they experience can be highlighted.

Consistently with the lack of ML-related education or training for designers, major issues are (i) superficial knowledge and difficulty in comprehensively understanding ML capabilities. Interestingly, Yang, Sciuto, et al. (2018) note that the interviewed designers developed a *"designer savvy,"* using abstractions to make sense of what ML systems can do in relation to users' utility, and this synthesis is very distant to ML educational materials. However, this did not represent an impairment in the collaboration with ML experts, as both professional figures are trying to figure out the best way to design for AI/ML, and they could do that together. Their relationship is only undermined by the actual (ii) possibility of interacting with them. Likewise, the lack of tools and techniques to handle ML for non-experts has also been overcome. Designers learned the strategies of their technical counterparts and adapted their methods and tools (e.g., data journey maps, service blueprints, conversation maps) to discuss with them (Zdanowska & Taylor, 2022), or to *"embrace data centricity"* (Yang, Sciuto, et al., 2018). In any case, (iii) current HCI methodologies are not feasible for real-world AI/ML projects because they do not capture the specificity and the complexity of this technology, missing how it is going to work (Zdanowska & Taylor, 2022). This clearly reflects in the (iv) prototyping phase. The same researchers state that low-fidelity prototypes, commonly used by designers to test traditional human-computer interactions (like the Wizard of Oz), should be based on in-depth knowledge of the systems' qualities and behavior in order to work (Zdanowska & Taylor, 2022). So, they claim that fast and high-fidelity levels should be reached to properly comprehend the functioning and implication of the imagined solutions. While no real indication can be derived about the best strategy to prototype for such systems, research on the topic is needed (Yang, 2018) as few experimentations are reported in the literature – among them (Van Allen, 2018; Holstein et al., 2020). In relation to a deeper understanding of AI and ML systems, (v) it is also challenging to clearly comprehend mental models before launch and (vi) to anticipate ethical considerations that, so far, rely on designers' empathy. Indeed, designers feel the desire to provide a human-centered perspective (Dove et al., 2017) and responsibility to ensure ethical outcomes (Zdanowska & Taylor, 2022). This is why they advocate good user research and even participatory design strategies, hoping to invert engineers' dominant perspective.

Then, the portrayed experiences of UX designers working in interdisciplinary teams generally revealed that they could figure out how to collaborate with their technical counterparts, especially by observing and understanding how they are used to working with ML and adapting their skills and tools to enter the discourse. Indeed, they can also prove their strategic role in driving the design of AI/ML projects (Zdanowska & Taylor, 2022). Therefore, a critical question about the actual necessity to properly educate and train designers to work with this technology emerges. A possibility, already risen by Yang, Sciuto, et al. (2018), is that professional designers with superficial knowledge about the technology and a self-learning process based on ML experts' approaches can find their place in interdisciplinary teams and provide helpful contributions when dealing with ML systems, but this might not be sufficient for them to unlock innovative solutions.

In fact, the testimonies reveal that UX designers finally adhere to a data-driven approach, or they are just users' advocates. Their successful integration in teams with multiple backgrounds (even if limited to the reported examples) and their accomplishment in bringing users' perspectives in the process of developing ML systems is, per se, a great step to demonstrate designers' predisposition and value in handling not fully explored territories. Yet, Dove et al. (2017) state that it is rare to see UX designers conceiving innovative ideas as they are most commonly requested to *"put lipstick on the pig."* Hence, in the belief that designers' contributions can be pushed towards more disruptive outcomes, it is still interesting to understand how basic ML knowledge might play a role in disclosing their real potential.

1.5 Human-ML interaction and experience

Another core perspective to achieve a more designerly comprehension of ML is through interaction and experience design. According to phenomenology, in fact, the world's experience through the body and situated action can help generate meaning and comprehension (Merleau-Ponty, 1945).

Intuitively, one might say that ML systems' unique and complex nature poses issues to traditional interaction and user experience (UX). They are *"the new UX,"* as Stoimenova and Price (2020) suggest, and new paradigms might be needed. Hassenzahl et al. (2021) define the artifacts integrating this technology as *"otherware,"* emphasizing their role as "others" in the interaction with people. They mean that ML-infused objects and services are no longer just tools to use as an extension of oneself. They have become entities with which to dialogue, collaborate, or delegate something. As both the AI (see section 3.1.1) and computer ethics (see 1.7) fields state, AI and especially ML systems grew to have their own agency, which is unusual for humans to recognize in non-animated artifacts. Although people might be inclined to project personalities in objects as a way to synthesize the qualities and value they perceive in them (Green & Jordan, 2002), with ML-infused products and services, the situation is reversed: the latter portray some kind of behavior that users can only acknowledge. This mainly depends on their inherent capability to *learn* from their physical or digital environment and autonomously adapt the way they work to achieve their goal better (further explanations of these possible misleading concepts can be found in section 3.1.2). The result is that their behavior can be uncertain and unpredictable as it evolves in the relationship with other autonomous entities in a complex environment (Van Allen, 2017), and, over time, users might receive different outputs to the same input.

Thus, it is compulsory to consider these peculiar qualities while designing for experiences and interactions. Indications about how to do it are emerging (Amershi et al., 2019; Giaccardi & Redström, 2020; Yang, 2020).

Often, though, AI and ML systems are just concealed in products or services as new or improved features. In cases like spam filters, recommendation systems (YouTube, Netflix, Spotify, Amazon, etc.), some photo editing tools (Adobe Photoshop), photo organizers (Google and Apple photo apps), self-orienting robots (Roomba), and many other, people are provided with new possibilities for optimization, but they do not know how they function, which might mean that they ignore relevant information. Not being able to fully comprehend how these things work can easily lead to misunderstandings and frustration. This is also the case when AI and ML capabilities are disproportionally advertised, especially when dealing with human-like traits (chatbots, smart speakers). Unclear communication and general unpreparedness of the lay public let people build high expectations, which ultimately get betrayed, like projecting human-human interactions or impossible magical powers on AI and ML.

An example might be the commercial campaign of Google Nest devices that promises to *"make your home a nest"* but actually portrays another limitation peculiar to the experience these products elicit. In an attempt to concretize the vision of ubiquitous computing (Weiser, 1994), they offer to implement digital functionalities directly within the physical world through multiple devices and touchpoints. However, they fail to create a holistic experience, which is instead scattered among different interfaces (physical, digital, and conversational). This produces disappointing interactions and limits the discoverability of the devices. So, people do not explore their true potential and just exploit their basic functions, such as setting alarms or checking the weather forecast (Sciuto et al., 2018; White, 2018).

Overall, AI and ML-infused artifacts have multiple qualities that make them uniquely challenging to design. Their current implementation demonstrates that several factors are not yet leveraged or are just starting to be addressed.

To summarize what to keep in mind and evaluate when designing for human-ML interaction, some of the insights derived from the Meet-AI research project (of which the author was part) can help. Funded by the Design Department of Politecnico di Milano and coordinated by Prof. Davide Spallazzo, it aimed to identify an appropriate method to capture the UX of such entities and ultimately produced an evaluation scale (AIXE).

The dimensions selected for building it (Sciannamè & Zavarrone, 2022; Spallazzo et al., 2022) represent a possible way to synthesize and articulate the complexity of products and services integrating AI and ML capabilities. They include:

- ***Intelligence:*** one of the focal features of AI-infused products, it entails the essential characters of ML systems, such as their ability to learn autonomously from their environment and adapt their behavior over time, as well as proactively take action or propose suggestions to their users being aware of the context in which they are inserted. Additionally, the level of accuracy and understanding they manifest and enable can contribute to the perception of this dimension.
- ***Trustworthiness:*** this is not an internal feature of AI systems, but how much people feel they can trust these artifacts considerably affects the way they interact with them. It can be said that a product is trustworthy when it is personally and socially acceptable and reliable. A major role is played by how data are handled, how much control people have, and how transparent the operations are.
- ***Conversational dimension:*** enabled by Natural Language Processing (NLP) algorithms, this feature is not included in all ML-based systems, but it is

definitely the most revolutionary in terms of interaction modalities. The quality of conversational interfaces mainly lies in the mutual comprehension and development of dialogue.

- ***Meaningfulness:*** being a long-debated issue in the HCI community, it can be configured in terms of understanding (Dourish, 2001) or satisfaction of psychological needs (Hassenzahl et al., 2013). Based on recently developed frameworks (Mekler & Hornbæk, 2019), products and services can be defined as meaningful if they manifest a definite purpose, a personal significance, a shared/cultural significance, if they generate experience, communicate a symbol, or exhibit a temporal quality. Basically, they should not fall into the category of gadgets or toys (Levinson, 1977), as many current AI and ML manifestations do.
- ***Pragmatic dimension:*** representing the qualities of products that support users in achieving their concrete goals, this dimension can include reliability, intuitiveness, helpfulness, or the possibility to customize the behavior of ML agents.
- ***Hedonic dimension:*** classically intended as the aspects that bring pleasure and engagement, it can manifest in artifacts integrating ML capabilities in the form of empathy and proactivity, simulating human comprehension through statistical predictions.

While being commonly relevant in traditional UX evaluations, two dimensions – namely the aesthetic and affective one – have been excluded from the construction of the scale because they resulted too difficult to assess in relation to AI-infused artifacts. This does not imply that they are not relevant. On the contrary, they can represent rich areas of exploration for designers who might find proper ways to define them according to their new design materials, like in the case of Wensveen's (2021) *Aesthetics of Intelligence*.
While this section depicted the challenges and opportunities that arise by looking at ML systems from UX and interaction design perspectives, the potential improvements to the design of artifacts including this technology are not limited to these fields. Further issues can be brought to light if the consideration of the relationship between people and ML systems passes from use to impact.

1.6 Ethical concerns of ML

AI and ML algorithms, models, and systems currently spreading into our everyday life result from a not-so-new technology that has now encountered favorable technical conditions to better express its potential. However, this is not the only reason for its flourishing period. It is thanks to the research investments of big tech companies, universities, and the availability of private and public capital – especially in the US and Chinese contexts – that AI had the possibility to become so pervasive. For this very reason, the purest technological experimentation and market-driven interests have been shaping the technology as we know it, and the discourses about AI long-term effects focus on issues such as how automation can increase wealth and improve industrial processes (Kulesz, 2018).

Concerns about this technology, or at least about how it is currently developed and deployed, naturally arise from human-related disciplines. Specifically, ethics plays an important role in questioning and reframing attempts within the field of AI, as the numerous ethical guidelines emerged in the last few years demonstrate. (Chapter 4 will provide a more in-depth discussion about them.)
As a matter of fact, relatively recently, the European Union has decided to enter the game of global leaders dominating the AI field, rooting its contribution on ethical issues, and trying to fill what is becoming a pressing gap. Indeed, among others, they recognized that *"the way we approach AI will define the world we live in"* (European Commission, 2018). The document, written by the High-Level Expert Group on Artificial Intelligence (2019b), highlights the main issue already in the title: we need ethics to guide us toward AI systems that can be *trustworthy*, implying that today they are not. The reason why this technology does not yet deserve people's trust may lie in two main aspects: (i) AI is still too much opaque, especially for a non-expert public; and (ii) it lacks the inclusion of human factors. Both these issues are causing undesired and unnecessary confusion, uncertainty, frustration, and, ultimately, mistrust, interpreted by the public in terms of *creepiness* (Fruchter & Liccardi, 2018).
In particular, in relation to (i), AI is commonly perceived as too technical and even arcane. Several authors from different disciplines refer that it is often addressed to as a magical or monstrous entity, as it results unfamiliar, unknown, and uncontrollable. Therefore, it produces not only concern but also fear (Antonelli, 2018; Dove & Fayard, 2020; Johnson & Verdicchio, 2017; Kulesz, 2018). This tropos of creating monsters around new technologies we cannot fully understand is recurring in human history, and AI makes no exception. Cinematography is full of films that portray AI-based entities going mad, obsessing, turning against humans, or pursuing agendas that harm life on Earth (e.g., *Matrix*, 1999-2021; *I, Robot,* 2004; *Jexi*, 2019, and many more). What is relevant is the evidence that there are widespread misunderstandings about this technology.
On the one hand, AI systems are not clear because, despite their name and theoretical procedural affinities, they operate in a way that is *"strangely different from how we commonly understand human intelligence"* (Kulesz, 2018) and communication based on a misleading parallel is not really facilitating comprehension by non-experts. On the other hand, a considerable problem of unclarity in terms of responsibility inevitably generates anxiety, and the resulting picture is even worse than the explosion of the Challenger Space Shuttle in 1986. That became an ethical case study, as accountability bounced between different professionals and was finally attributed to poor and unclear communication. In a scenario where ML is involved, responsibility is not only contested among different human parties, but the question also includes the technology itself, bearing its (obscure) decision-making capability. Johnson and Verdicchio (2017) go deeper in these two causes, highlighting that this situation is more precisely related to (a) the notion of *autonomy*, easily suggesting that machines can be out of human control, and to (b) *sociotechnical blindness*, concealing *"the essential role played by humans at every stage of the design and deployment of an AI system"*. The latter, in particular, is such a critical point that the European Commission revised its previous definition of AI, clarifying – from the very beginning – that *"Artificial intelligence (AI) systems are software (and possibly also hardware) systems designed by*

humans" (High-Level Expert Group on Artificial Intelligence, 2019a). It can be stated that people – more or less closely – follow each step of the development of ML systems. Thus, they have the power to affect them.

This lack of precise communication about the huge role human actors play in the development and deployment of AI systems directly links the discourse to the absent or superficial consideration of humans in using these systems, point (ii). Again, in the *Ethics Guidelines for Trustworthy AI,* we find the evidence of this problem, as they need to underline that AI systems have to be *human-centric*, both in a *utilitarian* way – meaning that it is necessary to maximize the benefits AI systems bring and, at the same time, limit their risks – and according to a sort of reversed-stated *reciprocity principle* – *"AI is not an end in itself, but rather a promising means to increase human flourishing"* (High-Level Expert Group on Artificial Intelligence, 2019b) – that implies that humans are the actual end of this technology, and not means to be exploited in the name of technological innovation. So, it becomes patent that *"the technical and the social are irreducibly entangled"* (Dove & Fayard, 2020), and the *technological mediation* of AI systems, in particular, can be highly impactful both on human perception and action. This is why so much attention has to be paid to the development, deployment, and use of this technology, which can be reasonably considered an agent in the mediation of human experience. Consequently, it needs to be regulated according to basic social and interpersonal principles. According to the EU, for instance, AI systems should be lawful, ethical, and robust – both from a technical and social perspective (High-Level Expert Group on Artificial Intelligence, 2019b).

Overall, for AI and ML-based technologies to have a beneficial impact on the world, there are plenty of requirements that can be identified and as many possibilities to implement them. What is needed is a comprehensive reframing of these systems. One that considers different disciplinary standpoints to create new spaces for intervention, as suggested in the following section.

1.7 ML as part of greater systems

Indeed, because of the tight relationships that AI and people share, considering that this technology always serves human purposes and is inevitably intertwined with human activities, another warning emerges from Johnson and Verdicchio's (2017) argumentation, which opens an insightful perspective.

Inspired by theories from the field of Science, Technology, and Social Studies (STS), they claim that AI systems should be interpreted as *sociotechnical ensembles.* It means that they are part of a combination of artifacts, human behavior, social arrangements, and meaning. They make sense in the relationship with a complex reality that they affect in positive or negative ways. Even more specifically, van de Poel (2020) sustains that AI systems are **a particular kind of sociotechnical system**. Like commonly known sociotechnical systems, they include:

- ***technical artifacts***, the products or services enabling some kind of activity;
- ***human agents***, all the people involved, from those who develop the system to the users, and all of those who are directly or indirectly impacted by it;
- and ***institutions***, the rules that people follow, such as moral, societal, and behavioral norms.

What makes AI sociotechnical systems different is the recognition of ***artificial agents***, the embedded AI systems themselves, which follow ***technical norms***, not defined by their own intentions (as they cannot have any) but by causal-physical rules. Acknowledging AI systems' agency is not just an ethical, intellectual exercise but a validation of their nature and capabilities, as Chapter 3 will explain more extensively.

Then, these *re*-definitions entail that it is impossible to develop AI systems with a narrow view of perfecting algorithms and models. Investigations about explainable artificial intelligence or XAI (Confalonieri et al., 2021) and interpretable ML (Molnar, 2019) tackle only a fraction of the ethical concerns around AI systems. For instance, these approaches focus on increasing the transparency of AI and ML systems by finding ways to clearly explain the process they undertake to get an output and allow their human developers or decision-makers to better understand the results' validity. Dealing with ontologies, knowledge graphs, accuracy levels, and all of those parameters that increase the explainability of these systems is, for sure, a strategy to make them more ethically acceptable, but in a very circumscribed way. These helpful investigations mainly take into account the technical artifacts and norms and partially the human agents involved (concentrating on the developers and perhaps the primary technical users, those who need to make decisions based on AI or ML systems' outputs). They do not handle the wider and more complex sociotechnical system into which they are or will be integrated. Acknowledging the easily identifiable trustworthiness problems that AI and ML's opaqueness might cause towards their developers and decision-maker users, they overlook their broader societal repercussions. **What does it mean to introduce these systems in people's everyday life?** Do they need to be informed about their presence? How to communicate the values and limitations that they imply? To address these general questions, the intervention of the computer scientist or software engineer on the algorithms is insufficient. Different kinds of initiatives should be undertaken by acting on people's environments and norms, ensuring the recognition of the multiple stakeholders and factors that constitute the sociotechnical system.

Hence, despite the essential work of AI and ML experts in shaping the future of technology, other professional figures should complement it with a more holistic mindset. Figures that are able to understand and act in a context that is determined by the network of relations within it, as theorized in the Actor-Network-Theory (Latour, 1996), where no clear distinction needs to be made between humans and non-humans, people and technology, science, and nature (Law, 2015), as they are all actors in an interconnected reality. Although it is beyond the purview of this thesis, this concept can consistently be extended to other natural entities that are inevitably impacted by people's artifacts and actions. In them, for example, one might find purpose or different perspectives to devise beneficial solutions that overcome human-centeredness in favor of whole ecosystems.

Then, framing AI and ML systems as sociotechnical systems assumes a deep significance in a designerly perspective of ML. It enables to embrace the complexity of the *artificial* world, in line with Simon's (1996) view, by enlightening multiple directions to pursue a meaningful development of ML systems. Indeed, by recognizing the agency, value exchange, and impact that different actors (or agents) have in relation to each other, the scope of design is expanded to multiple opportunities for intervention that do not rely solely on technical experts.

1.8 ML as a design matter

Looking at ML from non-computer science perspectives, the chapter illustrated that not only designers are not yet participating in the early development of AI systems, but they seem unprepared to effectively leverage their capabilities and foresee opportunities (Yang, 2018). As non-expert practitioners in the field of AI, they are experiencing all the above-mentioned difficulties, and they are not formally introduced to knowledge about AI or ML systems (Meyer & Norman, 2020), nor do they have means to facilitate their comprehension and the design process (Dove et al., 2017). The result is that they are deferring technical understanding to software engineers (Dove & Fayard, 2020), but in an environment where communication and collaboration are possible, as (some) designers have found ways to fit in the conversation.

Additional issues arise in the still poor and uncertain interactions with artifacts integrating the capabilities of this technology, which deployment also causes ethical concerns because a broader, systemic approach should be adopted when dealing with AI and ML systems (Johnson & Verdicchio, 2017; van de Poel, 2020).

In light of the claim that AI is the latest design material (Antonelli, 2018; Stoimenova & Price, 2020; Yang, 2018), the section aims to find a convergence for this multifaceted picture of ML as a design matter. Based on their systematic semantic analysis of design definitions from academics, practitioners, and general understanding, Auernhammer and Ford (2022) identified some traits that characterize design, which might help to frame the significance of this research in relation to the design discipline. A description of the way and reason why design should contribute to ML can be outlined using the intrinsic attributes that Auernhammer and Ford's (2022) study highlighted (in italics). Specifically,

design might help *make sense* and *give meaning* to ML by *crafting new ways* and *processes* to *solve problems* and aim for *purposeful change* in the materialization of the systems, in the disciplinary field, and society at large. Design *activities* within the process of developing ML-infused artifacts can thus be a *means* to satisfy the *needs* of people and their ecosystems (*nature*).

Therefore, in the following, the attention is brought back to the discipline of design, reflecting on the possible roles that designers and the educational context can play in steering ML toward the flourishing of life on Earth.

1.8.1 Designers' potential

Much could be written about the skills and roles designers can play. Yet, the scope of this section is not to make a comprehensive overview but to highlight their relevant qualities in light of the presented ML-related issues.

The current experimental and uncertain phase of ML development and deployment is a perfect target for designers, who are used to dealing with **fluid, ill-defined, and continually changing contexts** (Auernhammer & Ford, 2022; Meyer & Norman, 2020). Thanks to their long-standing relationship with technology, transformative influence, cross-disciplinary predisposition, system-level thinking, and empathy, they have the potential to simplify obscure, *"abstract, and scary"* concepts to make them *"banal and familiar"* (Antonelli, 2018), accessible and helpful for everybody at different scales.

1.8.1.1 Relationship with technology

The designerly mindset of some visionaries has contributed to the most important technological breakthroughs of our recent history. In 1977, Ken Olsen, co-founder of the US-based Digital Equipment Corporation (DEC), affirmed, "*There is no reason anyone would want a computer in their home."* Considering what computers had been until that moment, a niche instrument for few expert technologists, it was even a plausible position. Nevertheless, the same year, the so-called *"1977 Trinity"* of personal computers (the Commodore PET, the Apple II, and the TRS-80 Model I) came out, and what was a complex technology for the few became an essential tool for many purposes in our time. A similar fate had ARPANET: from a relatively small network connecting a few universities in the United States, it became the pervasive World Wide Web.

If these events might have been intuitions of enlightened people, the combination of technology and human-based knowledge, bringing together disciplines like computer science and psychology or cognitive science, fostered the emergence of new disciplines like human-computer interaction (HCI) and interaction design (Carroll, 2014), which proved very fruitful in transforming what technologies reserved to insiders into widespread, accessible tools.

The past shows that the technological advances that come to the market, even without a complete perspective of their possibilities, can get re-understood and re-imagined as the technology matures (Dove et al., 2017). That might also be the case for ML if designers *"immerse themselves in the possibilities of this technology, transforming it through their ways of seeing and thinking about the world"* (Hebron, 2016).

As anticipated in this chapter, if they better comprehend the capabilities of this technology, maybe strengthening collaborations with programmers (Hebron, 2016), they might be able to envision significant, engaging, and innovative scenarios that make the technological perspective thrive by combining it with art, culture, and design, capturing the wholeness of humanity (Antonelli, 2018) and understanding ML-related impact, side effects and the ways to counter them.

1.8.1.2 Transformative influence

Once curiosity and reflective inquiries lead to clear insights about the problem context (Weil & Mayfield, 2020), designers can offer a fresh pair of eyes and out-of-the-box thinking to experiment with ML and discover new, unique possibilities (Hebron, 2016). As seen, it is not unusual that after technical advances, design innovations follow. Designers have the capability to envision things that did not exist before, to find new forms and new purposes for technology, to radically re-imagine what it might be or do, and to bring them in peoples' lives (Yang, 2018). Basically, through design-driven

research and interpretation, they can generate **new interpretations of what might be meaningful for people**, leading to radical innovation (Norman & Verganti, 2014).
Paola Antonelli, senior curator of the Department of Architecture & Design and director of R&D at MoMA, NYC (2018), even defined design as the *"enzyme for any kind of innovation."* This is why they should be included in ML development teams throughout all the process. In any moment of the conjunct work, they can find occasions to propose new approaches that open to novel scenarios for real-world applications. Being aware of how to structure projects and keep consolidated exemplars as a reference to explore uncharted paths (Schön, 1983), designers can easily adapt to and be valuable resources for any other professional field (Zwick, 2006).

1.8.1.3 Cross-disciplinary and system-level thinking

Prepared to address all sorts of ill-defined problems (Cross, 1982; Rittel & Webber, 1973; Schön, 1983), designers are **naturally predisposed to listen, learn, interiorize, and leverage multiple disciplinary perspectives**. This makes them optimal candidates to work with, within different disciplines and at different scales (Norman & Stappers, 2015).
Additionally, they can creatively and purposefully tackle non-routine, changing, dynamic, open, complex, and networked situations (Dorst, 2011; Muratovski, 2016; Weil & Mayfield, 2020). They can handle challenges that span from *performance* to *systemic, contextual, and global* levels (Friedman, 2019; Meyer & Norman, 2020) in a continuously evolving world. Challenges they break down into more manageable elements and interventions (Weil & Mayfield, 2020), coordinately addressing broader issues thanks to a holistic sensitivity.
Some examples include AIGA's research detecting complex problems and technological-related matters among the most significant trends for design practice and education (Davis, n.d.); the redefinition of design programs in leading institutions expressly moving their focus to societal challenges, complex systems, and change (Voûte et al., 2020) *"for the betterment of society"* (Weil & Mayfield, 2020); and the attempt of influential academics to define a new design approach attacking fundamental *"problems involving a mix of human and societal needs where solutions involve technology,"* called *DesignX* to reinforce the underlying complexity (Friedman et al., 2019).
The fundamental skills required to cope with DesignX are a combination of multi-disciplinary expertise and a systemic comprehension of complex problems (Friedman et al., 2014). A preferable solution is then to create problem-based teams composed of people with different interests, values, and disciplinary backgrounds, useful to understand the issue at hand, who cohesively and collaboratively work together, merging their knowledge and methods (Friedman et al., 2019).
As new insights related to the problem emerge, these teams can change and evolve to develop specialized cross-disciplinary knowledge and find practical solutions to attack the core causes of problems instead of relieving the symptoms (Friedman et al., 2019).
As suggested by Meyer and Norman (2020), one constant could help the functioning of these complex initiatives, the organizational and managerial position of designers who can **keep together, mediate and coordinate the different perspectives with a holistic and systemic view**. To do so, they need to participate in key dialogues, like the one related to AI and ML.
This argument is also sustained from an ethical standpoint. One of the issues of AI systems is that they are the result of a too narrow vision – mainly comprehending computer scientists and software engineers as developers and big tech companies, with consequent economic interests, as deployers. Among the requirements of the European Guidelines to fulfill the fairness principle is the necessity to enable inclusion and diversity throughout the entire AI system's life cycle, which can be carried out in a twofold way. On the one hand, engaging people from diverse backgrounds, cultures, and disciplines in the design process of AI systems can ensure diversity of opinions and should be encouraged to avoid unfair biases. On the other, accessibility and universal design principles, as well as a beneficial involvement of all stakeholders, should be taken into account for the development of this technology (High-Level Expert Group on Artificial Intelligence, 2019b). All of this is at the very basis of design practice.

1.8.1.4 Empathy

Cross-disciplinary and systemic thinking, education, and practice are also pillars of what Norman (2023) defines Humanity-Centered Design. As an evolution of Human-Centered Design, he emphasizes the urgent need to amplify the scope of design problems, focusing no more on individuals but on the ecosystems of people, all living beings, and the physical environment (Norman, 2023), interconnected in systems of relationships and dependencies that need a long-term perspective to be tackled.
The designers' strategies for a Humanity-Centered Design approach consist of adopting a people-centered perspective, solving the right problems by identifying their root causes, understanding the system of connections characterizing the context, and doing *"small, simple interventions to tackle the most important problem"* (Norman, n.d.).
Empathy is the essential ingredient that allows this sensitive way of working. Friedman et al. (2014) recognize it as the **most important talent that enables designers to address these complex issues**. To actually embed people's needs and desires in the development of possible solutions, designers empathize with those who work or act within the system, those who approve it, and those who finally benefit or are impacted by the intervention, whether individuals, communities, organizations, political, or non-human entities.
Designers have an innate *"human advocacy,"* which allows them to balance personal concerns and broader values, principles, implications, and tradeoffs characterizing sociotechnical systems by naturally applying established theories from HCI and social sciences to real-world situations (Weil & Mayfield, 2020).
In the ML context, designers could use empathy not only to bring a humanity-centered perspective to the design of ML-infused sociotechnical systems. But, if aimed at the technology itself, as an artificial agent with kindred "thinking," it might also be the key to uncovering ML systems' unprecedented, unconventional, and pioneering possibilities. Hence, designers' empathy is the underlying core ingredient that unlocks their system-level thinking, cross-disciplinary predisposition, and transformative influence in the relationship with technology, making them ideal candidates to enter the ML discourse.

1.8.2 Design education: a bottom-up force for change

As soon as the European Union expressed its interest in being a global leader in the AI landscape, it identified education as a key area to prepare for future socio-economic changes and challenges (European Commission, 2018). They envisioned three levels in which AI literacy should spread. The first is *generalized* across society, the second focuses on *repurposing worker*s whose jobs might be automated, and the third is about *creating AI specialists.* Introducing AI and ML basic knowledge in design education may fit both in the first level, for its generalized scope to sensitize non-experts, but also in the third one, if we configure designers as key professionals in the development of AI-based artifacts. In this sense, as data scientists are essential to prepare the material to feed algorithms, designers could play strategic roles in finding uses and meaning for them.

The gap in designers' knowledge, language, skills, and competencies related to AI and ML has been highlighted in section 1.4 and, despite Dove et al. (2017), Yang, Sciuto, et al. (2018), and Zdanowska and Taylor (2022) report this does not compromise a viable collaboration with technical experts, they also open the possibility that some kind of education in this field might help them envision novel, implementable AI things for a given problem (Yang, 2020).

This necessity for designers to know how this technology works, its possibilities, and its limitations, also intercepts some requirements from design education literature. In particular, the urgency to develop *system-level thinking* is being stressed (Frascara, 2020; Voûte et al., 2020; Weil & Mayfield, 2020), and three are the main components to achieve it.

Undoubtedly, the complexity of sociotechnical systems requires **(i) more interdisciplinary collaboration**. Thus, discipline-based models need to be overcome in favor of an education that enables multi-competencies teamwork (Davis, n.d.; Frascara, 2020; Friedman et al., 2019; Meyer & Norman, 2020; Voûte et al., 2020). Indeed, it is not possible to rely on traditional disciplinary practices to face present and future challenges like the endangerment of life on Earth, and a multi-disciplinary approach might be key to fostering socially-desirable artifacts and positive innovation (von Schomberg, 2013). *"Therefore, there is the need to continuously "design" new design practices and education programs to enable a culture of design in which many people contribute to bringing about new and purposeful change"* (Auernhammer & Ford, 2022). Indeed, collaboration with people from other disciplines and cross-disciplinary skills are essential for (future) designers to be prepared for the real world (Tremosa, 2022), to actually get the bigger picture, to play and maintain a strategic position (Friedman, 2012), which leverages on their intrinsic qualities and is also coherent with the integrative nature of design.

In particular, to support contemporary education, **(ii) appropriate knowledge of technology** (Meyer & Norman, 2020) is crucial. This does not only emerge from our current panorama in which governments and universities push toward AI curricula but goes back to the origins of design. Gropius understood that technology profoundly determines how people live already in the 1920s, and Simon theorized it in close relationship with AI itself. Another indicator of the familiarity of design with technological disciplines can be found in (Christensen & Ball, 2019), who report that, in the last 40 years of publications in Design Studies, the majority of disciplinary affiliations come from Technology and Engineering (59.5%), almost doubling those from Applied Art, Design and Architecture (32.7%). Hence, design should understand technology at its core (Stoimenova & Price, 2020), even at the algorithmic level according to (Frascara, 2020). Although the definition of a proper and useful translation of technical knowledge might be disputable, what can be agreed on is that the design discipline should *"draw more extensively on knowledge developed in other established fields"* (Meyer & Norman, 2020).

Finally, **(iii) *"Design education might become much more theoretical"*** (Redström, 2020). New models of education should be explored (Friedman et al., 2019), and design approaches and processes should be actively designed, codified, and openly testable (Friedman et al., 2014; Redström, 2020). Therefore, designers (including students) should be aware of the methodologies they are using (Voûte et al., 2020), and shared vocabulary and tools would help to recognize design as a discipline (Friedman et al., 2019; Meyer & Norman, 2020). Additionally, Frascara (2020) and Meyer and Norman (2020) advocate the necessity to make ethics and design principles explicit in education. For sure, this echoes the *"nuanced views of AI as a human responsibility"* and the *"increasing effort to engage with ethical considerations"* reported in the AI100 study from Stanford University (Littman et al., 2021), but also the need for designers dealing with AI and ML-infused products and services anticipated in section 1.4 (Yang, 2018; Zdanowska & Taylor, 2022).

These premises justify the educational character of the research.

1.9 ML (for) Design. Objectives and research questions

"AI is helping us to solve some of the world's biggest challenges. [...] The way we approach AI will define the world we live in" (European Commission, 2018, p. 2). The efforts of the European Commission effectively highlight both the goal and the motivation of the research and outline the main disciplines involved: AI – in the form of its subset **ML** – as a **tool for design** to address the challenges humanity is facing. For proper orientation, the research develops underneath an **overarching ethical frame** to keep a human-centered perspective within an extended sociotechnical ecosystem, or *humanity-centered* approach in the words of Norman (2023).

The potential benefits of ML systems are undeniable and need to be exploited in a way that takes into account the preservation and promotion of life on Earth. Hence the current model, often moved by the urge to unlock new technological capabilities or by market interest and lacking a sense of the impact on society (Antonelli, 2018; Yang, Banovic, et al., 2018), needs to be overcome.

Indeed, if responsibly and meaningfully exploited, ML represents a valuable asset as it specifically addresses complex and ill-defined problems, preventing humans from logically framing the problem the machine has to solve.

These premises finally set the groundwork for framing the research and indicate its direction.

The research is positioned in the uncertain relationship between design education and the surrounding technological advancement. The piloting phase of the research, illustrated in this chapter, has been essential to gain a comprehensive view of the state of the art on the interconnections between AI and design. This has brought to light the foundational gap of designers' unpreparedness for dealing with AI and ML systems and the contribution they could make if they were more involved in the design process of these artifacts. Therefore, the basic doctoral research intends to **draw a conjunction line between the two disciplines, leveraging related ethics knowledge and combining designers' skills and sensitivity with real-world problems to be solved by integrating cutting-edge technologies (ML)**. As portrayed in Fig. 1.8, this path is still not much explored, unlike the many creative explorations of ML applications, like texts or images generators (ChatGPT, Midjourney, etc.), that both design and art fields are carrying out and are outside the boundaries of the research.

Its primary objective is to **provide a theoretical and practical contribution to design education by finding ways to translate and introduce core ML knowledge to design students.** Willing to foster a new approach to deal with ML, it targets and involves design students with basic design competencies to enable them to (i) envision consistent, meaningful, and responsible solutions integrating ML, (ii) handle ML as an asset to address current and future challenges in a human-oriented perspective, and (iii) set the basis for cross-fertilization and interdisciplinary communication between design and ML in the perspective of multi-disciplinary teamwork.

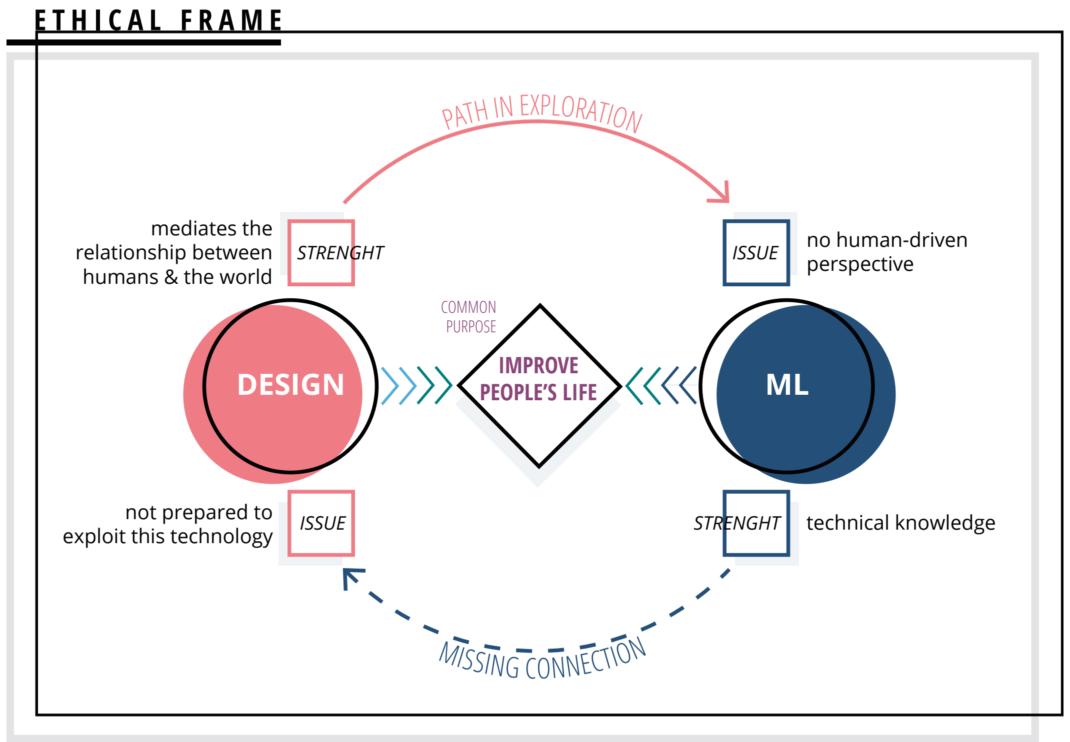


Fig. 1.8 | Research framework.

Then, to address these urgent issues, the investigation points to the roots of the problem, and its leading research question is intended to understand ***how to translate ML basic knowledge for design students?*** Here, the definition of design student is intentionally left open. In fact, whether the subject matter is intuitively relevant for students in UX, interaction, and digital design programs, it can also benefit the design of products, services, systems, spaces, and in general, any professional that could find a creative and valuable application for this technology. Thus, in the spirit of the EU's imperative to expand AI literacy among large groups of the population (European Commission, 2018), the research targets any branch of design education as a fertile ground for positive innovations. The only requirement for design students is their expected competence in understanding and handling a human-centered design process. For this purpose, master's level students are preferred, but those at the end of their bachelor's degree programs have also proven suitable.

Once the overall objective was clarified, the strategy to achieve it consisted in unpacking it into questions of incremental complexity, starting from core issues and building on their findings to address higher-level ones.

In a constructive approach, the process started from the foundations, exploring the fields of interest to identify:

>> RQ1: What can and has to be translated from ML and related ethics to design?

Then, the attention moved back to the intersection between design and ML to understand how the interdisciplinary bridge could be built and, in particular:

>> RQ2: How to frame ML knowledge for transfer?

Based on the theoretical assumptions from the first two questions, the investigation concentrated on the meta-task of design, looking for practical integration of ML and ethics knowledge in the design process, aiming to uncover:

>> RQ3: How can the theoretical constructs be operationalized into models and tools to be implemented and tested in educational contexts?

Finally, synthesizing all the insights and experiences from the previous phases, the research led to the envisioning and implementation of a possible solution to the initial, reframed problem:

>> RQ4: Which design education method can support the conceptualization of ML-infused solutions?

TO SUM UP

- Machine learning (ML) is currently the most common expression of artificial intelligence (AI). It has been identified as a *new design material* because ML systems are changing the interactions and meanings related to products and services.
- The attempt at a systematic literature review on the relationship between design and AI revealed a conspicuous lack of theory and previous research (only four relevant papers were identified).
- An exploratory and qualitative research investigates the potential for designers to deal with ML, highlighting the disciplinary affinities of AI and design, which include:
  - A common theoretical background.
  - A solution-focused approach to tackle wicked problems towards improving people's lives.
  - Abductive reasoning: achieving a goal without indications about what has to be built or which working principle should be applied. While ML systems work with examples, designers use metaphors in Schön conception.
- Designers are currently under-exploited in the development of AI/ML-infused artifacts. Literature reports that with self-learning processes based on ML experts' approaches, some designers can work in interdisciplinary teams. However, they have few possibilities to interact with experts, superficial knowledge, difficulty in comprehensively understanding ML capabilities, unfeasible HCI methodologies to deal with this technology, no prototyping modalities, and therefore no ways to comprehend mental models before launch or to anticipate ethical considerations. This would not allow them to unlock innovative solutions.
- ML systems have a unique and complex nature that poses issues to traditional interaction and user experience (UX). For instance, they have agency; they are unpredictable; they are based on unclear communication and multiple touchpoints but fail to create a holistic experience.
- They also raise ethical concerns because of their opaqueness (they are often considered monsters) and lack of human factors, as the role of people in the development process is often concealed, and impacts are not considered.

- ML, then, should be defined as a sociotechnical system, and to develop it towards meaningful directions, professional figures with a more holistic mindset should be involved.
- The current experimental and uncertain phase of ML development and deployment is a perfect target for designers. They have the potential to make it accessible and helpful for everybody by making sense and giving meaning to it thanks to their long-standing relationship with technology, transformative influence, cross-disciplinary predisposition, system-level thinking, and empathy.
- To create the conditions for designers to work with ML, the research focuses on education. This context is relevant both for the European Commission's goal to spread AI literacy and create AI specialists and for the field of design education, which seeks to develop system-level thinking, interdisciplinary communication and collaboration, appropriate knowledge of technology, and more awareness about design theoretical methodologies for 21st century designers.
- Thus, the research aims to **provide a theoretical and practical contribution to design education by finding ways to translate and introduce core ML knowledge to design students.** In particular, it aims to enable them to (i) envision consistent, meaningful, and responsible solutions integrating ML, (ii) handle ML as an asset to address challenges in a human-oriented perspective, and (iii) set the basis for interdisciplinary communication between design and ML in the perspective of multi-disciplinary teamwork.
- Then the main research question (***How to translate ML basic knowledge for design students?***) is tackled through the following sub-questions:
  - *RQ1: What can and has to be translated from ML and related ethics to design?*
  - *RQ2: How to frame ML knowledge for transfer?*
  - *RQ3: How can the theoretical constructs be operationalized into models and tools to be implemented and tested in educational contexts?*
  - *RQ4: Which design education method can support the conceptualization of ML-infused solutions?*

2. METHODOLOGY

Designers are expected to imagine new things,
not to study what exists today.

(Koskinen et al., 2011)

C H A P T E R T W O

Once the general context and objectives of the research have been introduced, this Chapter offers a **methodological overview** to depict its main features and organization. As suggested by (Rampino & Colombo, 2012a), different levels of the research are described to provide a comprehensive picture.

First, the **general nature** of the study – based on the context and requirements – is portrayed (2.1). Then, the overall **research process** to reach the objective defined by the main research question is outlined based on its theoretical references (2.2). Finally, the **specific strategies, methods, tools, and outputs** are illustrated in relation to the secondary questions (2.3).

Further details, better specifying the codified research procedures and instruments used, will be discussed in the following chapters.

2

2.1 Nature of the research

As previously introduced, the investigation unfolds in a quite uncharted territory for design. Building on scarce literature, theories, and previous experiences, and in line with the requirements of doctoral research, it has the objective to produce new knowledge at the edge of different disciplines. For this purpose, it systematizes and translates key concepts from computer science and ethics to deliver more accessible and design-oriented knowledge to design students. These premises configure the inquiry as **basic or fundamental research** according to Archer (1995). Paraphrasing the words of Buchanan (2001), it is meant to empirically examine some principles to develop theories that may have wide-ranging implications for the design discipline and bridge it to other branches of knowledge to provide means to understand and meaningfully exploit ML systems in the design of products and services.
Another patent and essential trait lies in the *"fusion of disciplines,"* as described by Muratovski (2016), that represents both the intent and the starting point of the research, which aims to provide sufficient and comprehensible knowledge for design students to work across disciplines in novel ways.
The **transdisciplinary** character derives from the necessity to frame and solve the complex problems affecting our societies, for which disciplinary competencies are not sufficient (Muratovski, 2016).
Indeed, the complex, multifaceted, and immature context – as there is a lack of theory and previous research – also calls for **exploratory** and predominantly **qualitative** research. As Creswell (2014) claims, *descriptions, interpretations, verifications,* and *evaluations* of the issue at hand can benefit the design research community. To do so, phenomena have to be explored and communicated in an attempt to develop new theories. For this purpose, quantitative measures can hardly be helpful as they would need a substantial amount of literature to provide direction for formulating research questions and initial hypotheses, and systematic reviews are not feasible. Instead, literature has been used to initially frame the problem (Creswell, 2014) and to inform research activities focused on natural settings to make sense of the situation in all its different layers and dimensions without aiming at simplifications (Muratovski, 2016).
Consequently, the inquiry results are **reflective** and **interpretative** (Rampino & Colombo, 2012b). It leverages field research methods (Koskinen et al., 2011) to infer information from direct observation and participation of the researcher. Thus, it is inevitably characterized by subjectivity and the acknowledgment of situated

outcomes that cannot be generalized in absolute terms. In this sense, it acquires the connotation of a **constructivist** approach: the central goal of the research is to understand (Robson & McCartan, 2015), identify subjective meanings, and inductively generate a comprehensive picture of the situation being studied by also relying on the participants' views and starting from open-ended questioning (Creswell, 2014).

2.2 Overall research process

In terms of methodology, the inquiry cannot be univocally labeled. Embracing P.J. Stappers's view, it is characterized by a research-through-design soul, as design activities and background here *"play a formative role in the generation of knowledge"* (Stappers & Giaccardi, 2017). Indeed, through gaining actionable knowledge from complex and ill-defined contexts, effectively reframing the problem, and iterating possible solutions, the final results are uniquely shaped and wouldn't have been the same if tackled from any other discipline's perspective.
From a procedural point of view, instead, it takes on the traits of action research, defined by Archer (1995) as a *"systematic investigation through practical action calculated to devise or test new information, ideas, forms or procedures and to produce communicable knowledge*." With the broad aim of understanding ***how to translate ML basic knowledge for design students***, it falls into the established tradition of action research applied to education (Bresler, 2021; Robson & McCartan, 2015) and the kind of investigation that represents and transfers knowledge to promote students' understanding, as portrayed by Gore and Zeichner (1991).
Artifacts (here intended as the result of a design process) are not the focal point of the research. They can concretize the research outputs, but their value lies in the materialization of hypothesis, knowledge, values, and know-how (Villari, 2012) to be tested in the real world. Hence, the focus is on the process that brought to them, which is mainly characterized by reflective practice. To put it in Schön's (1983) words, at the core of the investigation, there is a *reflective conversation* with the issue at hand that draws a clear connection between action research and research through design, as both the former and design practice rely on reflection-in-action and they share the same process (Swann, 2002). Particularly appropriate in complex and uncertain situations, this approach allows new findings to be obtained while addressing a reframed problem. In this way, a better understanding of the situation emerges from the attempt to change it and changes from the attempt to understand it (Schön, 1983). This continuously opens the space for further exploration in a spiraling flow.
Built on constructive premises, the research is highly iterative, and the process to address such a broad challenge assumes a cyclic structure, as Muratovski (2016) effectively illustrates (Fig. 2.1).
The research questions are increasingly comprehensive, moving from essential aspects of the translation to wider and more articulated constructs that, ultimately, outline an educational method for design students as the ultimate synthesis of the knowledge developed (Fig. 2.2).
While a detailed description of the research process, including the specific methods employed and outcomes obtained, is depicted in the following paragraph, the theoretical bases are explained here. Four main phases characterize the process of

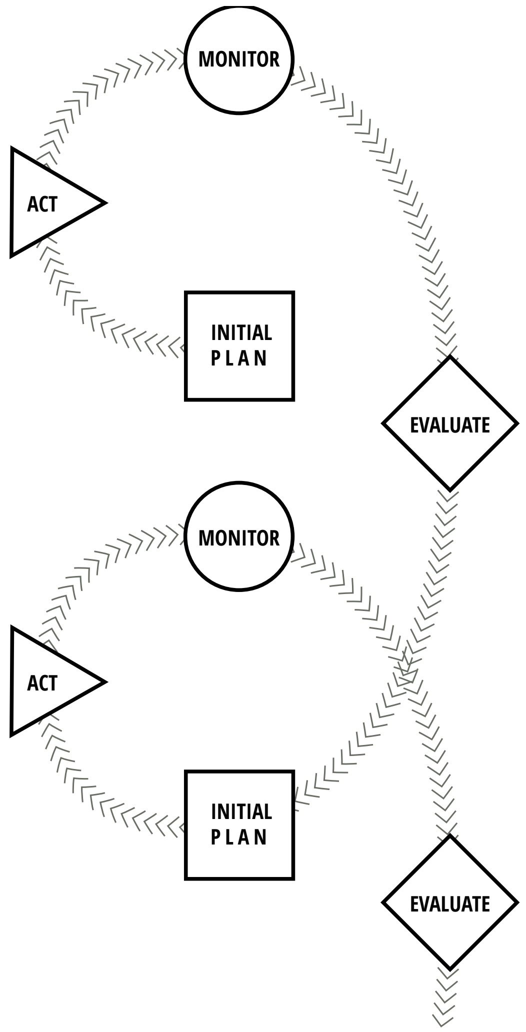


Fig. 2.1 | The process of applied research as represented by Muratovski, 2016.

addressing each question. To portray them, I refer to Muratovski (2016) and Swann's (2002) contributions.

Phase 1 > Planning. A preliminary understanding of the context was essential to develop a strategy to respond appropriately to a question. In this phase, the problem was identified, analyzed, and synthesized into possible solutions to translate ML knowledge for design education (usually characterized as a set of assumptions).

Phase 2 > Acting. The assumptions were then translated into practice and the research question into a project (Robson & McCartan, 2015). Therefore, the possible solution was designed, materialized, and executed to be tested.

Phase 3 > Observing. This crucial step allowed to understand whether the hypothesis was *affirmed* or *negated* by collecting and documenting the consequences of the action (Schön, 1983).

Phase 4 > Reflecting. Finally, the results were evaluated to determine if they could be satisfactory and to understand which kind of issues emerged. This is the phase in which the research findings were progressively uncovered, and new starting points for further investigation were produced, building the premises for the following research question.

Identification of the research context and gap

RQ1: What can and has to be translated from ML and related ethics to design?

PLANNING

Research Actions

Literature review (*snowball sampling procedure*)

Desktop research (*e.g. ML courses*)

Outcomes

Common educational strategies are not suitable for non-experts

ACTING

Research Actions

Desktop research

Case studies collection and auto-ethnographical experience (*16 strategies for public outreach*)

OBSERVING

Research Actions

Content analysis

Outcomes

Most recurrent elements to explain ML

REFLECTING

Research Actions

Data elaboration

Theoretical synthesis (*core concepts to understand ML*)

Outcomes

Key elements to translate

Encoder & Decoder

ML Suitability Matrix

Fig. 2.2 | Overall research process.

RQ2: How to frame ML knowledge for transfer?

PLANNING

Research Actions

Theoretical synthesis (*essential knowledge for translation*)

Outcomes

ML Designerly Taxonomy

Formats to explain ML capabilities (*ML Pills*)

ACTING

Research Actions

Online workshop (*ML Pills for Designers - 17 MSc DID students from Polimi*)

OBSERVING

Research Actions

Observation, Questionnaires, Discussion

Outcomes

Understanding of the preferred forms and language of translation

REFLECTING

Research Actions

Data elaboration

Theoretical synthesis (*translation of ML capabilities & requirements for ML-infused solutions*)

Content analysis (*ethical guidelines for AI*)

Expert reviews

Outcomes

ML Agents

Responsible Cycle for ML Design

Foundational assumptions for the educational experience

RQ3: How can the theoretical constructs be operationalized into models and tools to be implemented and tested in educational contexts?

PLANNING

Research Actions

Theoretical synthesis (*materialization of a holistic, flexible, and modular educational approach*)

Outcomes

STEP 1_Introductory Game to ML Responsible Design

STEP 2_ML Concept Building Blocks tool & VALUable by Design Expansion

ACTING

Research Actions

STEP 1_Online & in-presence workshop (*Introductory Game to ML Responsible Design - 4 PhD students from Polimi*)

STEP 2_*Superpowered museums* workshop (*3rd year bachelor students in interior design at Polimi*)

OBSERVING

Research Actions

Observation, Focus group, Questionnaires, Oral feedback

Outcomes

Strenghts and criticalities of the educational experiments

REFLECTING

Research Actions

Data elaboration

Theoretical synthesis (*materialization of a holistic, flexible, and modular educational approach*)

Outcomes

Insights for defining educational models, balancing guidance and students' agency

RQ4: Which design education method can support the conceptualization of ML-infused solutions?

PLANNING

Research Actions

Literature review (*preferable pedagogical methods*)

Outcomes

Flexible and modular educational models (*consistency, responsibility, integrated*)

ACTING

Research Actions

ML Hero Agency & VALUable ML Heroes (*Nantes - 18 Digital Design MDes students & 10 MDes UX Design students*)

VALUable ML Hero Agency *(Graz - 7 MSc Interaction and Media Design students, Madeira - 15 3rd year bachelor students in design)*

Condensed VALUable ML Hero Agency (*Polimi - 104 MSc Design & Engineering students*)

OBSERVING

Research Actions

Evaluation research through: observation, formative tests, peer evaluation, questionnaires, oral, and written feedback

Outcomes

Strenghts and criticalities of the educational experiments

REFLECTING

Research Actions

Data elaboration (*mixed method strategy*)

Expert reviews

Outcomes

Interdisciplinary educational method to enable the envisioning of ML meaningful solutions

This reflects two of the central features of action research as defined by (Robson & McCartan, 2015): the constant tension towards improvement (of understanding, practice, and, eventually, of the situation) and the need for involvement. The former is inherent in the ambition to introduce knowledge from other disciplines (and new to the design field) into design education to possibly achieve purposeful change (Auernhammer & Ford, 2022) in developing future technological applications, and it manifests in the constructive approach. The latter is related to dealing with a recent and little-explored matter. Indeed, the direct engagement of representatives of the core audience of the research was precious to collect insights. With a practice-based and participatory approach, theories have been turned into concrete things to be experienced in the field, and design students from different universities and backgrounds have been involved as participants in the experimental *action process* and called to evaluate it. In the following chapters, the details of these systematic yet qualitative activities are reported, emphasizing the *"empathic grasp"* the research aims to acquire, which goes beyond mere data analysis (Koskinen et al., 2011, p. 75).
Of course, this method presents some limitations. It cannot be deterministic and strongly depends on the researcher's perspective and on the context in which it takes place. Additionally, it produces *"local"* understanding (Koskinen et al., 2011) that hinders the generalization of the findings. For this reason, some mitigating measures have been adopted, such as providing rich and explicit details in the report of the activities (underlining bias identification) and triangulating data, methods (including quantitative ones), and perspectives (conducting expert interviews) whenever possible. Further specifications are provided in the next paragraph and the related chapters.

2.3 Planned actions and main outputs

As anticipated, each research question has been addressed with a constructive approach according to the four phases that characterize action research – in a more or less accurate manner. In the following, the research actions and methods employed are portrayed within a schema, while more methodological details are specified in the corresponding chapters. Additionally, the milestones to which each research question led and their connection to the next stage of the investigation are highlighted.

2.3.1 RQ1: What can and has to be translated from ML and related ethics to design?

Planning. The preliminary phase of the research was purely exploratory. Since only limited materials linking design education and practice to the development of AI and ML systems were available in 2019, in early 2020, a qualitative literature review was conducted based on a snowball sampling procedure. As this inquiry mainly instructed the initial research strategy (defining the research objective and the first draft of research questions), an extensive deepening into technical knowledge was necessary to have a wide-ranged perception of what ML is and what can and should be translated to design. Literature in both AI and related computer ethics fields was

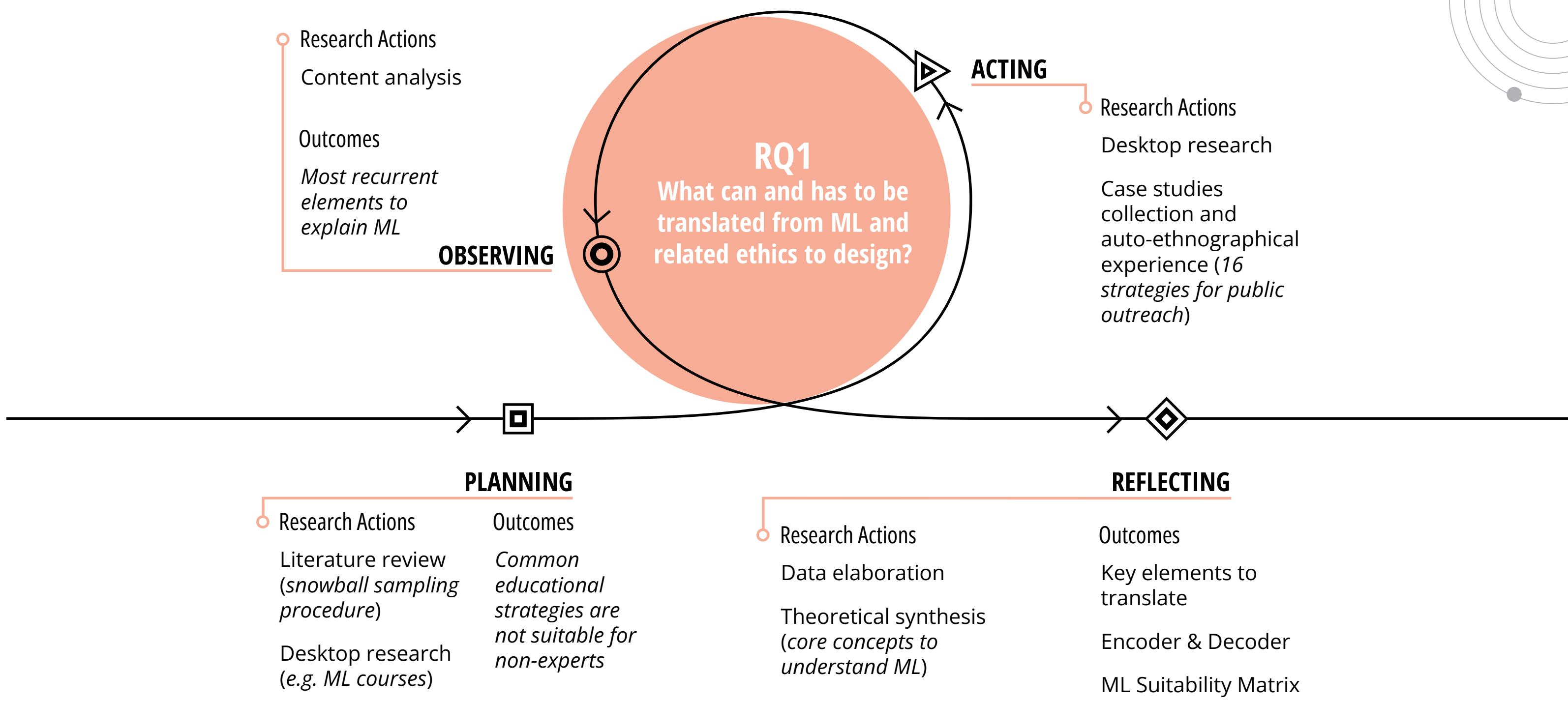


Fig. 2.3 | Outline of the first cycle of the research, responding to RQ1.

instead abundant. Therefore, the review firstly got to world-wide referenced authors and courses, and the newly updated and most broadly adopted textbook *Artificial Intelligence: A Modern Approach. 4th Edition* (Russell & Norvig, 2020) became the main reference. This study and further clarifications asked ML experts with informal interviews led to envisioning a necessary practical activity. Indeed, technical jargon and common educational strategies are unsuitable for non-experts, and it is quite fuzzy to convincingly grasp the core principles of AI, as (Russell & Norvig, 2020) stated.

Acting. At this early stage, the relation with action research was faint. The practical activity consisted of an auto-ethnographical experience to understand what strategies had been adopted to simplify AI and ML-related knowledge for public outreach. In desktop research to investigate non-technical dissemination, 16 case studies were collected, varying for target audience and format. Then, a content analysis was performed.

Observing. The first-hand experience of the outreach tools was essential to gain a broad perception of the most common explanations about what ML is, what it can do, and which are its major limitations that design and ethics can address. Annotations from the analysis also highlighted the most recurrent elements that could inform the knowledge translation.

Reflecting. The evaluation and elaboration of the data collected resulted in identifying the key contents and skills to translate – partially synthesized in the ML Suitability Matrix – and of the "minimum viable structure" to define ML systems – firstly represented with the *Encoder* and *Decoder* tools. This process triggered further reflection on the level of technicality designers need to acquire to deal with the development of ML systems. In the end, it was assumed that essential explanations of the core ML capabilities, using a designers'-friendly language, are the minimal and focal element for introducing ML to design students. No further technical information should be necessary for an introductory level. Altogether, these ingredients suggested the need for a theoretical systematization and were used to construct it and put it to the test in the following research stage.

2.3.2 RQ2: How to frame ML knowledge for transfer?

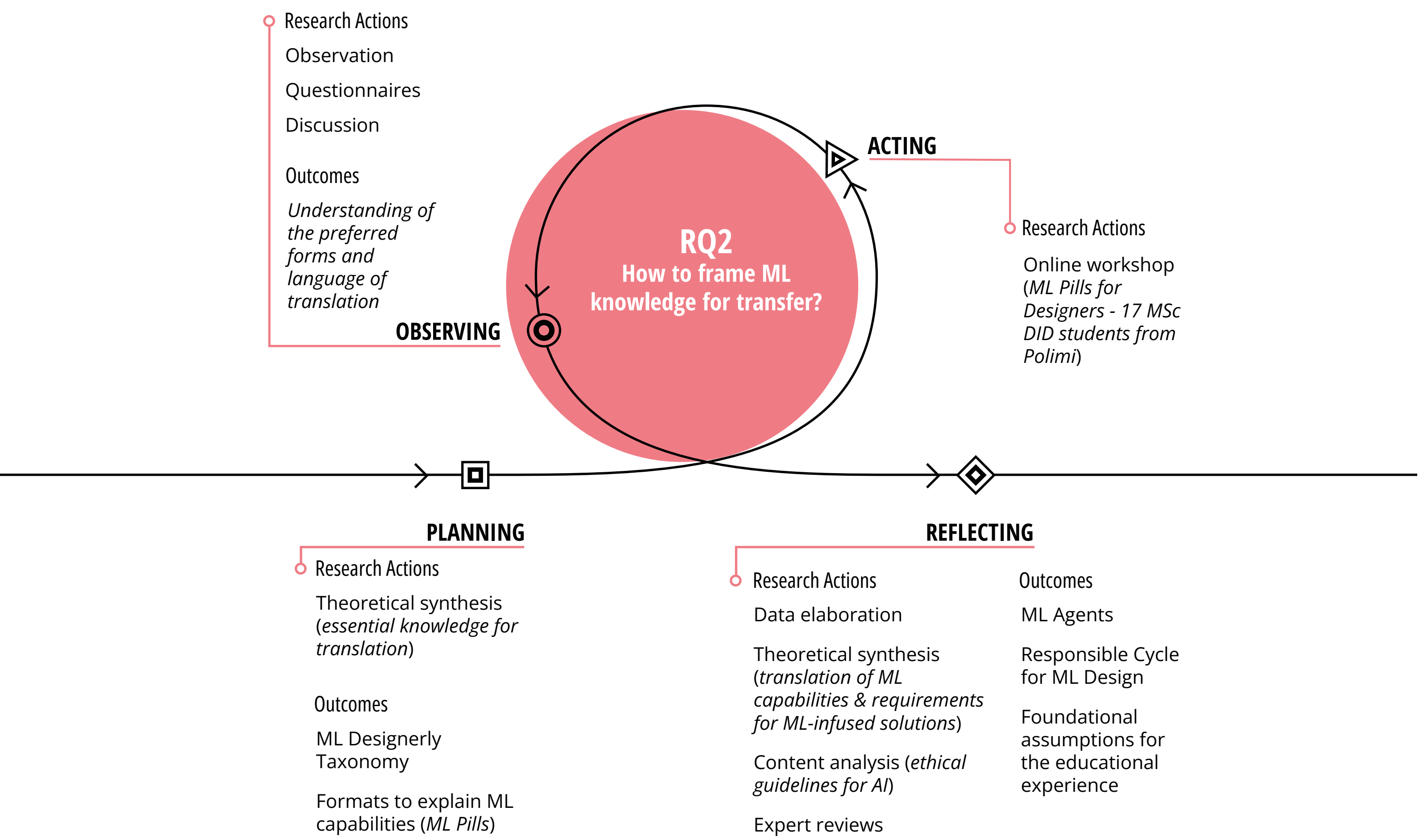


Fig. 2.4 | Outline of the second cycle of the research, responding to RQ2.

Planning. Moving from the felt necessity to synthesize, organize, and simplify the understandings from RQ1, a ML Designerly Taxonomy was developed and assessed through interviews with a ML expert and a UX design academic. It represents a beneficial systematization for both designers and ML experts and facilitates communication among and between these specialized figures. To complete the framing of knowledge, however, a better comprehension of how the target audience relates to ML was essential.

Acting. The theoretical inferences converged in a preliminary workshop involving 17 students enrolled in the Digital Interaction Design master program at Politecnico di Milano and aimed at understanding the best form for translating technical knowledge. For the occasion, some tools using four different languages to summarize ML's core concepts and introduce them in the design process have been developed and tested in a project-based activity through observation and questionnaires.

Observing. The observation, in particular, focused on discerning whether the identified contents and the forms and languages proposed to depict them could effectively enable design students to properly apply them in a project. The practical experience unsurprisingly revealed that combining the formats proposed (instead of selecting a preferred one) might be a more desirable option for the translation and that visual support and examples are crucial. Overall, it highlighted the helpfulness of a translation tool in a hands-on application.

Reflecting. This phase proved valuable as a theoretical and practical elaboration of the insights collected in the participatory research action. Indeed, the observation and the feedback received contributed to the evolution of the tools proposed into a more comprehensive one, ML Agents, to portray the capabilities of ML systems. Additionally, some room was detected to include ethical issues and complete the framing of ML knowledge. For this reason and to foster the envisioning of ML implications, limits, and potentials, a content analysis of the ethical guidelines collected in the AI Ethics Guidelines Global Inventory was carried out. Its purpose was to reflect on the values, risks, and possibilities for the responsible development and deployment of ML systems. It finally resulted in a *Responsible Cycle for ML Design*. At this point, it became clear that the design process had to present three key factors to envision meaningful ML-infused solutions. The first, intrinsic to design education and permeating the whole process, is a humanity-centered approach. Instead, the other requirements are (i) ***consistency*** with the possibilities the technology offers and (ii) ***responsibility***, respectively, tied to knowledge derived from ML and ethics. What further research should uncover is how to enable their achievement.

2.3.3 RQ3: How can the theoretical constructs be operationalized into models and tools to be implemented and tested in educational contexts?

Planning. Having identified the essential technical contents and ethical issues that design students need to start envisioning concepts including ML systems, a way to provide them with this knowledge had to be elaborated. Specifically, in accordance with the twofold soul a meaningful result should have, two possibilities were pictured for testing. One was promoting a holistic approach and merging all the disciplinary perspectives in a single educational activity. The other proposed two separate paths, each focusing on a different disciplinary requirement (consistency or responsibility). The tools to operationalize and support both possibilities were developed in this planning phase. The *Introductory Game to ML Responsible Design* (realized in a physical and a digital version) synthesized the entire process to design a ML-enabled system integrating ethics-driven moments for reflection and decision-making. To start exploring a more focused educational activity, the consistency path was deemed of primary importance, and a specific tool to guide and support the development of a consistent application for ML capabilities was conceived: the *Concept Building Blocks*. Then, in the perspective of modularity and the multidisciplinary findings of the

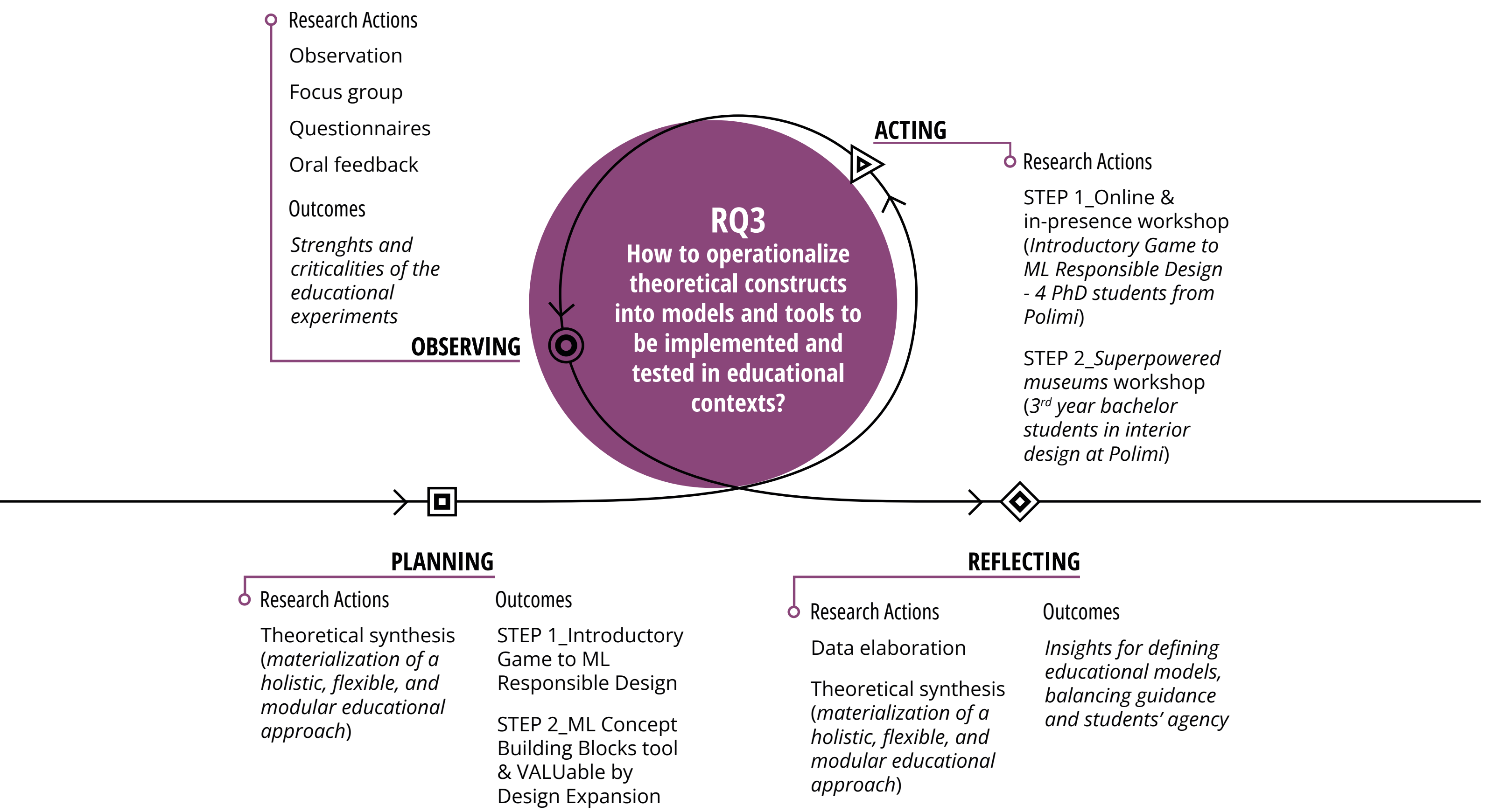


Fig. 2.5 | Outline of the third cycle of the research, responding to RQ3.

research, it was complemented by the *VALUable by Design Expansion*, which integrates an early definition of the intended impact and main principle to pursue and value-driven anticipation of risks.

Acting. The tools and the underlying assumptions needed to be tested. An in-person and an online workshop were organized to assess the effectiveness of the holistic approach and the *Introductory Game to ML Responsible Design* as a tool to foster it. They involved a total of four Ph.D. students in design at Politecnico di Milano, with some or no prior knowledge of the subject matter. Their closeness to the research's intended audience, analytical skills, didactic experience, and consciousness about design research theories and methods made them very helpful in building an insightful discussion and peer evaluation of the activity.

As a pilot study, the *Concept Building Blocks,* instead, was the protagonist of the *Superpowered Museums* workshop, held within a Final Synthesis Studio and addressed to 38 third-year interior design students from the bachelor degree program of Politecnico di Milano. They had to portray *science fiction prototypes* of ML-superpowered exhibition spaces enabling new and futuristic experiences.

Observing. Both experiments aimed to gather feedback and proved helpful in fueling subsequent reflections. For research purposes, the game experiences were carefully observed by the investigator, and they were followed by a focus group session where potentialities and criticalities of the tool and approach emerged. Instead, the project-based and future-oriented workshop was structured as a case study involving multiple data collection methods (Robson & McCartan, 2015): namely questionnaires, observation, and thematic analysis from delivered visual materials and final discussions.

Among the most relevant insights are the participants' successful embrace of the multidisciplinary perspective and the recognition of the quite prescriptive character of the introductory game, which might highly benefit from modularity and flexibility to serve different purposes. As well, the *Concept Building Blocks* tool proved effective in guiding and supporting interior design students (with no prior knowledge of the topic) in developing an exhibition concept consistent with ML.

Reflecting. The premises of this phase opened the space for reflections on the elaborated tools and the structured educational activities. As the format and language of the *Concept Building Blocks* were more finely articulated and supportive, it proved a suitable basis for creating a multilevel and modular tool. Moreover, the observation and feedback collection about the educational strategies adopted to test the preliminary tools were synthesized in helpful considerations for further formalization into didactic models.

2.3.4 RQ4: Which design education method can support the conceptualization of ML-infused solutions?

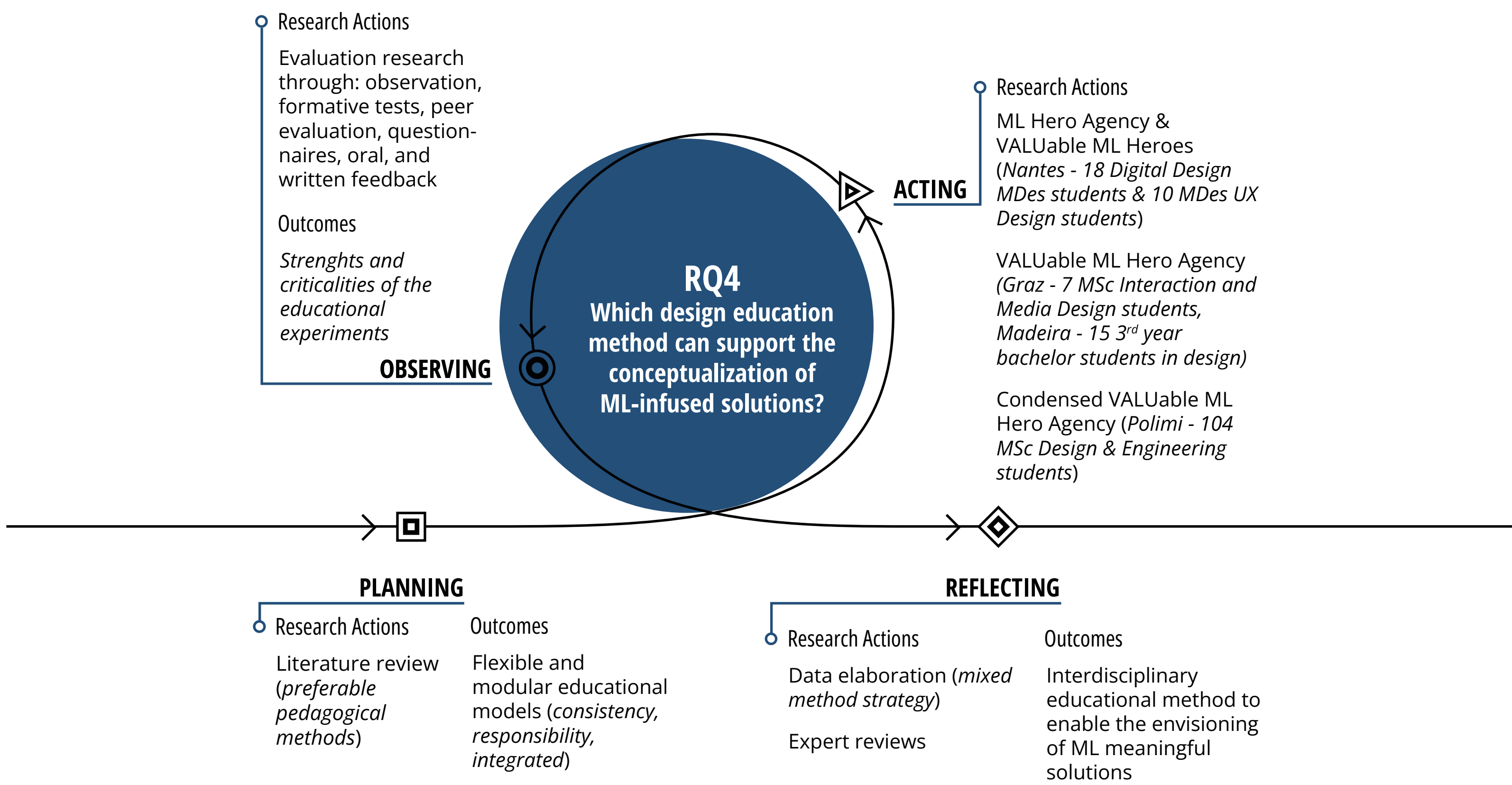


Fig. 2.6 | Outline of the fourth cycle of the research, responding to RQ4.

Planning. Previous experiments highlighted that active educational experiences with procedural information in a setting where design students have enough agency

to explore and make decisions are beneficial characteristics for the educational framework synthesizing the research outcomes. Building on this insight, a literature review on preferable pedagogical frameworks for design education and suggested practices to encourage the acquisition of complex contents was performed to properly define the constituent features of the educational method. A problem-based pedagogical model and Gagné's events of instruction (Sancassani et al., 2019) have been selected as primary references for constructing a didactic project. Equally essential are flexibility and modularity, which allow to address different targets and meet the necessities of multiple didactic contexts. Finally, the theoretical structure for the educational method was portrayed.
For research purposes, two main educational modules were developed as workshops of (minimal) variable lengths, later combined into a comprehensive one. One aims to design-driven ML-infused solutions *consistent* with ML capabilities and unfolds in a two-day time span. The other focuses on a *responsible* approach to the conceptualization of products and services integrating ML and requires, at least, half a day.

Acting. The two strategies have been tested as single modules and, eventually, consequentially combined in an evaluation research frame (Robson & McCartan, 2015). Specifically, they were materialized and iterated in four workshops organized in different design educational contexts. The first experimentation, at the École de Design Nantes Atlantique, included two separate workshops. One focused on the development of technologically consistent concepts: the *ML Hero Agency,* and addressed a class of international students from the first year in the digital design master program. The other encouraged a responsible approach to the design of ML-infused solutions: *VALUable ML Heroes,* involving first-year master students in UX design. Subsequent iterations of the workshops merged the two approaches in a consequential way. The *VALUable ML Hero Agency* workshop was then tested at FH Joanneum University in Graz with first-year students of the interaction and media design master programs, at Universidade da Madeira with third-year bachelor students in design, and – in a condensed version – also at Politecnico di Milano, with master students in the first year of the Design & Engineering program.

Observing. The workshops were intended as multiple case studies. The convergence of formative and research objectives was helpful in identifying common evaluation strategies analyzed with mixed methods for the triangulation of qualitative and quantitative data from observation, formative tests, peer evaluation, questionnaires, and oral and written feedback. Overall, the results showed positive signals (e.g., students' enthusiastic comments on the approach and tools, their perceived interest and comprehension of the presented subject, and their capability to perform the required tasks at satisfying levels), with minor issues to be addressed in subsequent iterations. However, the limited time in the condensed version had a negative impact on the experience, which served to highlight the most relevant requirements for the educational models.

Reflecting. Elaborating on the strengths and weaknesses that emerged from the workshops, an educational method was framed to set some significant steps for teaching and learning how to envision meaningful ML solutions merging different disciplinary perspectives. This synthesizes the findings of the experimental research, marking the final contribution of the doctoral research to design knowledge. Of course, it can inspire exploring alternatives based on the same focal points. Further implementation strategies and scenarios in which the developed tools and models can be applied have been conceived, and opportunities for future improvement have been envisioned.

2.3.5 Transversal research iteration: expert interviews

A triangulation of observers' points of view was finally conducted to acquire different perspectives, validate the main assumptions that have guided or emerged from the research, and infer richer and less subjective conclusions. For this purpose, semi-structured interviews with expert design researchers and educators were organized as a conclusive research action in the iterative spiral cycles of three key passages of the qualitative action research strategy. Specifically, they relate to (i) the thematic context of the research, (ii) the framing of the translation of ML knowledge (RQ2), and (iii) the educational method structure (RQ4).
The interviewees are all distinguished and experienced design researchers and educators, working at the edge of the boundaries of the design discipline and familiar with digital technologies as design tools. Specifically, they include Jodi Forlizzi (School of Computer Science, Human-Computer Interaction Institute, Carnegie Mellon University, Pittsburgh, USA); Marianne Graves Petersen (Department of Computer Science, Aarhus University, Denmark); Peter Gall Krogh (Department of Digital Design and Information Studies, Aarhus University, Denmark); Johan Redström (Umeå Institute of Design, Umeå University, Sweden); John Sharp (School of Art, Media, and Technology at Parsons The New School for Design, New York, USA); John Zimmerman (School of Design, Human-Computer Interaction Institute, Carnegie Mellon University, Pittsburgh, USA).
Their precious contribution to the research consisted of providing their personal position on thematic and methodological issues regarding the translation of ML knowledge for design education. Not to influence them to focus on any specific topic, the semi-structured questions were deliberately broad, leaving the discussion to be driven by the experts' sensitivity.
In particular, the interviews developed in a maximum of one-hour time frame according to the following outline.
The interviewees were provided in advance with a two-page document anticipating the research summary in case they wanted to get a more accurate picture. At the beginning of the interview, ice-breaking questions were meant to dive into the research topic and understand the respondents' position. They could sound like:

> *What kind of relationship do you foresee between designers and the development of products and services integrating ML systems in the near future?*

> *What do you think about introducing ML in design education?*

Then, the foundational assumptions of the research were briefly introduced. Namely, the purpose (enabling designers to envision meaningful ML applications), articulated in the two main disciplinary strands identified (the technological and the ethical one), and the main theoretical premises elaborated: the *ML Design Taxonomy* and the *Responsibility Map for ML Design*. Consequently, the core practical assumptions emerging from RQ2 were challenged with the questions:

> *What do you think about using ML capabilities as the focal point in translating and communicating ML knowledge to design students?*

> *What do you think about leaving students free to explore solutions to ethical concerns with minimal hints and related education?*

Finally, a quick overview of the translation of these premises in an educational method was presented to collect the expert's free impressions on issues like intended learning outcomes (ILOs), format, structure, covered topics, tools, expected outputs, and so on. In particular, they were asked:

> *What weaknesses, strengths, or opportunities can you spot in the approach proposed for the educational models?*

The results of these interviews inform the *reflection phase* of the identification of the research gap, RQ2 and RQ4. Therefore, they are respectively portrayed in apposite boxes at the end of chapters 4 and 6, titled "In experts' words," because quotes from the interviews are used as much as possible to depict their feedback.
In general, consonant and slightly nuanced positions related to the subject matter emerged. The perception of the current relationship between designers and AI, though, presents some differences based on the European or American context, and the relevant formative objectives change based on the disciplinary background of the respondent (design or computer science). Overall, interesting comments and suggestions fed reflection to further develop and improve the educational experiences and can be read in the appropriate sections of the contribution.

3. ML IN TRANSLATION. SUBSTANCE AND BOUNDARIES

I see AI as a tool. When designers master that tool, they can expand their ability.

(Antonelli, 2018)

C H A P T E R T H R E E

Delving into RQ1 (***What can and has to be translated from ML to Design?)*** ML is primarily analyzed from the perspective of computer science (3.1). In the attempt to understand how it can be defined and presented, its evolution and most recursive concepts are investigated as explained in its disciplinary context. The inherent difficulties and issues they raise for non-specialists are then discussed with the support of computer ethics. This fueled the hypothesis that a direct transfer of ML knowledge would not be significant for designers and translation work is needed to make it easier and clearer to grasp.

As simplification strategies for non-experts are already in place, they have been explored in a case study analysis of 16 examples of ML knowledge outreach, differing in format and target audience to have a good picture of the possibilities (3.2). A content analysis based on two coding cycles and an auto-ethnographical investigation allowed the extrapolation of information about the addressed topics and their overarching categories and useful insights related to the most effective language and communication strategies to operate a designerly translation.

In the reflective stage of the constructive process to answer RQ1, the findings from the exploratory and analytic activities were further elaborated and led to the definition of the **core elements to describe ML systems**, supported by the development of ML *Decoder* and *Encoder* and ML *Suitability Matrix* (3.3).

3

3.1 Exploring a disciplinary perspective on ML

The journey to get familiar with ML could only begin at its cradle, trying to immerse in the perspective of computer science and positioning this discipline as a subset of AI. From the start, inherent difficulties emerged. As shown in Fig. 3.1, AI and ML-related literature has exponentially increased in the last decade, reaching over 300 thousand publications – in English – in 2021 (Zhang et al., 2022). And the discipline dates back more than 70 years. Inevitably, over time and countless experiments in AI and ML, a vast and multifaceted landscape of perspectives has emerged, and it is not easy to find agreement on the topic.

On the contrary, the wide-ranging knowledge produced several ramifications, and it is difficult to find consistent systematizations, hence clarity in this domain. Instead, the current hype connected to the new possibilities enabled by recent technological developments has increased the circulation of misconceptions. Indeed, artificial intelligence and machine learning have become buzzwords spreading in all areas of human life, sometimes improperly.

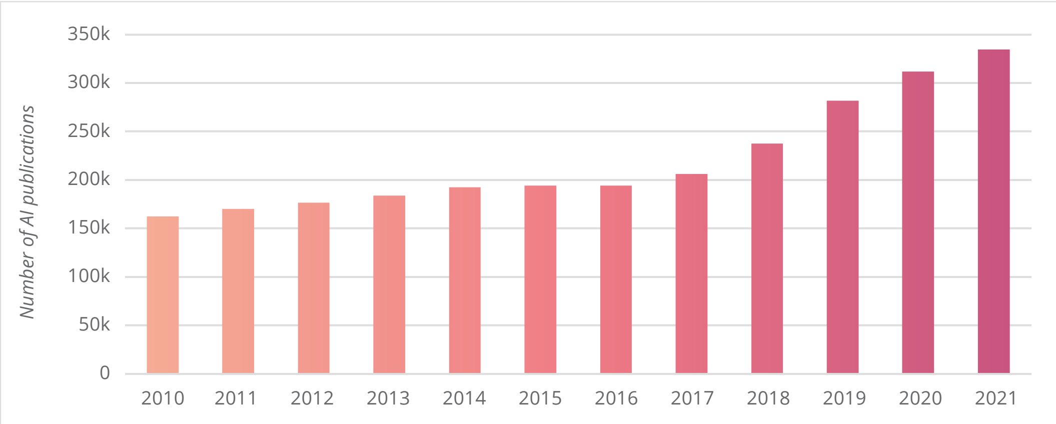


Fig. 3.1 | Number of AI publications in the world. From 2010 to 2021, the total number of AI publications doubled, growing from 162444 in 2010 to 334497 in 2021. Source: AI Index Report (Zhang et al., 2022).

This is why, to find a reliable entry point to the discipline, the researcher has embarked on a web-based search of the most influential sources and authors to get introduced to machine learning. She started to study textbooks and online courses like *Machine Learning* by Andrew Ng, Stanford University, on Coursera (which has now been separated and expanded into a ML Specialization). For a better understanding of ML, the scope of the inquiry was also extended to related subjects. It included courses like *Basic Statistics* offered by the University of Amsterdam and *Programming for Everybody (Getting Started with Python)* by the University of Michigan through the Coursera platform. However, the primary reference for technical explanations has been the foundational textbook: *Artificial Intelligence. A Modern Approach* (Russell & Norvig, 2020), the most-adopted AI textbook – by over 1500 schools in 134 countries or regions (Norvig, 2022). It has been considered of utmost relevance also because of its very recent update (the new edition came out at the same time as this part of the research was carried out), the familiar and precise tone they use, the clear structure, and the breadth of its contents.

In the following sections, some insights from this investigation are presented to give an overview of the context to be translated, highlighting key issues and challenges for the knowledge transfer. It is not meant to be a comprehensive presentation, but it synthesizes focal points according to the author's perception.

3.1.1 From humans to agents. The problematic evolution of AI definition

The concept of artificial intelligence eludes a precise definition. Turing (1948) even defined it as emotional, as it mainly depends on a subjective perspective of what people are willing to recognize as intelligent. Undoubtedly, it is perceived as a distinctly human trait, and the origins of artificial intelligence are inevitably shaped by the relationship with human beings, as its deep connections with philosophy, neuroscience, cognitive psychology, and linguistics demonstrate. The primary reason scientists and engineers invented computers was to assist people in solving difficult problems more easily and quickly (e.g., Turing's machines cracking German Enigma codes during World War II). The more complex the problems, the more the computers had to be as efficient as humans in dealing with them. Therefore, people had to craft machines capable of helping them in what they usually do. Intuitively, they concluded that these machines should behave like people to handle their tasks effectively.

At least, these are the premises for the summer workshop that gave birth to the discipline at Dartmouth College in 1956. A group of ten researchers, led by John McCarthy (computer and cognitive scientist at Dartmouth College), aimed to identify the strategies *"to make machines use language, form abstractions and concepts, solve kinds of problems now reserved for humans, and improve themselves*" (McCarthy et al., 1955). Indeed, they expected to delineate all the aspects of human learning and intelligence so precisely that they could then replicate the same processes in machines. However, a twofold interpretation has revolved around the concept of artificial intelligence. According to one part of AI researchers, including McCarthy, it is synonymous with creating a super brain capable of simulating human behavior. For others, like Engelbart (engineer and inventor in the field of HCI), AI is a way to augment and amplify human potential. It is more of an instrument than a substitute for people (Winograd, 2006).

This dichotomy is well synthesized by Russell and Norvig (2020), who distinguish two pairs of contrasting interpretations of artificial intelligence. One reflects the above-mentioned distinction between an accurate simulation of *human* performance and a more formal definition of *rationality*. The other focuses on the topic as an internal *thought* process or an external manifestation of *behavior*.

The varying combinations of these concepts constitute four possible approaches to AI. The (i) *"Turing test approach"* recognizes AI systems as intelligent because they *act humanly,* which can be measured by their capability to effectively communicate by processing natural language (NLP), store information (knowledge representation), answer questions and draw new conclusions (automated reasoning), adapt to new circumstances and identify patterns (ML). The (ii) *"cognitive modeling approach"* sustains that AI systems should *think humanly*, showing internal thought processes similar to those that psychology, cognitive and neural sciences study in people. The (iii) *"laws of thought approach"* transfers irrefutable reasoning processes derived by logic or probability to describe how AI systems should *think rationally.* Finally, the currently prevailing interpretation is the *(iv) "rational agent approach,"* according to which AI and ML systems should *act rationally.*

This conceptualization, which the authors use as a unifying theme throughout the textbook, identifies these systems as agents able to perceive their environment through sensors and act upon it through actuators. Its simplicity makes it an effective synthesis because it is easily measurable and independent from fuzzy and ambiguous references to human intelligence. Additionally, as anticipated in Chapter 1, this explanation is also functional for theories developed outside computer science, including ethics and design, which entail the non-human active role these systems can have. To further validate its efficacy, it is the foundation of the precise definition of AI systems that the High-Level Expert Group on AI (a diverse group of experts appointed by the European Commission in 2018) agreed on:

> ***Artificial intelligence (AI) systems are software (and possibly also hardware) systems designed by humans that, given a complex goal, act in the physical or digital dimension by perceiving their environment through data acquisition, interpreting the collected structured or unstructured data, reasoning on the knowledge, or processing the information, derived from this data and deciding the best action(s) to take to achieve the given goal. AI systems can either use symbolic rules or learn a numeric model, and they can also adapt their behaviour by analysing how the environment is affected by their previous actions.*** (High-Level Expert Group on Artificial Intelligence, 2019)

From this articulation, assumed as a **working definition for the study**, the essential concepts that can be retrieved include the acknowledgment of AI systems as agents that (1) acquire data from their (digital or physical) environment, (2) analyze and (3) process them, and finally (4) act within the same environment to achieve the (0) complex goals for which people expressly designed them. Not explicitly, they also introduce ML systems' capability to adapt over time by analyzing the environmental reactions to their actions.

Indeed, being a subset of AI, Russell and Norvig (2020) define **ML systems as a particular kind of agents, embedded in computers, that are able to improve their performances by making observations of the world.** Once again, identifying ML systems as rational agents allows one to distance from informal and misleading indications. While Arthur Samuel (1959), one of the pioneers of ML, talked about it as *"the programming of a digital computer to behave in a way which, if done by human beings or animals, would be described as involving the process of learning,"* the definition evolved into a sort of mathematical formula as Mitchell (1997) dictates: ***"A computer program is said to learn from experience E with respect to some task T and some performance measure P, if its performance on T, as measured by P, improves with experience E."***

Evidently, over time, the necessity to evolve early naive theories into more rational and unequivocal explications prevailed. The present thesis completely embraced the **rational agent interpretation** as an enlightening node for the translation of ML.

Nonetheless, the underlying ambiguity of artificial intelligence still manifests reminiscences of human mimicking concepts in various explorations that still circulate nowadays. For instance, IBM (2019) outlines AI systems as *"simulating human intelligence and thought processes,"* while the entry that the Oxford Dictionary provides reads: *"the theory and development of computer systems able to perform tasks normally requiring human intelligence, such as visual perception, speech recognition, decision-making, and translation between languages."* And the more one approaches non-authoritative sources, the more this phenomenon is amplified, paving the way for dangerous misconceptions, especially – but not only – when they get to non-experts.

3.1.2 Clarifying distinctive and misleading concepts

The interdisciplinary transfer of ML-related knowledge to the design field should identify and emphasize the unique features of ML systems. Still, much attention is needed to the language used and the possible misunderstandings that can arise. For sure, the metaphorical parallel between humans and AI can be an intuitive and useful instrument to simplify complex and not yet well-defined issues about this technology, but Johnson and Verdicchio (2017) highlight that it is a semantical trap that can trick both non-experts and AI researchers themselves. The latter, in fact, may attribute different meanings to words of common use when these are applied to AI and ML systems, which make them responsible, to a certain degree, for the way their work is communicated and understood by non-experts in the private sector, policymakers and those affected by this technology in their daily lives.

According to Johnson and Verdicchio (2017), it is not only a matter of terminology, but it indicates a more profound semantic gap, which might cause miscommunications that can influence future research in the field. The problem emerges especially in relation to qualities that usually describe people but, when referred to AI systems, imply concepts that cannot be directly transferred in everyday language, and autonomy is the emblem. Undoubtedly, it is the most remarkable trait distinguishing AI and ML systems from traditional programs. In the field of AI, **autonomy** is a metaphor for expressing the capability of these systems to perform a task without step-by-step programming. For example, when dealing with natural language processing (NLP), writing precise instructions and rules for determining how a machine should formulate a response would be an immense effort and almost impossible task for their developers. What ML systems do is instead acquire thousands of examples of text (perceiving their digital environment), identify the patterns that constitute the basis of language, and *autonomously* infer the rules (processing) to produce a satisfying result (acting) for the goal they have been given. So, the autonomy of AI and ML systems lies in the fact that human programmers must determine (i) an objective, identify (ii) a possible model for the AI system to be able to perceive its environment, and (iii) provide it with a comprehensive picture of the context itself (e.g., with a huge set of examples). In these terms, the autonomy of AI is quite far from the concept of *"[1] the right or condition of self-government, or [2] freedom from external control or influence; independence"* (Oxford Dictionary), as well as from all the inferences of machine uprising that science fiction and media portray. AI and ML systems actually depend on people at every step of the process (Fig. 3.2). Current AI-related narrative lacks a clear and open depiction of the roles played by human actors, from programmers to users, including deployers, investors, decision, and policymakers. The discussion is often affected by *"sociotechnical blindness,"* as (Johnson and Verdicchio (2017) put it.

And it also extends to other key qualities that separate ML systems from any technological artifact, such as **adaptability**. Being the capability of ML systems to evolve over time by analyzing how the environment responds to their actions within it and trying to improve their performances based on past experience, it is not entirely a synonymous with unpredictability and black box. Once again, the context in which ML systems work is generated, and sometimes even controlled, by people. If undesired and unexpected outcomes emerge, it is not because of the free will of technology but most likely a manifestation of problems in the human domain. Thus, blaming the monster instead of the creator (Latour, 2011) is a sterile attitude. Acknowledging the role of people in creating, selecting, and providing data and instructions to ML

Fig. 3.2 | Human intervention in the development and deployment of AI systems.

systems, these can become much more predictable than they are advertised to be (Johnson & Verdicchio, 2017). This awareness can even give rise to initiatives to prevent or limit negative impacts by acting on what is in total control of humans. It is the case of Interpretable ML (Molnar, 2019), which works to modify the modalities in which models are built, also trying to identify the reason why the outcomes do not correspond to people's predictions.
Overall, this aura of misunderstanding built on poor communication and half-truths inevitably creates the myth of AI as magic or a monstrous entity. It is a natural reaction to what is felt to be unknown or over-optimistic claims (Dove & Fayard, 2020), but technological solutions and change will always be the consequences of human actions, not some magical AI intervention (Kulesz, 2018). This is why **a thorough reflection on how things are presented, at the level of language, contents, and consciousness about human roles, should be at the base of an adequate translation of ML-related knowledge to non-experts.** An additional focus might be on making designers aware of their agency and their responsibility in steering the development and behavior of AI and ML systems, no matter how their qualities are commonly portrayed.
Nevertheless, besides properly grasping the nature of these systems, the inquiry proceeded towards identifying the level of knowledge design students should acquire to effectively deal with ML systems.

3.1.3 Theoretical foundations and their suitability for designerly comprehension

With variations in length, details, and clarity of explanations, ML theory is often presented according to the same categorization of types of learning. Namely, *supervised learning* requires that ML agents identify a function to get from input to output by observing examples of input-output pairs provided by people; in *unsupervised learning*, the agents are required to learn patterns without specific feedback; while *reinforcement learning* is based on a trial-and-error approach, and learning happens, like in animal training, by giving rewards or punishments to the agents (Russell & Norvig, 2020). Then, the discourse moves into increasingly technical details that go from learning problems like classification, regression, and clustering (more details about them will be provided in the following) to learning methods and formulas that draw much from other disciplines, like statistics.
Even though no generalizations can be made as some design universities offer more technical programs than others and design students might be more or less mathematically inclined according to their personal preferences, it is safe to assume that it is not the language design education uses to prepare their reflective practitioners (Schön, 1983). As a matter of fact, beyond this high-level theoretical construction, the practical applications and implications of ML, which might be of utmost interest for design as an applied discipline, are still ill-defined as ML represents a substantially different paradigm for traditional programmers. As Russell and Norvig (2020) admit, *"We are still in the early stages of defining a methodology for ML projects; the tools and techniques are not as well-developed."* Informal conversations with professors of information engineering further confirmed that very little practical knowledge is taught in ML courses, and no emphasis is placed on the actual comprehension of what lies beneath formulas.
Moreover, as anticipated in section 1.4, the professional experiences of designers working with AI and ML proved that formal conceptualizations – like supervised or unsupervised learning – are not helpful for communicating with experts in real life. They mainly treat ML as a black box, in which they are only interested in grasping what it can or cannot do (Yang et al., 2018). Designers are also used to working with simplified tools that require little or no programming skills, like platforms or apps to create websites, mock-ups, interactive prototypes, etc. In this way, their education and practice can focus on what they excel at, like designing for people's experiences and leaving technicalities to experts.
Therefore, the study of ML with its disciplinary perspective resulted in **no materials for a direct translation**. Indeed, a replica of an already demanding ML curriculum would be of no use to designers, who should collaborate with ML experts and complement their knowledge and skills, not substitute them in their job.
As Yang and colleagues (2018) framed well, designers need abstractions, exemplars, and simple insights about ML capabilities. They should be provided with a different, simplified narrative to grasp ML to the core and, possibly, unleash novel, innovative interpretations. Then, a practical exploration of communication strategies for non-specialist outreach was undertaken to understand the essential ingredients to enable this process.

3.2 Non-technical ML knowledge dissemination

3.2.1 Case studies collection strategy

The inquiry on how ML is communicated within its field gave the researcher a general background on the topic. Still, it did not produce relevant insights to guide the translation. Therefore, a broader perspective, including ethical concerns, was necessary for a more comprehensive understanding of the subject matter.
Consequently, to gather more information about practical issues and possibilities for bringing ML-related knowledge to a different audience, the context of exploration should change, and more empirical research could help. Then, an exploratory case study analysis (Muratovski, 2016; Robson & McCartan, 2015) was conducted to better understand how ML is currently disseminated to a non-expert public. For this activity, the author played both the role of the researcher and the recipient of the outreach instruments under investigation, as her background is consistent with the intended target audience of the doctoral research (bachelor's and master's degree in interior design). Inevitably, the strategy highlights the subjectivity of the inquiry, as the outcome reflects only the researcher's perspective. Yet, the exploration was meant to acquire a first-hand starting point for building hypotheses (in the reflecting phase of the constructive approach) to be tested later in the iterative research process with representatives of the potential target of the knowledge translation.
A web-based desktop research was conducted to identify different case studies of tools for disseminating AI and ML-related knowledge to non-experts, especially those targeting designers or the general public. The aim was to collect quick and

easy-to-navigate examples, differing in format and target audience, to have a good picture of the possibilities. Sampling saturation was reached in 16 case studies. Items with similar characteristics to those already analyzed were not included as they were not adding information to the research. For this reason, it does not provide a comprehensive map of the existing tools nor a selection of the best ones.

For each case study, qualitative content analysis and auto-ethnography were performed, and the following data were retrieved: author or provider (*who*), title (*what*), key concepts portrayed, target audience, format or tools employed, and personal notes.

The researcher manually conducted the content analysis due to the varied nature of the case studies. For the same reason, it included a first cycle of *descriptive coding* to capture the basic topics covered in the case studies and a second *focused coding* intended to categorize coded data according to thematic similarity (Saldaña, 2009). The auto-ethnography activity, instead, aimed at taking notes about the case study's appropriateness of language for designers, particularly (un)effective communication strategies and helpfulness for potentially integrating AI/ML capabilities in a project.

3.2.2 Case study analysis. Results and discussion

Table 3.1 synthesizes all the information gathered throughout the investigation. The majority of the retrieved case studies (eight) is meant for a general audience; five are suitable for any kind of designer; one specifically targets service designers; one addresses UX professionals and project managers; and one is open to anyone who wants to engage in basic ML programming. The latter was included as it requires no technical background and only gives a general overview of how ML systems work, with no ambition to train future programmers. However, most of the available educational resources are specifically intended for aspiring computer scientists or engineers. Though, these are not relevant to the inquiry because, despite the simplification of contents, they retain the structure of disciplinary courses.

Almost all the selected case studies are available online at the indicated links. Only the physical book *Bestiario di intelligenza artificiale* (Ammagamma, 2021) has partial contents on the authors' website. Overall, they present a diverse range of formats for disseminating AI and ML-related knowledge: from the most traditional but carefully curated book, web-based content, and talks to more interactive online courses, games, toolkits, engaging experiments, and videos, also produced in a docuseries. The variety implies different levels of involvement in the learning process. YouTube original series *The Age of AI* (YouTube Originals, 2019) is the most emotionally engaging, while the online courses by Kaggle and Reaktor in partnership with the University of Helsinki (Kaggle, n.d.; Reaktor & University of Helsinki, 2018), and the online IBM games (IBM, 2018) inevitably require active participation.

Interestingly, the content analysis revealed that, regardless of the differences, all case studies used concrete examples to portray AI and ML systems, which could be an index for a universally and transversally comprehensible language. In some cases, these were used as first-person experiences (IBM, 2018; WIRED, 2021) to directly grasp functioning or ethical concepts, while in (YouTube Originals, 2019), examples prominently drive the narrative and lead to the introduction of AI-related explanations.

Additionally, the second cycle of *focused coding* was useful for synthesizing the presented contents in twelve general categories – namely, *AI/ML Applications, Capabilities, Demystification, Programming references, Definitions, Limitations, Values, Design suggestions, Qualities, Prototyping, Implications, and UX* – of which more precise information can be found in Tab. 3.2. Most frequently, the analyzed case studies included references to some technical or programming-related content, revealing the strong link with the practical development of AI and ML systems. This was not necessary for Google Experiments (Google, n.d.), as they do not have explanations, and for the resources that proved to be most design-oriented (Piet, 2019)'s toolkit and the talks by (Clark, 2019; Holbrook & Lovejoy, 2017). The latter was added as it only includes *label*, which is borderline as a technical keyword.

Other recurrent topics are ML limitations and capabilities, immediately followed by applications, qualities, and values. All give relevant information about what AI and ML systems can and cannot do and, for this reason, are essential elements in becoming familiar with this technology. However, some clarification is due. In particular, ML Applications and Capabilities are challenging to differentiate. The former implies the operationalization of ML for a specific function in a real context. Instead, the latter portrays the current skills these systems have, but these often overlap, as in the case of *forecasting algorithm* or *marketing recommendation*, where they are indistinguishable.

Moreover, Qualities identify all the inherent characteristics of AI and ML systems like *adaptability, autonomy,* and *probabilistic* nature. Limitations and Values respectively indicate the weaknesses they have at the moment and the benefits they can generate if responsibly designed. Definitions and Demystification are additional ways to clarify AI and ML notions, especially in relation to human beings, and they appear in about half of the case studies. Not surprisingly, all the cases targeting designers also provide design suggestions of various kinds, from supporting a human-centered perspective to understanding how to approach the challenges of this technology. Finally, only few examples explicitly tackle issues related to AI and ML Implications, Prototyping, and UX, reflecting their inherent complexity and non-central role for an introductory understanding.

Helpful indications for the designerly translation emerged from the auto-ethnography annotations. Even though they did not employ particularly original formats, Clark (2019) and Holbrook and Lovejoy's (2017) argumentation was the most effective. In both cases, design was the protagonist of a perspective shift, demoting AI and ML as tools to be carefully molded. Not by chance, the authors are professional designers. Most other cases use familiar and immediate language, but even the attractive illustrations fail to equal the effectiveness of these ones. Uniquely presenting practical examples like Google's AI Experiments (n.d.) or IBM's games (2018), instead, seems to limit the opportunity to have a more holistic grasp, but a thorough explanation was probably not their explicit intention.

In terms of communication strategies, three are the main paths not to follow. The first concerns the *breadth of content*. Intuitively, focusing on a narrow perspective gives limited means for designers to exploit, and often it is combined with a closer development perspective that is not necessarily helpful in envisioning new applications. At the same time, *trying to cover multiple facets* of this complex discipline

Who	What	Target	Format/Tools	Addressed categories	Key Concepts	Concrete examples
Ammagamma	Bestiario di intelligenza artificiale [Artificial intelligence bestiary]	*General public*	• Illustrated book	Applications, Capabilities Demystification, Programming references	#0; #1; #forecasting algorithm; #marketing recommendation engine; #image analysis; #natural language processing; # optimized planning; #model predictive control; #robotic process automation; #business intelligence; #if; #intelligence; #human intelligence	Yes
Futurice	The Intelligence Augmentation design toolkit	*Service designers*	• Toolkit (cards, canvases)	Definitions, Applications, Capabilities; Limitations; Values; Programming references; Design suggestions	#ML definition; #ML use cases (#predict; #personalize; #recognize; #uncover structure); #designing for failure (#confusion matrix); #designing for learning (#feedback loop; #interaction loop); #designing for the worst (#training data diversity; #filter bubble); #dictionary (#algorithm; #AI; #artificial neural network; #big data; #classifier; #data mining; #deep learning; #feedback loop; #intelligence augmentation; #ML; #model; #negative example; #pattern recognition; #positive example; #predictive analysis; #training data); #customer identification; #human needs & values; #customer jounrey map; #smart service concepts (#ML touchpoints; #ML interaction - #predict; #recognize; #personalize; #uncover structure; #unexpected bugs - #bias in bias out; #loose cannon; #filter bubble; #echo chamber; #chucky; #network crash; #pretty bird; #breach; #flat line; #hydra; #uncanny valley-; #smart service storyline; #confusion matrix; #great power great responsibility)	Yes
Google	AI Experiments	*General public*	• Concrete and interactive examples of AI applications	Capabilities; Applications	#writing; #drawing; #learning	Yes
Google Design	Design is [Smart]	*Designers*	• Video / talk	Applications; Capabilities; Demystification; Limitations; Prototyping; Design suggestions; Qualities; Values; Programming references	#application examples; #features; #prediction (#classification; #regression; #clustering; #sequence prediction); #supervised learning; #overfitting; #confusion matrix (#true/false positive; #true/false negative); #unsupervised learning; #reinforcement learning (#actor; #environment; #reward); #myths (#AI==ML; #AI monolith - specific piece of a larger system; #Human-in-the-loop - AI needs humans at every step; #Data aren't neutral; #no perfection - often not even best UX); #human centered design (#human needs; #uniqueness; #cost; #WOZ; #precision - recall trade-off; #co-learning; #adaptation; #longitudinal research; #labels; #efficiency; #ML as creative process); #responsibility; #feedback loop; #inclusiveness (#differences)	Yes
Google PAIR	People + AI Guidebook	*UX professionals and product managers*	• Website + worksheets + glossary + explorables + tools	Applications; Implications; Values; Design suggestions, Limitations; Capabilities; Programming references; Qualities	#applications and implications; #AI adding value; #right expectations; #benefit vs technology; #errors; #good data practices; #tradeoffs; #transparency; #safety; #familiarity; #third-party sources; #model confidence; #understanding vs completeness; #automation; #feedback; #supervision; #agency; #dataset; #reward function / binary classifiers; #training datasets; #people's mental models; #adaptivity; #co-learning; #explainability; #trust; #confidence level; #control; #user needs + defining success; #data collection + evaluation; #mental models; #explainability + trust; #feedback + control; #errors + graceful failure	Yes
Helsinki University & Reaktor	Elements of AI	*General public (education of 1% of EU citizens)*	• Free online course	Applications; Implications; Definitions; Qualities; Demystification; Programming References; Values	#broad applications and implications; # AI definition (#autonomy; #adaptivity); #misleading "suitcase" words; #game tree; #minimax algorithm; #probability; #naive Bayes classification; #ML; #GANs; #Nearest neighbor classifier; #Regression; #Neural networks; #CNNs; #future&society	Yes
IBM	IBM Design for AI	*Designers & developers - techincal guides for non-technical people*	• Website + links to external resources	UX; Design suggestions; Qualities; Capabilities; Values; Definitions; Applications; Programming references; Limitations	#new interaction modalities; #role of design (#purpose; #value; #trust); #characteristics of AI (#understands; #reasons; #learns; #interacts); #meaningful relationship for human enhancement (#initiating; #experimenting; #intensifying; #integrating; #bonding); #AI ecosystem; #design thinking; #AI definition; #AGI; #narrow intelligence / weak AI; #AI use cases (#accelerate research & discovery; #enrich interactions; #anticipate and preempt disruptions; #recommend with confidence; #scale expertise and learning; #detect liabilities and mitigate risk; - ML tasks #computer vision; #NLP; #natural language understanding; #text to speech; #robotics; #ML); #example based; #ML definition; #learning; #supervised; #unsupervised; #reinforcement learning; #deep learning & neural networks; #data; #data collection (#sampling; #acquiring data; #data completeness); #data organization (#consolidating data; #consistent data; #data richness); #de-biased data; #AI design ethics; #accountability; #value alignment; #explainability; #fairness; #user data rights; #conversation	Yes
IBM AI Research	Learn + Play	*General public*	• Games	Applications; Programming references; Demystification; Limitations; Capabilities	#WordBot; #CLEVER score; #human intervention; #backdoors; #biases; #image recognition; #activation clustering; #conversation	Yes
Josh Clark	AI is Your New Design Material	*Designers*	• Lecture	Design suggestions; Qualities; Values; Limitations; Capabilities; Applications; Demystification	#role of designers with ML; #first generation new tech; #consistent use of tech (#things machines are uniquely good at); #amplification of human potential; #ML as a design material; #strenghts & weaknesses of ML; #pattern recognition; #recommendation; #prediction; #classification; #clustering; #generation; #casual uses of ML; #low-level production work (#time-consuming; #repetitive; #detail-oriented; #error-prone; #joyless); #questions to trace opportunities (#smarter questions; #new questions; #unlock new sources of data; #source invisible patterns); #"grain" of ML; #machines are weird (#insights from the unexpected; #set expectations and channel behavior accordingly); #narrow domains; #opaque logic; #people mental model; #transparency; #agency; #probabilistic; #signals vs absolutes; #machines reinforce normal; #bias; #values & behavior shift; #tech not neutral; #time for wild ideas	Yes
Kaggle	Intro to Machine Learning	*Programmers wannabes*	• Free online course	Programming references; Limitations	#models; #Decision Tree; #dataframes; #model building (#define; #fit; #predict; #evaluate); #underfitting; #overfitting; #Random Forest	Yes

Tab. 3.1.a \| Synthesis of the case studies analysis of ML outreach strategies.

Who	What	Target	Format/Tools	Addressed categories	Key Concepts	Concrete examples
Lingua Franca	Polytopal	*Designers*	• Web-based Handbook, • Principles and Elements	Definitions; Design suggestions; Programming references; Limitations; UX; Prototyping; Values; Qualities;	#AI definition; #ML definition; #(right) problem selection; #observing human behavior; #data collection; #labeling & construction; #authentic &synthetic data; #visualization & exploration; #design tradeoffs (#coverage; #salience; #transparency; #malleability; #instrumentality); #human interaction; #choosing technologies; #errors; #feedback & self-improvement; #prototyping; #wizard of oz; #iterative design; #agency; #accountability; #ethics; #architecture (#learning; #accuracy); #dynamics (#co-adaptation); #intuition (#trust); #embodiment (#human-like); #augmentation; #errata (#probabilistic systems) ; #bias (#in data collection; #in user experience; #complex feedback cycles; #as social responsibility); #transparency (#explainability); #assortment; #candidate; #clarification; #comparison; #correlation; #evidence; #forensics; #guard rails; #history; #intent; #latent space; #mark; #model card; #multi-modal; #override; #re-engagement; #signal; #variadic; #verifier; #warm-up	Yes
Nadia Piet in collaboration with MOBGEN | Accenture Interactive	AI meets Design toolkit	*Designers*	• Toolkit (cards, canvases)	Definitions; Qualities; Process; Capabilities; Design suggestions; Limitations; Values; Implications; Prototyping	#AI definition; #ML definition; #narrow AI; #general AI; #super AI; #deep learning; #reinforcement learning; #supervised learning; #unsupervised learning; #ML process; #ML tasks (#regression; #clustering; #classification; #dimension reductions; #testing & matching; #association rules; #multivariate querying; #density estimation; #GANs); #user-centered problem solving; #tech-driven opportunity spotting; #data-driven opportunity spotting; #impact matrix; #value proposition design ; #assessing feasibility; #framing task + plotting model; #probability; #user research & feedback; #mental models; #success & failure; #ethical & experiential concerns; #testing; #user trust & transparency (#explainability; #expectations; #failure & accountability); #user autonomy & control (#machine teaching & user feedback; #user controls & customization; #data privacy & security); #value alignment (#computational virtue; #bias & inclusivity; #ethics & (un)intended consequences)	Yes
Oxford Internet Institute in partnership with Google	The A-Z of AI	*General public*	• Illustrated guide	Definitions; Limitations; Capabilities; Programming references; Applications; Values; Demystification; Qualities	#AI fundamentals (#artificial intelligence; #bias; #datasets; #machine learning; #Turing test); #making AI (#GANs; #Human-in-the-loop; #image recognition; #knowledge; #learning; #neural networks; #open source; #quantum computing; #speech recognition; #zeros & ones); #society & AI (#climate; #ethics; #fakes; #journalism; #you); #using AI (#predictions; #robotics; #uses; #virtual assistants; #Watson; #x-ray)	Yes
R2D3	A Visual Introduction to Machine Learning	*General public*	• Visual representation	Qualities; Definitions; Programming references; Limitations	#statistical learning; #features / predictors / variables; #training; #machine learning; #decision tree (#fork; #branches; #split point; #tradeoffs; #recursion; #accuracy; #test); #overfitting	Yes
Wired	Computer Scientist Explains Machine Learning in 5 Levels of Difficulty	*General public*	• Outreach video	Definitions; Qualities; Demystification; Capabilities; Programming references	**Child:** #ML definition; #patterns; #examples; #learning; #human vs machines strengths (#amount of data; #number of examples to learn) | **Teen:** #recommendation; #predictions; #human vs machines strengths (#amount of processable data; #number of examples to learn; #creativity vs #past experience) | **College student:** #feature engineering; #supervised learning; #unsupervised learning; #reinforcement learning; #deep learning; #still defining what good ML can be built	Yes
Youtube	The Age of AI	*General public*	• Docuseries	Definitions; Demystifications; Qualities; Capabilities; Applications; Values; Programming references; Limitations	**1. How far is too far** #AI definition; #demystification (#generalized AI); #human like AI; #object recognition; #emotion AI; #digital avatar; #human-AI collaboration; #ML definition; #human enhancement; | **2.Healed through AI** #speech recognition; #image recognition (#recognize medical conditions); #voice synthesis; **3.Using AI to build a better human** #learn from experience; #reinforcement learning; #trust; #human enhancement and augmentation; | **4.Love, art and stories: decoded** #AI as human-like companion; #computer vision; #generative algorithms; #creativity; #deep learning language model; #self-driving cars (#OxTS Global Positioning System; #Lidar sensors; #Vehicle 2 Vehicle communication); | **5.The "Space Architects" of Mars** #generative design; #AI farming; | **6.Will a robot take my job?** #automation; #common sense; #taking over tasks that people don't want to do (#burdensome; #risky); #robots; #big data; #efficiency; #forecasting; #optimization; #AI + humans; | **7.Saving the world one algorithm at a time** #preservation; #pattern recognition; | **8.How AI is searching for aliens** #anomaly detection; #good values	Yes

Tab. 3.1.b | Synthesis of the case studies analysis of ML outreach strategies.

only by giving unconnected hints results in a scattered overview that risks being unhelpful. The third direction to avoid *echoes the deceiving relationship between AI and human beings*, which proves faulty in giving a solid idea of what AI and ML truly are.
On the contrary, immediate examples and activities (Kaggle, n.d.; Reaktor & University of Helsinki, 2018) as well as explicitly bridging AI and ML to the design discipline and process (Clark, 2019; Holbrook & Lovejoy, 2017; Piet, 2019), are successful strategies towards enabling designers to master this technology, also with a view to its integration into a project. However, I personally felt that YouTube Originals's docuseries (2019) was the most compelling example of outreach, also considering designers as a target. Indeed, it achieves to explain what AI is in a comprehensible way and pays attention to discourage misconceptions. Most importantly, it also manages to sensitize and inspire the audience with thoroughly presented use cases that demonstrate how AI and ML development can be steered towards good impacts. This is particularly relevant because abstract dissertations about what should be done to have responsible and ethical solutions in the future are not required when you have concrete proof that it is already possible.
Summing up, this research activity proved helpful in getting a comprehensive picture of possibilities that set the ground for reflections about the contents and the modalities of the translation of ML to design education.

Categories	Frequency (out of 16 case studies)	Included codes
Applications	11	#forecasting algorithm; #marketing recommendation engine; #image analysis; #natural language processing; #optimized planning; #model predictive control; #business intelligence; #writing; #drawing; #learning; #application examples; #narrow domains; #design thinking; #accelerate research & discovery; #enrich interactions; #anticipate and preempt disruptions; #scale expertise; #detect liabilities and mitigate risk; #robotics; #WordBot; #climate; #fakes; #journalism; #virtual assistants; #Watson; #x-ray; #digital avatar; #recognize medical conditions; #human-like companion; #self-driving cars; #generative design; #AI farming; #automation; #burdensome activities; #risky activities; #preservation
Capabilities	12	#recommendation; #image analysis; #image recognition; #natural language processing (NLP); #natural language understanding; #optimized planning; #automation; #personalize; #uncover structure; #pattern recognition; #personalize; #uncover structure; #writing; #drawing; #learning; #prediction; #classification; #regression; #clustering; #sequence prediction; #understanding; #reasoning; #interacting; #ML tasks; #computer vision; #text to speech; #conversation; #generation; #ML tasks; #dimension reduction; #testing & matching; #association rules; #multivariate querying; #density estimation; #GANs; #Turing test; #speech recognition; #human vs machines strengths; #number of examples to learn; #human vs machines strengths; #amount of processable data; #creativity vs past experience; #object recognition; #emotion AI; #human-AI collaboration; #voice synthesis; #self-driving; #generative design; #efficiency; #forecasting; #optimization; #anomaly detection
Definitions	9	#AI definition; #ML definition; #dictionary; #deep learning; #neural networks; #artificial general intelligence (AGI); #super AI; #narrow intelligence; #weak AI; #reinforcement learning
Demystification	8	#intelligence; #human intelligence; #myths; #AI==ML; #AI monolith; #Human-in-the-loop; #Data neutrality; #no perfection; #misleading "suitcase" words; #human intervention; #machines are weird; #machines reinforce normal; #knowledge; #you; #human vs machines strengths; #amount of data; #number of examples to learn; #predictions; #amount of processable data; #number of examples to learn; #creativity vs past experience; #still defining what good ML can be built; #demystification; #generalized AI; #human like AI
Design suggestions	7	#designing for failure; #confusion matrix; #designing for learning; #feedback loop; #interaction loop; #designing for the worst; #training data diversity; #filter bubble; #customer identification;#customer journey map; #smart service concepts; #ML touchpoints; #ML interaction #smart service storyline; #great power great responsibility; #human centered design; #human needs; #longitudinal research; #right expectations; #benefit vs technology; #third-party sources; #model confidence; #understanding vs completeness; #supervision; #people mental models; #defining success; #user control; #graceful failure; #role of design; #initiating; #experimenting; #intensifying; #integrating; #bonding; #first generation new tech; #consistent use of tech; #ML as a design material; #trace opportunities; #smarter questions; #new questions; #new sources of data; #invisible patterns; #insights from the unexpected; #channel behavior; #time for wild ideas; #(right) problem selection; #observing human behavior; #choosing technologies; #tech-driven opportunity spotting; #data-driven opportunity spotting; #framing task; #plotting model; #user research; #machine teaching; #customization
Implications	3	#implications; #impact matrix; #(un)intended consequences
Limitations	12	#failure; #errors; #training data diversity; #filter bubble; #unexpected bugs; #bias in bias out; #loose cannon; #filter bubble; #echo chamber; #chucky; #network crash; #pretty bird; #breach; #flat line; #hydra; #uncanny valley; #overfitting; #underfitting; #confusion matrix; #true/false positive; #true/false negative; #cost; #precision-recall; #tradeoffs; #de-biased data; #backdoors; #weaknesses; #opaque logic; #bias; #tech not neutral; #coverage; #salience; #transparency; #malleability; #instrumentality; #data collection; #user experience; #complex feedback cycles; #social responsibility; #common sense
Programming references	13	#zeroes and ones; #if; #algorithm; #big data; #classifier; #data mining; #model; #negative example; #positive example; #training data; #labels; #good data practices; #dataset; #reward function; #binary classifiers; #confidence level; #data evaluation; #game tree; #minimax algorithm; #naive Bayes classification; #GANs; #Nearest neighbor classifier; #regression; #CNNs; #data collection; #sampling; #data completeness; #data organization; #consolidating data; #consistent data; #data richness; #authentic & synthetic data; #CLEVER score; #backdoors; #activation clustering; #models; #model building; #define; #fit; #predict; #evaluate; #Random Forest; #labeling & construction; #visualization & exploration; #architecture #errata; #assortment; #candidate; #clarification; #comparison; #correlation; #evidence; #forensics; #guard rails; #history; #intent; #latent space; #mark; #model card; #multi-modal; #override; #re-engagement; #signal; #variadic; #verifier; #warm-up; #GANs; #open source; #quantum computing; #decision tree; #fork; #branches; #split point; #tradeoffs; #recursion; #accuracy; #test; #feature engineering; #generative algorithms; #OxTS Global Positioning System; #Lidar sensors; #Vehicle 2 Vehicle communication
Prototyping	3	#prototyping; #wizard of oz; #iterative design; #testing
Qualities	11	#features; #supervised learning; #unsupervised learning; #reinforcement learning; #actor; #environment; #reward; #uniqueness; #co-learning; #efficiency; #ML as creative process; #adaptivity; #co-learning; #autonomy; #characteristics of AI; #AI ecosystem; #example based; #things machines are uniquely good at; #low-level production work; #error-prone; #joyless; #"grain" of ML; #signals vs absolutes; #self-improvement; #learning; #accuracy; #dynamics; #co-adaptation; #embodiment; #human-likeliness; #probabilistic systems; #ML process; #assessing feasibility; #statistical learning; #predictors; #variables; #training; #patterns; #learn from experience
UX	2	#new interaction modalities; #meaningful relationship for human enhancement; #human interaction;
Values	10	#intelligence augmentation; #human needs & values; #responsibility; #inclusiveness; #differences; #AI adding value; #transparency; #safety; #familiarity; #agency; #explainability; #trust; #user control; #future&society; #purpose; #meaningful relationship; #human enhancement; #AI design ethics; #accountability; #value alignment; #fairness; #user data rights; #amplification of human potential; #strengths; #values & behavior shift; #intuition; #augmentation; #value proposition design; #ethical & experiential concerns; #failure; #user autonomy; #data privacy & security; #computational virtue; #inclusivity; #ethics; #good values

Tab. 3.2 | Categories correspondences from content analysis.

3.3 Key elements for the translation

Reflecting on the outcomes of the disciplinary exploration of ML and its non-technical dissemination, this technology appears much less frightening than it is usually depicted. The key is to discern the fundamental information to grasp the essence of what ML really is among the many layers that constitute this subject matter, recognize that doing things the way they have always been done might not be the best strategy (especially if it leads to ambiguous results), and find a path towards simplicity. In line with Antonelli's (2018) view, the mission is to *"show that AI is not a monster from outer space"* by communicating it in the **simplest and clearest way possible, despite its internal complexity**. These are the premises for a designerly translation of ML, which should be configured as the latest tool designers have to make things better.
To reach this goal, though, the investigation proved that some systematization is needed. Hence, to answer RQ1 – *What can and has to be translated from ML and related ethics to design?* – a synthesis is done by identifying the core structure defining ML systems and their peculiar qualities and organizing related knowledge.

3.3.1 *Synthesizing the core of ML systems*

The connotation of **ML systems as agents** (Russell & Norvig, 2020) has meaning in both computer science and ethics disciplinary fields. It offers an unequivocal definition as opposed to the multiplicity of still ambiguous interpretations that refer to human likeliness as a measure for describing them. Keeping this in mind as an effective synthesis, just little essential information needs to be outlined to understand the core of ML systems. In particular, **(i) the inputs** the system needs to perceive its environment, **(ii) the outputs** expected for it to reach the given goal, and **(iii) the kind of processing** it might be doing to obtain the output given the input are required to identify the system as an AI agent in general. To precisely characterize it as a ML agent, then, one should also reflect on **(iv) whether it learns from experience**.
Additionally, to outline the bare minimum characterization of ML systems as agents, it would be necessary to recall Johnson and Verdicchio (2017) and van de Poel's (2020) definitions of **sociotechnical systems including AI**. Accordingly, the role of people should be defined and, at least, the **(v) need that the system addresses** must be identified. Implying both the motivation and the goal that people need it to achieve would reflect the basis of a human-centered approach but not a more comprehensive systemic perspective. The choice to keep the synthesis mainly functional would be justified in terms of simplicity. Indeed, this simplified overview would allow to effortlessly depict the basic constituents of a design solution integrating ML. In a sense, instead of perceiving ML systems as opaque entities, these elements are key to deciphering them in an intuitive and non-specialistic way.
For this reason, they have been framed in two sets of instructions: the *Decoder* and the *Encoder* (Fig. 3.3). The former can be used to *decode* or try to understand the functioning and make sense of an existing ML-infused product or service. It includes questions like: *Can you describe the task the system is performing? What may be the inputs? What are the outputs? Can it learn from experience? Can the system respond to a human need?* Analogously, the latter can help set out one's idea guided by similar questions. Namely, *What human need are you responding to? What task do you need the system to perform? Do you need it to learn from experience? What are the inputs of your system? What outputs do you expect?*

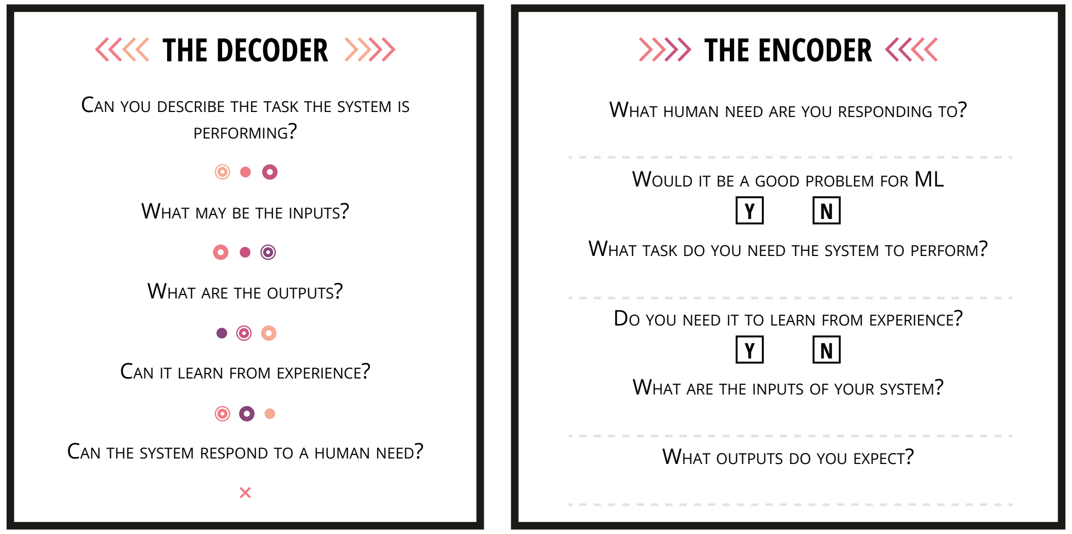


Fig. 3.3 | Decoder and Encoder.

An additional consideration could concern **whether it would be (or not) a good problem for a ML system to address**. However, this implies a deeper understanding of the qualities of this technology. Indeed, implementing ML requires more time, costs, and effort than traditionally programmed solutions. Thus, more careful reasoning at the beginning of the design process would allow these resources not to be wasted and more coherent solutions to be undertaken.
This is why understanding whether a defined problem suits ML's inherent characteristics should complement the above-illustrated foundational framing. At this scope, two parameters should drive the evaluation: the **value** the intervention might have for people and its **consistency** with ML capabilities, as synthesized in Fig. 3.4.

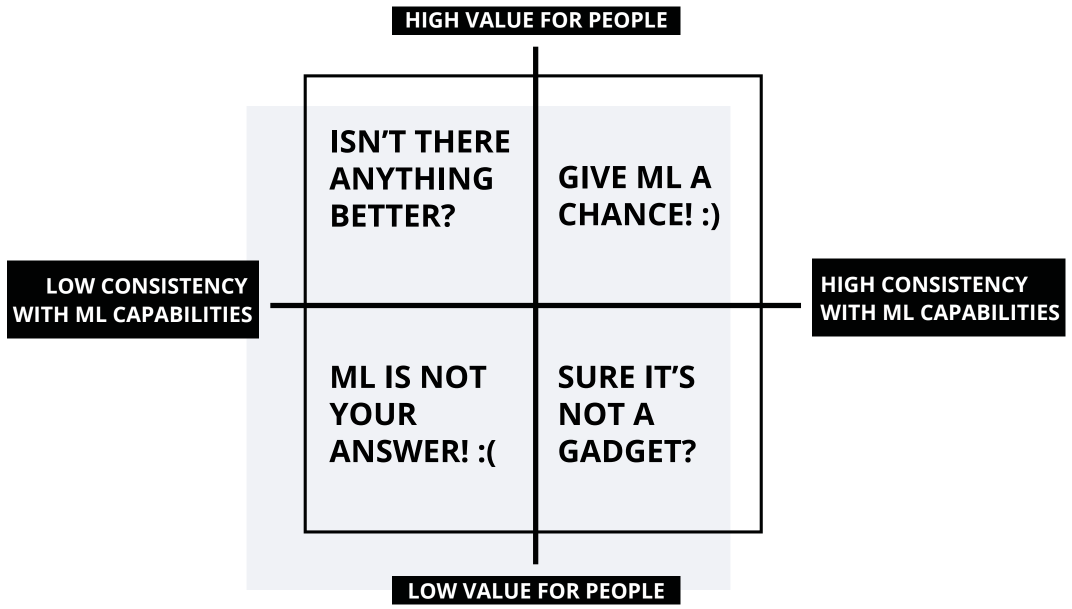


Fig. 3.4 | ML Suitability Matrix.

Intuitively, low consistency and low value for people suggest that better solutions can be found without ML; the low consistency and high value quadrant indicates that different technologies or non-technological interventions might be better; while high consistency and low value might depict the category of products and services that Levinson (1977) defines as toys or gadgets. Ideally, an idea would be worth pursuing when both high consistency and high value for people are met. To support the reflection about what is valuable and consistent in relation to ML systems, the main qualities and issues have been summarized (Fig. 3.5), using the lessons learned from the analyzed case studies as a reference (in particular: Futurice, 2017; Holbrook & Lovejoy, 2017; Clark, 2019; Piet, 2019). The collected insights do not constitute prescriptive guidelines but are food for thought to point to relevant concerns one should be aware of and address early in the design process.

VALUE FOR PEOPLE	HIGH	LOW
Nature of the task the ML system has to perform	Repetitive, burdensome, risky, time-consuming	Effortless, satisfying, enjoyable, enriching
Enabled possibilities	It inspires, empowers and/or augments human capabilities, enabling people to take action	It is superfluous, and doesn't add meaning to human life/ experience
Risk/benefits trade-off	Its foreseeable benefits highly overcome concerns	It may threaten important principles (human autonomy, fairness, and intelligibility) without commensurate benefits.

ML CAPABILITIES	CONSISTENT	NOT CONSISTENT
Step-by-step rules	Impossible or arduous	Possible, easy
Adaptable vs predictable	Adaptable and/or proactive	Consistent and predictable
Data for training	Great amount of examples available or easy to collect	Not enough examples are available, or hard to collect them
Actionability of the output	The output enables an action (decision)	The output enables only insights on data (predictions)
Relationship sought among the data	Correlation, mutual connections	Causation, why has something happened

Thumb rule: *If you cannot teach a human "intern" to do it, a machine cannot either! (for now)*

Fig. 3.5 | Reflection points for the ML Suitability Matrix.

TO SUM UP

The chapter addresses RQ1: ***What can and has to be translated from ML to Design?*** and looks at the ML discipline for answers.

- The concept of artificial intelligence eludes a precise definition, but the relationship with human beings inevitably shapes the origins of AI. Two are the most common interpretations. For some researchers, AI is synonymous with the simulation of human behavior. For others, it is a means to augment and amplify human potential.
- The currently prevailing interpretation defines AI systems as rational agents that (1) acquire data from their (digital or physical) environment, (2) analyze and (3) process them, and finally (4) act within the same environment to achieve the (0) complex goals for which people expressly designed them. As particular kinds of agents, ML systems are characterized by the capability to improve their performances with experience. These are also the working definitions driving the research.
- A disciplinary perspective on ML does not offer materials for direct translation for designers. While they need to understand what ML systems are capable of, computer science is full of technical indications and misleading concepts (like the semantical trap of the parallel between humans and AI, which also tricks AI experts).
- Then, an exploratory case study analysis of 16 communication strategies for non-specialists was conducted to understand better how ML is currently disseminated to a non-expert public. It highlighted that providing immediate examples and activities and explicitly bridging AI and ML to the design discipline and process can successfully favor designers' understanding but also sensitize and inspire them toward the good impacts that this technology might have. Instead, narrow perspectives on specific contents, scattered overviews of multiple facets of the topic, or echoing the deceiving relationship between AI and human beings should be avoided.
- The overall disciplinary investigation on ML brought the identification of fundamental concepts for understanding it. First, a **ML system** should be defined as an **agent**, which is **part of a broader sociotechnical system**. Thus, it can be recognized by the task it performs, its inputs, its outputs, and whether it learns from experience and meets human needs (as synthesized by the *Decoder* and *Encoder* tools). Additionally, one should be able to identify if ML is suitable to address a given problem, based on the value the intervention might have for people and its consistency with ML capabilities (as portrayed by the *ML Suitability Matrix*).

4. FRAMING ML KNOWLEDGE FOR TRANSFER

ML UX education and toolkits must provide new means that help UX teams develop a tacit understanding of ML - that is, not simply teaching how ML works, but empowering them with enough technical literacy to be able to ideate creatively yet practically, and to better collaborate with ML experts.

(Yang, 2017)

C H A P T E R F O U R

Once the basic contents for the translation have been identified, a structure and a language to communicate them need to be defined. Hence, ***how to frame ML knowledge for transfer?*** (RQ2).

A *Designerly Taxonomy of ML* was developed to **systematize ML knowledge** by combining designers and ML experts' approaches to designing ML systems, creating the basis for interdisciplinary communication (4.1).

To practically experiment with the theoretical assumptions and investigate the **most preferable forms to introduce ML knowledge to design students**, an online workshop – *Machine Learning Pills for Designers* – was organized in collaboration with a Digital and Interaction Design master's thesis student (4.2). As the combination of the four languages proposed (*definitions, metaphors, case studies, and practical examples*) could be most effective, *ML Agents* were elaborated as a new synthetic tool to translate ML capabilities: *classification, regression, sequence prediction, generation, clustering, and action selection* (4.3).

An ultimate layer for framing ML knowledge lies in the ethical compound. To support the design of responsible ML applications, a value-driven approach and a *Responsible Cycle for ML Design* are proposed, further information about relevant values risks and possibilities was gathered through a systematic exploratory content analysis of ethical guidelines for AI (4.4).

Finally, the foundational assumptions for the educational models are synthesized in two main requirements that ML solutions should have in addition to a human-centered perspective: **(i) consistency with ML capabilities, and (ii) responsibility** (4.5).

4

4.1 Systematizing an ill-structured discipline

The previous chapter highlighted disciplinary difficulties in presenting ML knowledge in an unambiguous way and the need for simplification to reach a non-expert audience. For the purpose of a designerly translation, the foundational contents identified in the previous chapter were used as references. However, also the structure and interconnections to functionally communicate them are important to define ***how to frame ML knowledge for transfer?*** (RQ2). Thus, the researcher, acting as a knowledge broker (Meyer, 2010), directed the investigation toward building a bridge across the disciplines involved.

4.1.1 Defining a method for the systematization

Per se, collecting recurring topics in the dissemination of ML-related knowledge is not enough to make it actionable. Indeed, Yang (2018) stated that teaching designers how ML works is insufficient to make it a design material. As Law (2002) – cited in (Meyer, 2010) – sustained, **translation is about connecting, moving, and shifting something to a new place, modality, or form, only retaining something**. This confirms that ML knowledge, as originally crafted in and for computer scientists, cannot fit with designers' necessities as it is, and theoretical elaborations for design education should introduce new elements to unlock ML potential as a design material. In her doctoral path, Yang (2018, 2020) identifies a clear research gap to enable a designerly understanding of ML systems. She underlines the need for a taxonomy based on a set of abstractions uniquely focused on matching user values and the contextual capabilities of this technology. This construct would differ from those elaborated in data and computer science and coincide with the purposes of this research.
These premises and the information gathered in the first stage of the constructive inquiry led to the development of the *ML Designerly Taxonomy*.
Kundisch et al.'s (2022) *Extended Taxonomy Design Process* (Fig. 4.1) is used as a reference to portray the driving method.

I. Problem identification and motivation. As introduced, the *ML Designerly Taxonomy* is intended to link ML knowledge and its application in the real world with a human-centered approach. It primarily addresses designers as actors that could join the ML discourse bringing a fresh and potentially innovative perspective, but – as a

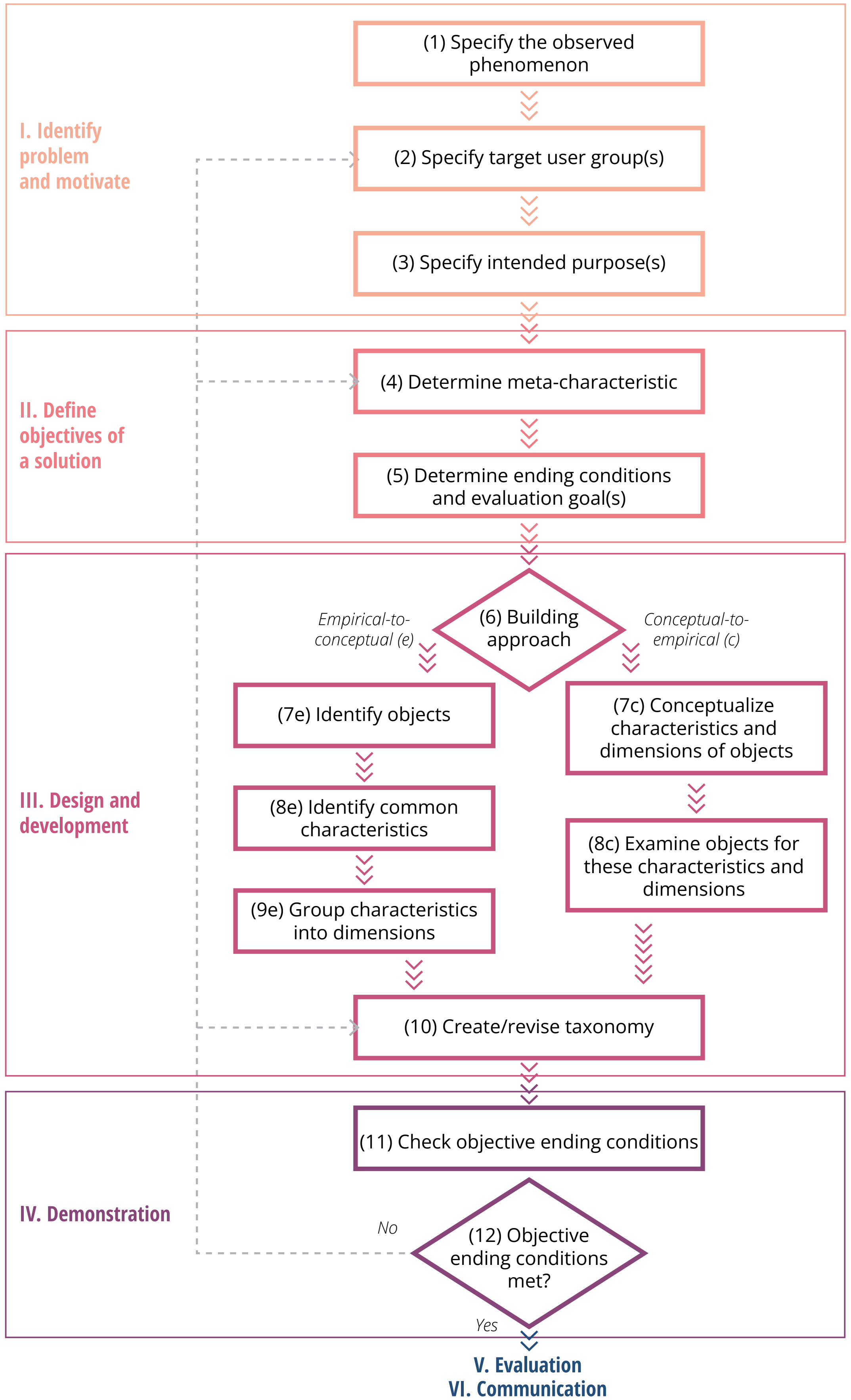


Fig. 4.1 | Adaptation of Kundisch et al.'s (2022) Extended Taxonomy Design Process.

bridging tool – it might also enable communication with ML experts. Indeed, dealing with a still undefined phenomenon, the theoretical constructs and relationships to be identified and structured through the taxonomy could be helpful in both fields.

II. Objectives definition. To combine designerly and technical approaches to ML, different levels of knowledge should constitute the meta-characteristics of the taxonomy. For instance, after identifying the challenges that designers can encounter when working with ML, Winter and Jackson (2020) synthesize them into categories reflecting different knowledge typologies. They list conceptual, operational, and technical knowledge, referring respectively to the capabilities and limitations of ML systems, the tuning of their models, and the evaluation of their performances. In addition to these ML programmers' issues, they point out aspects related to the UX, prototyping, and interaction of ML-infused products, labeled as development culture, design methodology, and interaction design. Although their work did not provide a unifying structure, they recognized the multi-faceted context that characterizes ML systems and inspired the construction of a more connected systematization.

To bring designers and ML programmers together, the *ML Designerly Taxonomy* expands the connection of *user values* and *ML contextual capabilities* – as suggested by Yang (2020) – to reach ML *system processing modalities*. In this way, it provides a familiar entry point to both intended targets as a preferable way to learn complex topics (D'Ignazio, 2022). On the one hand, it aims to depict *human values* as abstractions steering the intended impacts of real-world applications. On the other, *system processing modalities* frame the different technical approaches for implementing ML from a disciplinary standpoint, while contextual capabilities materialize both perspectives. The path from one side to the other is composed of a multilayered structure. From the top, the taxonomy includes the following levels: Conceptual Knowledge, to understand the potentialities of ML systems in relation to human capabilities; Designerly Knowledge, to identify the concrete opportunities designers have to exploit ML capabilities; Operational Knowledge, related to current applications of ML systems; Technical Knowledge, including the technical principles and processes underneath; and Operative Knowledge, implying the understanding or mastery of ML systems' functioning, which can extend to the fine-tuning of ML systems' models.

This theoretical scaffold is intended for both designers and ML experts to identify relevant applications of ML systems that are consistent both with the technology capabilities and people's values. Ultimately, the taxonomy would be successful if it is concise, comprehensive, and robust enough (subjective ending conditions) to determine how ML-infused solutions work at a very basic level and what they do to benefit people (evaluation goal).

III. Design and development. As the reason for the taxonomy development is to fill a gap for systematized knowledge, an empirical-to-conceptual approach has been selected. The same sources employed for the explorative analysis of disciplinary knowledge and the initial organization of the contents used for ML outreach (described in Chapter 3) have formed the theoretical basis for the systematization. In particular, for technical matters, Russell and Norvig's (2020) *Artificial Intelligence: A Modern Approach* remained the primary, most comprehensive and better organized

reference. Still, other texts and courses have also been used to corroborate their explanations or find alternatives. For instance, they include (Hebron, 2016; Google Developers, n.d.; Reaktor & University of Helsinki, 2020). Once the features of each level were outlined and clear, a targeted web search was conducted to identify any missing elements and ensure a complete overview.
Deductive and mainly inductive approaches have been employed, starting from the only universally agreed categorization of ML systems in supervised, unsupervised, and reinforced learning. Then, additional key elements, concepts, and qualities were collected and iteratively organized to fit the defined meta-characteristics levels and build connections. Before reaching a satisfying outcome, several operations of renaming, merging, splitting, swapping, and adding elements have been performed. The main doubts concerned the positioning of Conceptual or Designerly knowledge as extremity of the taxonomy; the internal organization of the Technical Knowledge level to create a consistent path from designers- and ML experts-related matters and vice versa; the categorization of Technical and Operative knowledge in a way that could be suitable for designers to understand current capabilities of ML systems; and the completeness of the depicted information. The results of this process are described in the next section (4.1.2).

IV. Demonstration. To check the formal validity of the taxonomy, the indications of Nickerson et al. (2013) were followed. First, the objective conditions for it to be considered a taxonomy were verified:

- all the elements included in the taxonomy were thoroughly examined and questioned in its latest version,
- it does not seem necessary to merge, split, change, or add any dimension or feature,
- at least one object can be classified under any dimension and feature,
- every dimension is unique,
- and the features are mutually exclusive and collectively exhaustive within their dimensions.

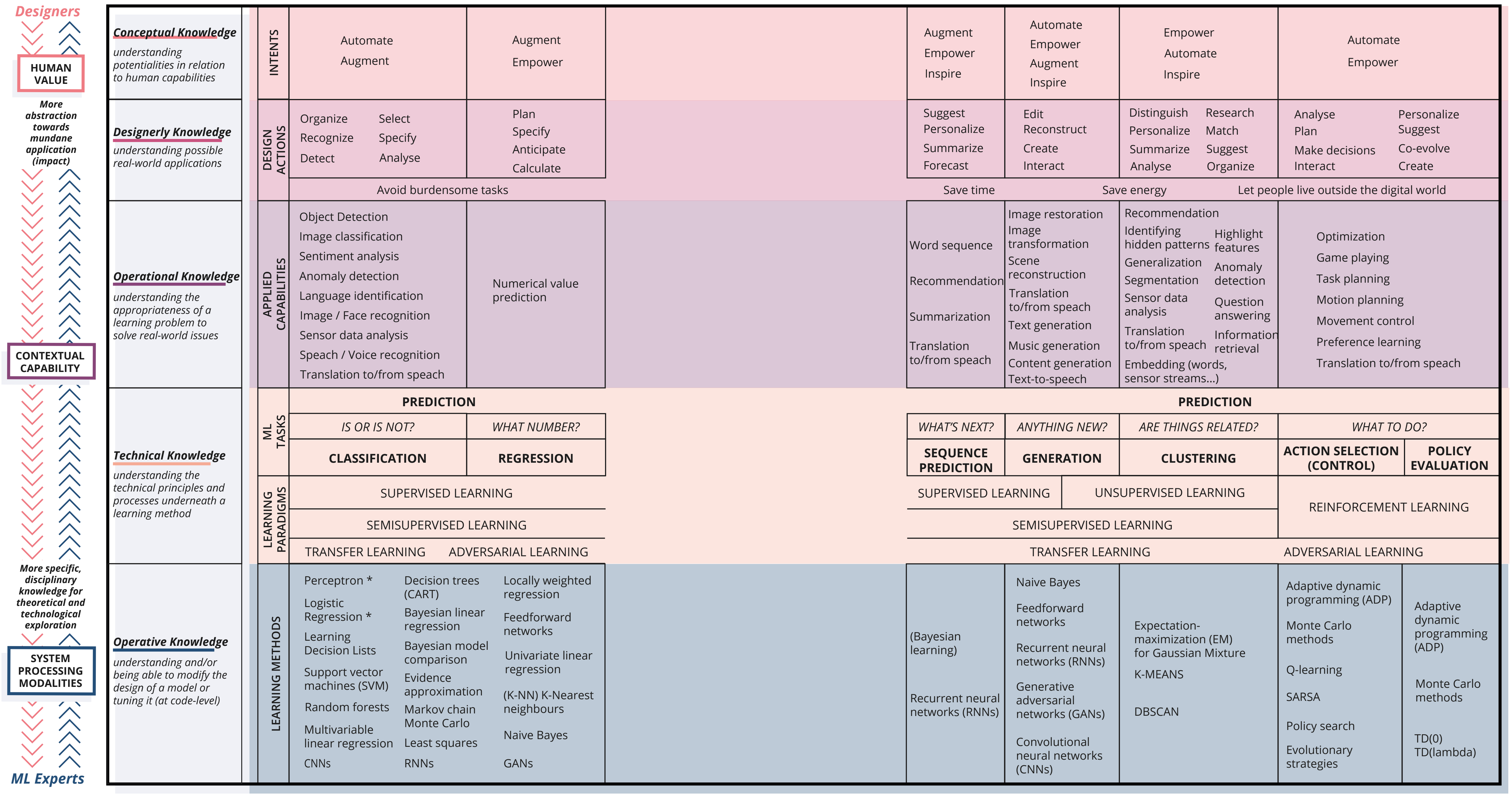
Nonetheless, it is not a hierarchical taxonomy but a faceted one (meaning that there is no need for a predetermined order of facets or dimensions). It has enumerative facets or mutually exclusive possible features (Glushko, 2020). So, in order to depict the relationships between the items in each level, in the final visualization (Fig. 4.2), they are repeated when needed. For instance, the Conceptual Knowledge level presents only four possible values (Augment, Automate, Empower, Inspire). As they are quite broad, they can apply to different ML capabilities (columns), and some cells might seem redundant.

Then, according to the fulfillment of subjective ending conditions, the taxonomy can be applicable. In fact, to the extent of the author’s subjective inspections, it resulted *concise*, presenting six dimensions; *robust*, portraying enough features to clearly differentiate different possibilities; and *comprehensive*, being able to classify random examples encountered. Moreover, even if they were not considered essential goals to achieve, the taxonomy can also be *extendible* (e.g., new dimensions or features might be added to include all AI systems possibilities, not just ML) and *explanatory* (without providing every detail of a ML-infused solution, they can give a hint about their nature, namely what they can do and how they do it in general terms).

V. Evaluation. To evaluate the effectiveness of the taxonomy, two strategies were adopted. The first, inviting experts for non-structured interviews to assess the taxonomy's usefulness, coherence, and quality, was not very successful. Indeed, without active relationships with ML engineers and UX designers having some experience with this technology, and complicit with the pandemic period, only one software engineer and one UX designer gave their availability to review the taxonomy. Overall, they found the approach interesting and with good potential, but few indications were given in terms of necessary modifications. Coming from an expert ML perspective, some suggestions for a more consistent classification within the Operative and Technical Knowledge were applied without major changes to the overall structure.
To compensate for the limitations of the first round of evaluation, a more practical approach has been identified. The taxonomy was operationalized as a theoretical structure underneath didactic activities and tools to be assessed while fulfilling its scope. An increasing number of taxonomy levels were included in workshops aimed at envisioning ML-infused solutions, and they will be presented in the following sections and chapters. Within the given timeframe and with the researcher's resources, it was possible to involve only design students. Then, the Operative Knowledge level still has to be tested in practice.

4.1.2 A designerly taxonomy of ML

The *ML Designerly Taxonomy* (Fig. 4.2) is ultimately a **synthetic theoretical construction** that has the potential to bring ML capabilities to practice **bridging human-centered and technical perspectives**, thus connecting designers and ML programmers. The link between *human values* and *system processing modalities* is constructed in five different levels of knowledge that constitute the meta-characteristics of the taxonomy. In turn, they take the form of dimensions (six in total) articulated into features, the main elements depicting a dimension.
At first, the taxonomy building consisted in gathering all relevant keywords that might represent the related knowledge level, to determine the most appropriate way to depict them. The construction process started from the only consolidated and recurrent differentiation of ML systems into supervised, unsupervised, and reinforcement learning, defined as *learning paradigms* by Russell and Norvig (2020). They respectively indicate systems that are given pairs of inputs and expected outputs (labeled datasets) to determine the inner function that correlates them; systems able to identify hidden patterns with no or little knowledge about how the output should look like (unlabeled datasets); and systems that are put into an environment and trained in a trial-and-error modality, with rewards and punishments instead of datasets. Additional variants, possibly transversal to any kind of ML system, are transfer learning (that consists in the application of the knowledge acquired from solving a problem to address a different but related one), adversarial



* classification only

Fig. 4.2 | ML Designerly Taxonomy.

learning (confronts one system with another that intentionally seeks to sabotage it). Semisupervised learning, in contrast, exploits the potential of supervised and unsupervised learning. What they all do is predicting something based on the perception of their environment.
ML systems can be further defined according to the outputs they are expected to reach or *learning problems* as Russell and Norvig (2020) call them. Basically, learning problems should represent what ML systems do to achieve their objective, but the terminology is misleading as it seems to shift the focus to the problem to solve instead of the capability. For this reason, having proved the term's ambiguity while discussing with colleagues, the author opted for *ML tasks* to name the related category. The specific definitions of the included features will be provided in section 4.3.
Generally, supervised learning is represented by classification and regression, while unsupervised learning by clustering. However, some ML systems present specific capabilities that distinguish them from the mentioned ML tasks, even

though they should formally fall under one of these categories. Hence, to create a more comprehensive taxonomy, able to differentiate current ML capabilities and applications more intuitively according to their underlying peculiar qualities, two ML tasks were added: sequence prediction (Holbrook & Lovejoy, 2017) and generation (Clark, 2019). The first is characterized by sequential and historical information as input, while the second produces new outputs based on the training examples. Density estimation could have been an additional unsupervised ML task, though, as it would overlap with classification and generation in terms of practical applications, it has been excluded. Moreover, two reinforcement learning tasks can be detected: policy search or evaluation and control or, more intuitively, action selection. In the former, the behavioral strategy is known. In the latter, the system has to learn it.

ML capabilities, declined as tasks, are a focal node of the taxonomy, as they determine both which models must be selected to enable them and what they can do in practice. Therefore, according to this categorization, further dimensions in both ML experts and designers' directions were populated. Operative knowledge was declined in terms of learning methods or the general functioning principle underlying ML systems. In fact, going into details about models or algorithms would have been overwhelming, as they are too specific and constantly evolving. While most current ML learning methods can be clearly differentiated based on their ML tasks, the distinction between classification and regression is blurrier. So, after the suggestion of the interviewed ML expert, the related learning methods are kept together, as they can be easily adapted to perform classification or regression tasks. Those that only apply to classification problems are marked with an asterisk.

Similarly, current applications of ML systems were retrieved and arranged under the dimension of *contextual capabilities*. Once again, the intent was to portray a good variety of possibilities while maintaining general comprehensibility. An example is given by the different ways in which generative systems can be used with visual contents, like image restoration, image transformation, and scene reconstruction.

Going upward in the taxonomy, the last two levels represent the researcher's interpretation of the translation work. With increasing abstraction, they identify how ML systems can connect to people's lives. As Manzini (2006) noted, we live in a fluid world, made of actions and interactions more than objects. And these actions are what we should focus on when designing. Hence, whatever artifact we can imagine, we will start envisioning it from the experience it should foster. For this reason, the Designerly Knowledge level has been depicted in the form of *design actions* that build on the underlying capabilities and applications, at the same time abstracting and freeing the materialization of ML systems from technical presumptions.

Finally, the Conceptual Knowledge level summarizes recurrent general capabilities attributed to ML systems. Here, synonyms like enhancement and augmentation have been carefully analyzed and selected to include the overarching intents for which ML systems could be employed.

4.1.3 Possible applications and impacts

The taxonomy can be considered a first step in the translation of ML knowledge to the design discipline. In particular, it attempts to overcome the issue of theoretical introductions making design students and professionals understand what ML is and how it works without sensitizing them to grasp the novel opportunities it offers – as described by Yang (2018). A possible remedy is provided by the practical application that it is meant to enable. Specifically, the upper part of the taxonomy allows the connection between the qualities of the technology and real-world applications. Indeed, once ML capabilities are clear, one can freely explore potential applications combining design actions and intents. Otherwise, what ML can currently do might not need to limit the creative process, as it may result in a natural solution to an independently framed intention that eventually requires ML to be realized.

Hence, as a theoretical construction, the taxonomy perfectly aligns with the overarching research objectives. In fact, if properly materialized, it could support the **(i) envisioning of consistent and meaningful solutions exploiting ML capabilities**, the **(ii) handling of ML as an asset to address challenges in a human-oriented perspective**, and the **(iii) setting of the basis for interdisciplinary communication between design and ML.**

Intuitively, the taxonomy could also be employed to categorize and make sense of current ML applications, a more conventional use for this kind of theoretical elaboration. Despite being at the basis of the construction of the taxonomy itself, it is beyond the scope of the research. However, it might positively impact the common comprehension of ML-infused solutions by providing a human-centered perspective.

4.2 Assessing forms and contents for the translation: ML Pills for Designers workshop

In line with the action research approach, a practical experiment was organized to assess the theoretical assumptions about the preferable ways for designers to handle ML knowledge. Secondary research and the analysis of ML outreach strategies revealed that highly technical information is not necessary for designers to grasp what ML can do. However, there are different possibilities to translate ML knowledge, both in terms of content and forms. The following experimental didactic activity was intended to test them with potential recipients of the research. The construction of the activity itself has also been a first testing ground for the taxonomy as a tool to support mutual understanding when discussing the possibilities of ML and to frame ML knowledge for learning. Both a synthetic overview (pages 100-101) and a detailed description (in the following sections) are provided to explain the workshop activity and its significance for the research.

4.2.1 Methods

4.2.1.1 General framing

A one-day workshop – *ML Pills for Designers* – was organized in collaboration with Oriana Arnone, a MSc student in Digital and Interaction Design (DID) at Politecnico di Milano. She was developing her master's degree thesis aimed at understanding how to enable designers to envision products, services, and experiences integrating ML. Towards this objective that the two researchers shared, the workshop specifically intended to assess (i) what **forms and languages for framing ML knowledge** are

ML Pills for Designers

Contextual information

WHAT	One-day hands-on workshop (8 hours).
WHEN	6 March 2021.
WHERE	Online (pandemic restrictions).
WHO	The author in collaboration with Oriana Arnone, a MSc student in Digital and Interaction Design (DID) at Politecnico di Milano.
STUDENTS INVOLVED	17 DID students (Politecnico di Milano) in their second year or recently graduated. Voluntary participation with no academic credits.

Research Rationale

RQ UNDER INVESTIGATION	***RQ1: How to frame ML knowledge for transfer?***
RESEARCH OBJECTIVE(S)	(i) Identify **preferable ways for designers to handle ML knowledge** in terms of forms and languages. (ii) Assess the effectiveness of the **synthetic explanation** of ML systems for transfer (input-processing-output core) (iii) Test the **appropriateness of the taxonomy** categorization of ML tasks to foster understanding and application of ML systems.
TARGET AUDIENCE RELEVANCE	MSc students in Digital and Interaction Design were targeted because of their good reflective skills, experience with a human-centered design approach, familiarity with designing technology-based solutions, but NO formal education in ML. Their profile makes them consistent recipients of the educational contents proposed and valuable resources for feedback.

Methodological framing

EXPLORATION STRATEGY	Students' self-exploration of ML capabilities proposed in different forms (*ML Pills*) and through the synthetic lens of the *Decoder,* and immediate application of the acquired knowledge to envision a ML-infused solution, guided by the *Encoder*. The students worked in groups of 2-3 people. Each group only received one form of ML Pills to explore a set of 3 ML capabilities, among which they had to use one (or more) to respond to a specific design brief. This process was iterated twice to make each group familiarize with two different forms of knowledge transfer (for future comparison) and to maximize the possibilities for each type of ML pill to be tested. This allowed a direct collection of feedback on the different forms and languages for the translation of ML knowledge, and an indirect evaluation of the understanding and operationalization of the conveyed information through the practical activity.
DATA COLLECTION	• Observation • Questionnaires (before, during, and after the educational activity) • Students' delivery of ML-infused concepts
RESEARCHER'S ROLE	Participant observer (facilitating the educational activities while gathering feedback and insights).

Structure of the educational activity

ILOS	***Knowledge***	• Understand **basic ML capabilities** and infer their potentialities
	Skills	• Approach ML as a **design material** • Identify **relevant problems** to be solved with the application of ML and imagine meaningful ML-infused solutions
	Values	• Understand **ML as an asset** for design • Maintain a **human-driven approach**
EXPECTED IMPACT(S)		**Build basic awareness** on ML to instill curiosity toward the topic and a possible specialization.
CONTENTS		• Demystification • Definition • ML process • Human involvement • ML capabilities / potentialities
TOOLS	***Knowledge transfer***	• *ML Pills* (cards explaining ML capabilities in the form of Definitions, Metaphors, Case Studies, and Practical Examples) • *Decoder*
	Design activities	• Predefined Miro board (for procedural information) • *Encoder* (for supportive information)
OUTPUT		Two concepts of ML-infused solutions to increasingly complex briefs.

Findings

KEY INSIGHTS	(i) The presented contents seemed appropriate. All forms might be suitable and (unsurprisingly) best if combined. • Primary importance of examples (both in the form of *Practical Examples* and *Case Studies*). • *Metaphors* worked well as abstractions to facilitate designers' comprehension and exploration of ML. • *Definitions* were the most challenging to grasp because they lacked visual or referential components, and their language was not designer-friendly enough. (ii) The synthesis of ML systems to their core structure was naturally apprehended and proved useful for the participants to present their ideas concisely. (iii) ML capabilities depicted by ML tasks enabled students to envision new solutions by intuitively applying them in their design process.
ISSUES FOR FUTURE INVESTIGATION	Despite the human-centered approach, the solutions were mostly individualistic and could easily fail in social contexts. The introduction of **ethical reflections** could prove beneficial.

more suitable to depict it as a design material for students, (ii) the **effectiveness of the agent synthesis of ML systems** (as portrayed in the *Decoder* and *Encoder* tools, section 3.3.1) to make sense of and communicate a ML-infused solution, and (iii) the **appropriateness of the taxonomy categorization** of ML tasks to show the possibilities and stimulate the envisioning of ML-based solutions.
To do so, insights were collected through the participant observation of the researchers, questionnaires, and the delivery of ML-infused concepts for subsequent qualitative analysis.
The workshop – based on voluntary participation and giving no academic credits – was held online (the only possible choice due to the pandemic context) on 6 March 2021. It engaged 17 DID master's students from Politecnico di Milano who were attending their second year or had recently graduated. The target audience was selected to have recipients with good reflective skills, experience with a human-centered design approach, and no formal education in ML. The participants were recruited via email, sharing information about the program and objectives of the workshop and providing a link to a registration form. This was also intended to get a profile of the class, asking for name, email, bachelor background, and their subjective relationship with ML knowledge through the questions:

- How much do you think you know what Machine Learning is?
- Do you feel confident you can properly include ML systems in your projects?
- To which extent do you believe ML-related knowledge can benefit your professional future?

A hands-on nature characterized the workshop as it was aimed to test the previously presented theoretical assumptions in practice. It was articulated in different phases. As an ice-breaking activity, the participants were asked a few questions through a Google Form to understand their personal perception and understanding of ML and its possible relationship with design practice before being exposed to the workshop contents. This questionnaire was meant to measure the influence of the workshop activities on the subjective beliefs and comprehension of the participants. However, due to the scarce response rate of the second iteration, it was not possible to infer such data.
Subsequently, a concise presentation introduced ML, providing information about how it can intersect with the design process and their mutual value, the focus and definition of ML, the demystification of common beliefs (including the description of the ML process and human involvement), and the workshop scenario. Indeed, the core educational experience to transfer knowledge about ML capabilities and potentialities consisted in two practical design activities that the participants had to carry out with the support of tools specifically designed to assess (i) possible forms and languages for the translation, (ii) the *Encoder* and *Decoder* synthesis of ML systems, and (iii) the ML tasks as outlined in the *ML Designerly Taxonomy*. They were intended to implement a freshly learned ML capability into a design concept to respond to a given brief. While apprehending and putting into practice ML capabilities, the participants were also required to assess the tools that supported these activities. Then, they had to present their ideas for a peer evaluation voting session.

To conclude and wrap up the workshop, the researchers explained the ML tasks according to the different perspectives proposed in the tools (presented in detail in the next section) and launched a final questionnaire to collect feedback, comments, and suggestions about the educational activities and materials.
Besides delivering the introductory and summarizing presentations, the researchers acted as workshop facilitators, being available for all the groups for assistance, further explanations, and reviews. On one side, this was necessary to support the participants in their activities. On the other, it allowed them to directly gather feedback and insights.

4.2.1.2 Multifaceted tools to depict ML tasks

To unlock the imagination of design students, the researchers agreed to provide them with basic explanations of the ML tasks depicted in the Technical Knowledge level of the *ML Designerly Taxonomy*, which was under development at the time. In fact, while reasoning about the most appropriate content to make designers grasp what ML systems can do and derive ways to exploit them, this simplification seemed the most appropriate.
Not all the ML tasks were included in the translation. Policy evaluation, in fact, is best suited for ML experts who want to improve the performances of their algorithms, and it was difficult to find applications for less specific contexts. Instead, classification, regression, sequence prediction, generation, clustering, and action selection, were deemed suitable to trigger the participants' creativity.
To investigate the preferable form and language to communicate them to design students, four possible alternatives were identified from the previous steps of

Fig. 4.3 | Example of a Case Study card. Graphics by Oriana Arnone.

the research (the case studies analysis of outreach strategies in particular) and materialized in the form of digital cards, the ML Pills, for a comparison.
They include **(1) definitions**, as the most common way to explain complex concepts; **(2) metaphors**, as universal processes of comprehension (Umbrello, 2020) and timeless tools that designers use to address and have a feel for new problems (Schön, 1983); and examples of ML applications. As the analysis of outreach strategies demonstrated, they are widely employed to disseminate knowledge. In particular, these were differentiated into **(3) case studies** and **(4) practical examples**, both presented with three exemplars for a more nuanced comprehension.
Case Studies pills (Fig. 4.3) provided a brief analysis of an existing ML application, following the synthetic structure identified to describe ML systems and specifically outlining a short description of the artifact (*what*), motivations and purposes at the base of its development (*why*), its basic functioning (*how*), the necessary *inputs* and resulting *outputs*. Additionally, the title of the ML-infused solution and a link to deepen the comprehension of the example were included.
Practical Examples (Fig. 4.4), instead, just introduced *what* the artifact was about and *how* it worked in a few words. The central educational part consisted in trying the ML task in a first-person, engaging, interactive experience with the application at hand through the link embedded into a "try me" box on the card.
Definitions pills (Fig. 4.5) were, at the same time, the most traditional and challenging ones. They tried to explain a ML task in a comprehensible and familiar way for designers while maintaining the precision and rigor of formal definitions like those provided by Russell and Norvig (2020). Further simplification was offered through the synthesis of ML tasks into easy-to-remember questions or keywords in the style of Holbrook and Lovejoy (2017). Additionally, concrete capabilities, which later developed into the taxonomy's *applied capabilities* and *design actions* dimensions, were suggested.

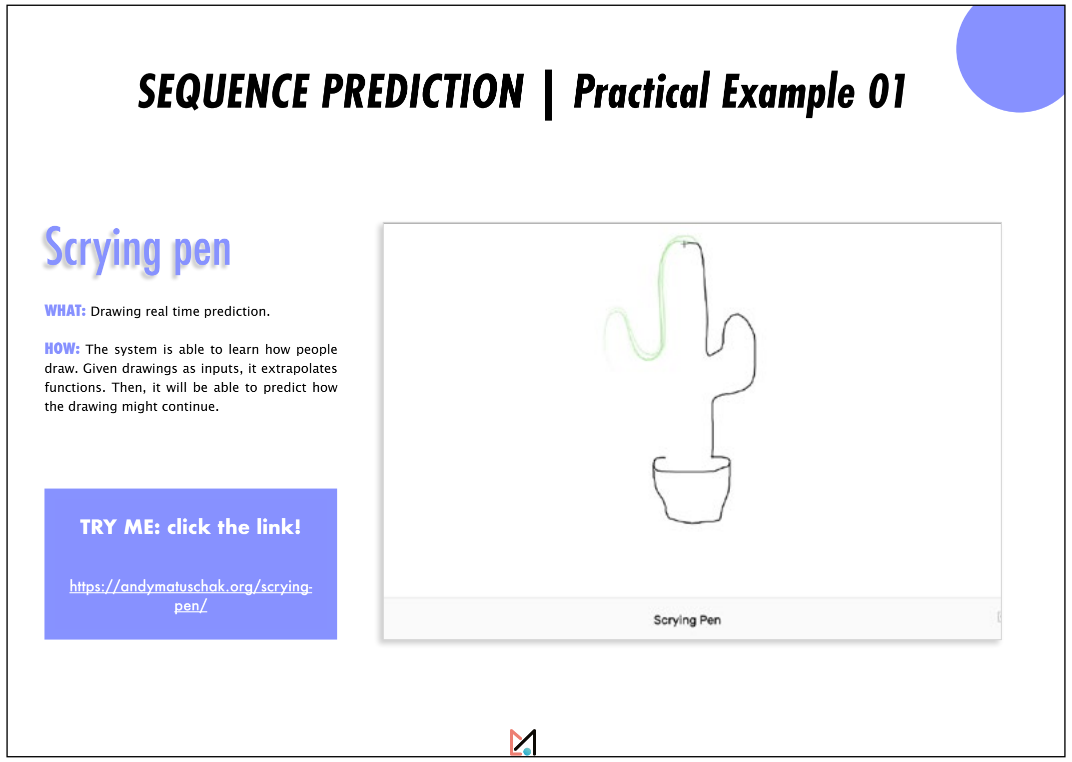


Fig. 4.4 | Example of a Practical Example card. Graphics by Oriana Arnone.

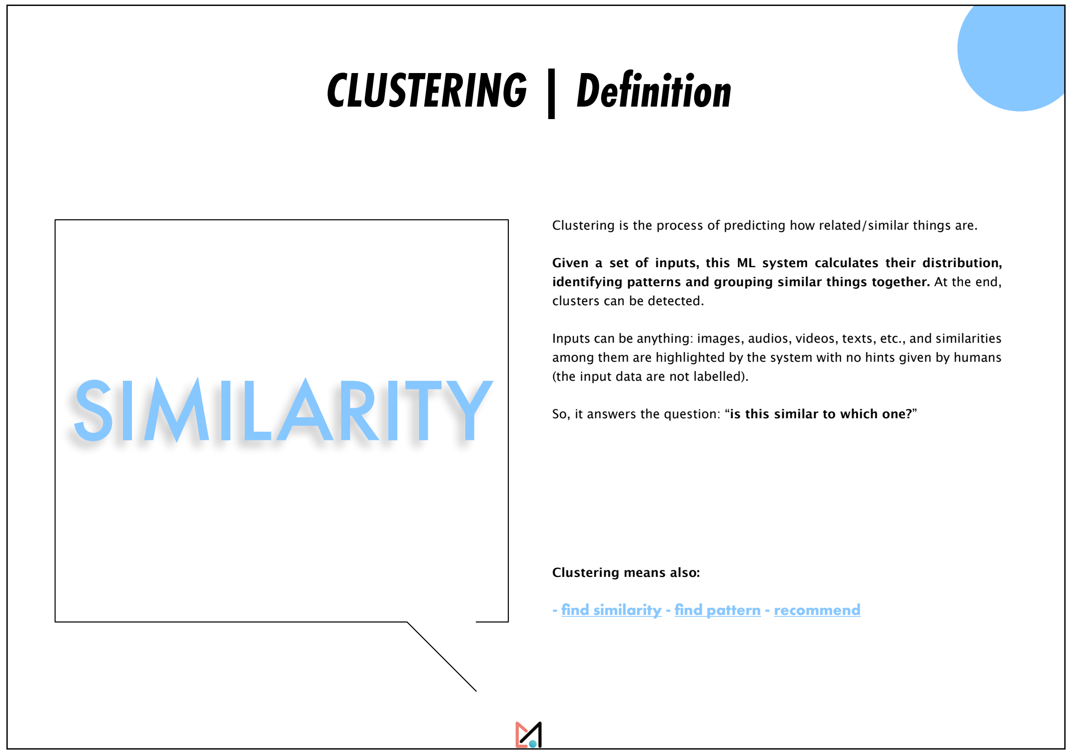


Fig. 4.5 | Example of a Definition card. Graphics by Oriana Arnone.

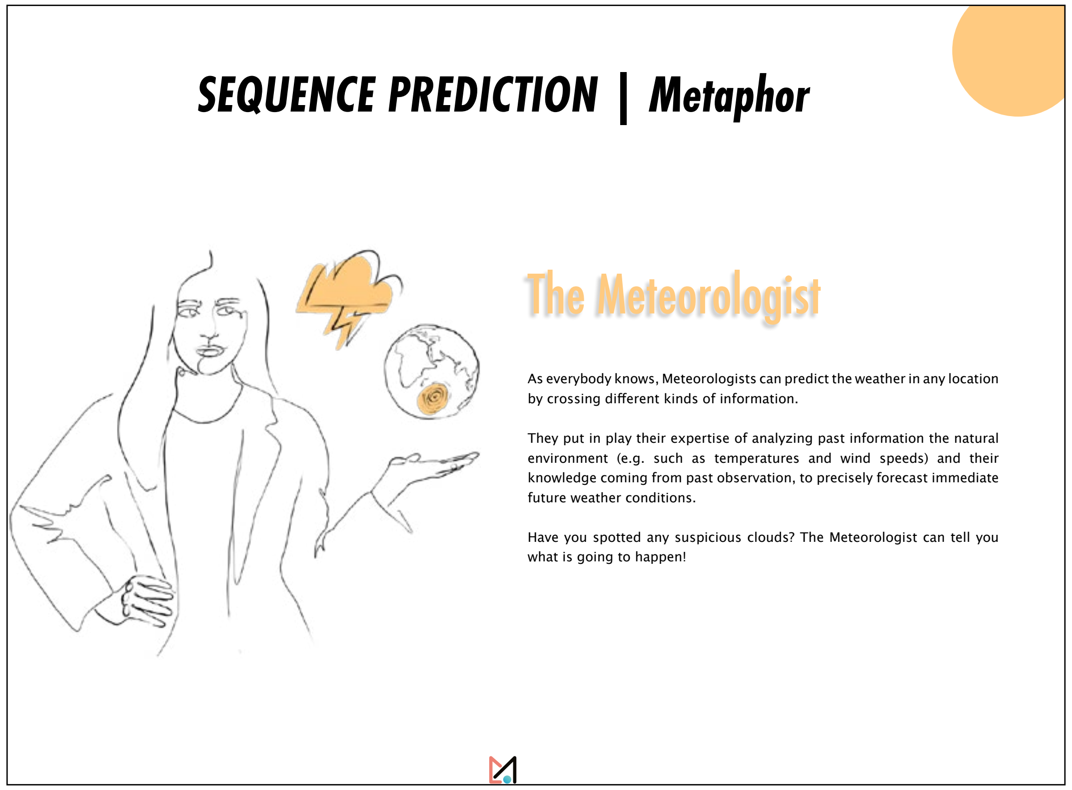


Fig. 4.6 | Example of a Metaphor card. Graphics by Oriana Arnone, illustrations by Sara Sciannamè.

Finally, Metaphors pills (Fig. 4.6) tried to translate ML task definitions in a more intuitive, visual, and evocative way, similar to what Ammagamma (2021) and Dove and Fayard (2020) did in their experimentations. In this case, the metaphorical counterpart was personified in specialized professionals whose job description carefully echoed the related ML task definition without explicit references. It was up to the reader to guess how the characterizing traits of the metaphor applied to ML systems.

4.2.1.3 Workshop scenario and design briefs

The workshop background narrative was quite simple, but it helped to give context and motivation for the design activities. In this scenario, the mayors of eight (fictional) cities wanted to obtain the bonuses that a supranational body was awarding to any city introducing technologically advanced interventions (integrating the provided ML Pills) positively perceived by their citizens as improving their quality of life. For these purposes, each mayor hired a group of designers to design innovative solutions that met the citizens' needs and were aimed at public well-being.
The 17 participants, divided into seven groups of two and one of three components, played the role of the designers assigned to the cities: Chelmis, Eyfield, Kirma, Las Wis, Leska, Pelya, Tinray, and Zranta. (The names were randomly generated.) They were asked to respond to two increasingly challenging issues by using the ML Pills and formulate concepts combining ML capabilities and human needs.
The first design brief (Fig. 4.7) consisted in finding a ML-enhanced solution to improve the life of commuters with and on public transport.
While the second one (Fig. 4.8) required an intervention in citizens' free time. In particular, it focused on offering improved museum experiences, in terms of services or cultural fruition, as a means to add quality to citizens' lives.
For both briefs, the design teams were expected to generate original ideas that could make their citizens happy, as these qualities would have been evaluated in a voting session after the presentation of all concepts. For this peer evaluation, all the

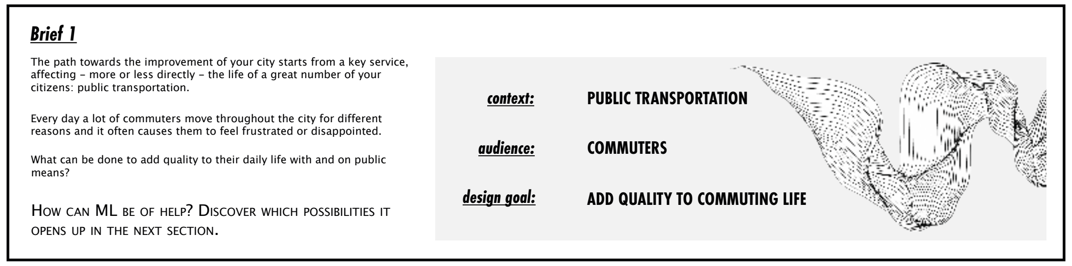


Fig. 4.7 | Presentation of Brief 1.

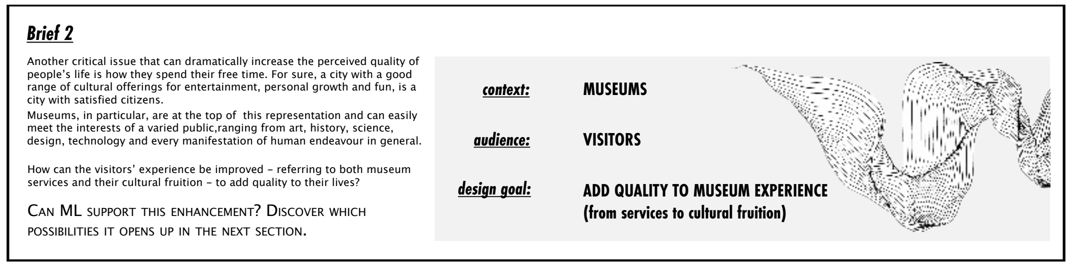


Fig. 4.8 | Presentation of Brief 2.

participants played the part of the citizens of these cities. They were given fifteen votes they could distribute as they preferred among their colleagues' ideas to assess them in terms of originality and happiness the envisioned solutions would provide. Two separate votes were taken for each design brief using a feature of the Miro platform – used to support all the workshop activities.

4.2.1.4 Testing strategy

To assess the effectiveness of the different forms and languages to translate ML capabilities for design students, each of the four formats had to be experienced alone. Combining them could have influenced the perception of the ML task explanation by providing multiple angles for understanding.
Moreover, to maximize the possibilities for each type of ML pill to be tested, each group had to use two of them.
For these reasons, the design activities were organized as follows. To respond to the first brief, the students had at their disposal only three ML tasks, i.e., classification, regression, and sequence prediction, explained by one type of ML pill. The researchers considered these tasks as more manageable to introduce the topic, and they would be used to assess a relatively simpler or more traditional design problem (improving the experience of the transportation system).
To address the second brief, all the groups discovered a different ML pill presenting the remaining ML tasks: generation, clustering, and action selection. These were more appropriate for improving the museum experience, as more creative solutions could be imagined.
For each brief, every ML pill was used by two groups and, overall, all the combinations were tested as shown in Table 4.1. For the two that had to be duplicated, combinations of written and exemplar explanations were preferred (precisely, Case Studies + Metaphors and Practical Examples + Definitions).

	TINRAY City	PELYA City	CHELMIS City	LAS WIS City	LESKA City	KIRMO City	EYFIELD City	ZRANTA City
B 1	*Case studies*	*Practical examples*	*Definitions*	*Metaphors*	*Case studies*	*Metaphors*	*Practical examples*	*Definitions*
B 2	*Metaphors*	*Definitions*	*Practical examples*	*Case studies*	*Definitions*	*Practical examples*	*Case studies*	*Metaphors*

Tab. 4.1 | Combination of groups and ML Pills in Brief 1 (B1) and Brief 2 (B2).

The design activities to address each brief were divided into two phases. The first consisted of knowledge transfer. Each group had one kind of ML pill presenting the ML tasks and 20 minutes to explore them and grasp their capabilities. This phase was supported by the *Decoder*, which could help the participants to outline the main characteristics of ML systems in case they were not clear enough.
After having made sense of the ML tasks, a first checkpoint required them to assess the tested tools according to the questions:

- *How effective are the tools in communicating the ML tasks?*
- *Are the contents clear?*

- *Are the contents sufficient to understand the ML task?*

After this, the second phase was more properly dedicated to concept development. Indeed, learning about ML tasks was fostered through their inclusion in the design process. As Stolterman (2008) suggested, they do not provide prescriptive information about how to approach the design of ML-infused solutions. Instead, they *prepare-for-action*. They give hints to reflect and find possible solutions in challenging design situations. Hence, this phase comprehended 25 minutes to explore the problem space and brainstorm ideas; 5 minutes to frame the preferred one in terms of context, audience, and design goal; and another 20 minutes to describe the concept by specifying its title, abstract, and basic structure by using the *Encoder*.
Again, an evaluation of the ML Pills' effectiveness to support the design activity followed. It included the questions:

- *Did the tool give you proper knowledge to apply ML tasks?*
- *Is the tool enabling you to exploit ML tasks for your design goal?*
- *Do you think the tool encouraged you to think out of the box?*

Further information was gathered through the participant observation of the researchers and the discussions triggered throughout the workshop, the assessment of the delivered concepts, and a final questionnaire primarily focused on the provided tools.

4.2.1.5 Intended learning outcomes and impact

To conclude, thanks to the *ML Pills for Designers* workshop, design students were expected to understand basic ML capabilities and infer their potentialities; approach ML as a design material; identify relevant problems to be solved with the application of ML and imagine meaningful ML-infused solutions; understand ML as an asset for design; and maintain a human-driven approach.
In terms of impact, the didactic experimentation was expected to **make the participants build basic awareness of ML and instill curiosity** towards the topic and a possible specialization.

4.2.2 Results

4.2.2.1 Pre-workshop

To frame the context of the workshop, it is useful to have an overview of the participants' knowledge baseline and predisposition about ML. As anticipated, all 17 students were at the end or had finished their path in the Digital and Interaction Design MSc program at Politecnico di Milano. This entails that they were used to prototyping and developing projects enhanced by digital technologies but had no formal education in ML. Only some had participated in a workshop on AI and data a few weeks before. However, no one felt very knowledgeable on the subject (Fig. 4.9). On a 4-point Likert scale, six of the participants assessed their level of knowledge at a 3, the majority (9) answered with a 2, while two affirmed they had no ML knowledge at all. The same applied to their confidence about being able to include ML systems in

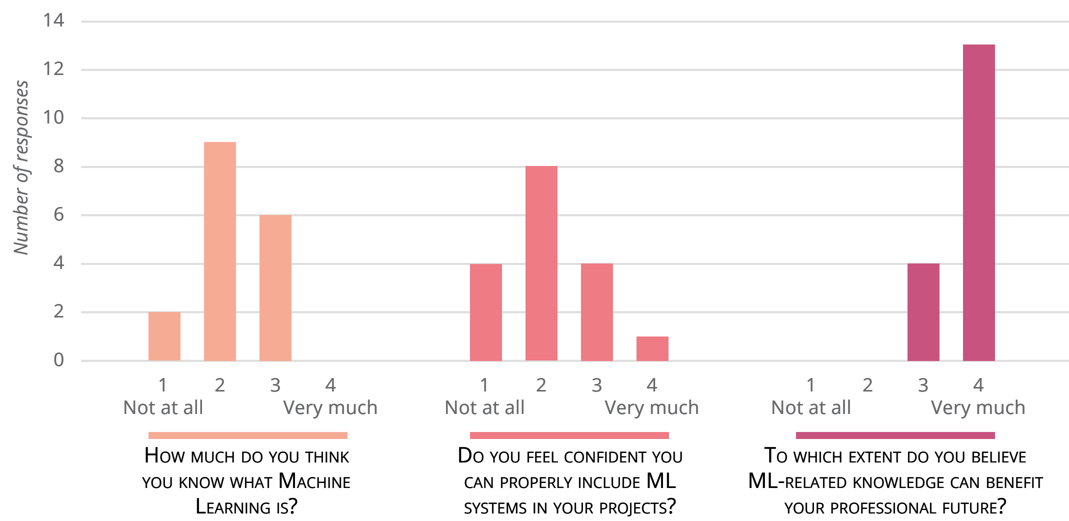


Fig. 4.9 | Preliminary responses of all the 17 participants to the registration form.

their projects (even though the question might be confusing because the moment to which it refers is not clear).
Still, they had good intuitions about how ML capabilities could or could not manifest in current products and services. In the brief questionnaire that they took before the beginning of the workshop, most of them correctly identified if the proposed items were ML or not (Fig. 4.10). The trickiest artifacts mislabeled as ML systems were *"A system predicting your BMI (Body Mass Index),"* a mathematical formula described with the intendedly misleading word *predict*, and *"A navigation system suggesting you a route to cover the less distance from your starting point to a selected location,"* although the correct counterpart was also present (*"A navigation system suggesting you the best route according to real-time traffic conditions"*).
In general, the participants expressed a positive attitude toward the subject matter. They felt optimistic about the usefulness of acquiring ML-related knowledge for their professional future (Fig. 4.9). Mostly, they were aware of the pervasiveness of this technology and saw it as a great opportunity or a tool (Fig. 4.11). In Fig. 4.12, it is also visible that the majority recognized the relevant role ML is going to play in design projects. Though, they seemed aware that the process might be long.
This positive predisposition toward the subject was probably a consequence of voluntary participation. Nonetheless, it created excellent premises for the workshop.

4.2.2.2 Analyzing the delivered concepts

Despite having only 70 minutes to get familiar with the three ML tasks presented, brainstorm, and outline an idea integrating (at least) one of them for each brief, all groups managed to deliver their concepts going through all the planned learning and design activities. Overall, the results were satisfactory, as all the developed ideas addressed problems that might benefit from the implementation of ML capabilities. Tables 4.2 and 4.3 portray the concepts using a code to identify them (those terminating with 01 refer to ideas responding to the first brief, 02 to the second), a brief description with the related title in bold, and a synthesis of their structure based on the inputs, tasks, and outputs as depicted by the groups in the *Encoder* and presentations.

Code	Concept	Input-Task-Output	ML Pill	ML Task	(i) ML solution		(ii) ML task consistency		(iii) Added value / Happiness			(iv) Originality		
					R1	R2	R1	R2	R1	R2	Votes	R1	R2	Votes
CH01	**Building an emotional connection with different places.** Help commuters discover the emotional qualities of different areas based on locals' perceptions.	(i) Locals' likes, dislikes, feelings, and preferences regarding a place - (t) Sequence prediction - (o) Destinations and activities to discover	*Definitions*	Sequence prediction	3	3	2	2	4	3	**33**	4	4	**36**
EY01	**Green and safe.** Suggest the best option(s) of transportation according to safety (crowdedness) and ecofootprint perspectives.	(i) CO2 consumption, time of the transport (day, month, part of the year), amount of people taking the same transport, length of the path, duration of the journey - (t) Regression - (o) Suggestion of best transportation option	*Practical examples*	Regression	4	3	3	3	3	3	**23**	3	3	**27**
KI01	**Release the stress of workers on the train.** Decrease workers' stress level by adjusting the air composition (scents) and temperature based on personal data.	(i) Personal daily health data - (t) Sequence prediction - (o) Stress level value decreased	*Metaphors*	Sequence prediction	3	3	1	4	3	3	**29**	3	4	**27**
LW01	**Unexpected events warning system.** Learn students' daily routines and warn them in case of disruptive events on their trip.	(i) Route information and users' daily routine - (t) Classification - (o) Warning notification	*Metaphors*	Classification	4	3	4	3	3	2	**16**	3	2	**18**
LE01	**My personal transport detector.** Improve transportation experiences by letting commuters know how crowded transports are and giving them options.	(i) Number of people (based on tickets, Google Maps info, maybe cameras) - (t) Regression + Sequence prediction - (o) Report on how crowded transports are	*Case studies*	Regression + Sequence prediction	4	4	3	3	3	3	**13**	3	2	**15**
PE01	**Commate.** Giving companionship to lone commuters by suggesting content (videos, podcasts, jokes) based on people's mood.	(i) Emotion, facial, activity recognition through camera + users' preferences - (t) Regression - (o) Video and/ or audio entertainment (background music, podcast, jokes, stories, etc.)	*Practical examples*	Regression	3	4	1	1	4	3	**30**	3	3	**33**
TI01	**Route optimization of public transportation in the city.** Reduce people's frustration when taking public transport by suggesting the best (quickest, less crowded) route to get to their destination.	(i) Train position and timings - (t) Sequence prediction - (o) Accurate delay (i) Live feed of people and cars at key points - (t) Classification - (o) Crowdedness of the station	*Case studies*	Sequence prediction	3	4	4	3	3	3	**12**	1	1	**5**
ZR01	**On-time.** Suggest real-time alternatives if there is an anomaly in the regular route. Relieve commuters from planning options.	(i) Commuter's commuting preferences and routine - (t) Sequence prediction - (o) Suggestions of best transportation alternatives to get to a destination	*Definitions*	Sequence prediction	3	4	3	3	3	3	**17**	2	2	**15**

Tab. 4.2 | Synthesis of the concepts generated for Brief 1.

Code	Concept	Input-Task-Output	ML Pill	ML Task	(i) ML solution		(ii) ML task consistency		(iii) Added value / Happiness			(iv) Originality		
					R1	R2	R1	R2	R1	R2	Votes	R1	R2	Votes
CH02	**Muzo - Your personal museum journey.** Suggest a personalized journey inside the museum based people's time, level of attention, curiosity, energy, and interests.	(i) Visitor's time, energy (age, tiredness), interests, curiosity level, attention (during the visit) - (t) Clustering - (o) Museum journey	*Practical examples*	Clustering	4	4	3	3	4	3	**23**	4	3	**15**
EY02	**Frames of music.** Generate a musical composition based on people's voices that can be more inclusive by activating more senses.	(i) Tone, pitch, volume - (t) Generation - (o) Musical composition	*Case studies*	Generation	4	4	4	4	4	3	**42**	3	4	**40**
KI02	**Children's paradise to play with art.** Interactive kindergarten spaces that allow children to paint virtually (digital projections) with their body movements according to an artist's style.	(i) Abundant artworks in the same style, skeleton tracking data - (t) Generation - (o) Interactive generated artworks and ambient effects	*Practical examples*	Generation	4	4	4	4	4	3	**26**	3	3	**31**
LW02	**Data visualizing system.** Expand the experience of virtual tours by letting visitors take photos of artworks they liked and creating publicity.	(i) Photos - (t) Generation + Clustering - (o) Trends, Publicity, Photo	*Case studies*	Clustering	4	4	3	3	2	2	**26**	3	2	**28**
LE02	**Community builder for art lovers.** Creating a community of art lovers based on their preferences and behavior in museums.	(i) Time spent in front of artworks, audio guides data, manual, museums visited, profiling (age, education, ...) - (t) Clustering - (o) Groups of people with similar characteristics and preferences about the artworks	*Definitions*	Clustering	4	4	4	4	3	3	**31**	2	3	**29**
PE02	**NetGallery - live museum discovering network.** Museum network that offers a recommendation-based experience, proposing contents from other museums coherently with visitors' preferences, literature, artists, and artworks info.	(i) Literature and people's opinion on certain artworks and artists + your preferences about some specific artworks - (t) Clustering - (o) Real-time suggestions of similar artworks/ literature from other museums	*Definitions*	Clustering	4	4	4	4	3	3	**41**	2	2	**34**
TI02	**Exhibition of ourselves.** Display that involves visitors creating new art representing their preferences and mood.	(i) Cookies, self-identified mood, art pieces - (t) Clustering + Generation - (o) A new piece of art based on the collective preferences of the group	*Metaphors*	Clustering + Generation	4	4	4	4	3	3	**14**	3	3	**14**
ZR02	**P.I.Y (Paint It Yourself).** Give the possibility to create artworks based on people's personal tastes and add them to the museum collection.	(i) Favorite artist, a person's aesthetics, favorite colors - (t) Generation - (o) A new work of art seen from the personal filter of the visitor's perspective	*Metaphors*	Generation	4	4	4	4	3	4	**27**	3	3	**19**

Tab. 4.3 | Synthesis of the concepts generated for Brief 2.

Fig. 4.10 | Results of the quiz about recognizing ML capabilities in current products and services. Response rate 15 out of 17 students.

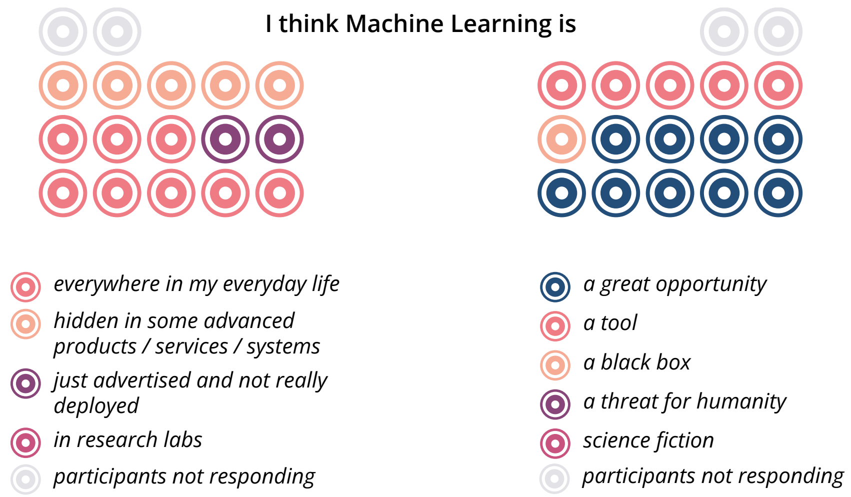


Fig. 4.11 | Participants' preconceptions about ML. Response rate 15/17.

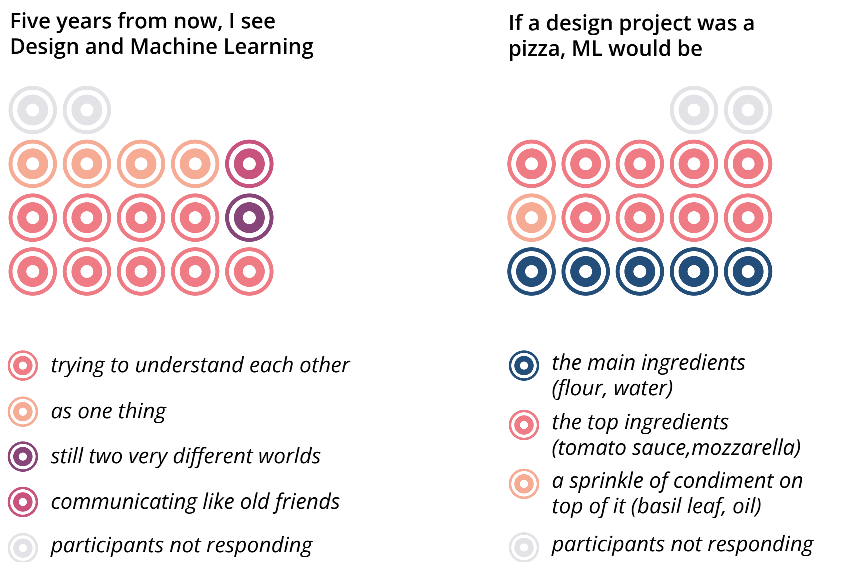


Fig. 4.12 | Preliminary perception of the relationship between ML and the design field. Response rate 15/17.

The first brief called for utilitarian solutions and, for the greatest part, the groups aimed to optimize the organization of commuting. Instead, the process iteration to respond to the second brief showed more creative, experiential, and art-based interventions. This was probably due to factors like the requirements and contexts of the briefs themselves, the different ML tasks at disposal, and the influence of the voting session and discussion after the first round of presentations, which appreciated more the proposals that appealed to people's emotional side.
Further insights have been inferred by evaluating the delivered concepts. The researchers operated independently, based on four parameters. (i) Whether the conceived solution could actually benefit from the integration of a ML system; (ii) the consistent use of the ML task; (iii) the added value for people; and (iv) the overall originality. These were assessed on a 4-point Likert scale (1: Not at all; 2: Slightly; 3: Moderately; 4: Very much). The results of the judging activity and those related to the theoretical premises, tools, and students' approach to the design of ML-based solutions are presented in the following. Tables 4.2 and 4.3 will serve as references for detailed information.

4.2.2.3 Decoder and Encoder to synthesize ML systems

From an external perspective, based on the observation of the Miro boards, reviews, and groups' presentations of the ideas, the participants had no problem handling the synthetic elements used to outline ML systems (inputs, tasks, outputs, and the concept of learning from experience). They used them to frame their thoughts in a natural and coherent way.
The *Decoder* and the *Encoder* supported the students' comprehension and envisioning of ML systems in multiple ways. They were employed to encourage reflection while understanding the provided ML knowledge and the suitability of the idea to be addressed by ML systems. One group even used the *Decoder* analytically by answering in writing the questions about the ML tasks to be apprehended (Fig. 4.13).
Additionally, the *Encoder* proved effective in summarizing and better specifying the developed concepts because it allowed focusing on the essential elements to communicate what a ML system does. Sometimes, it even proved more

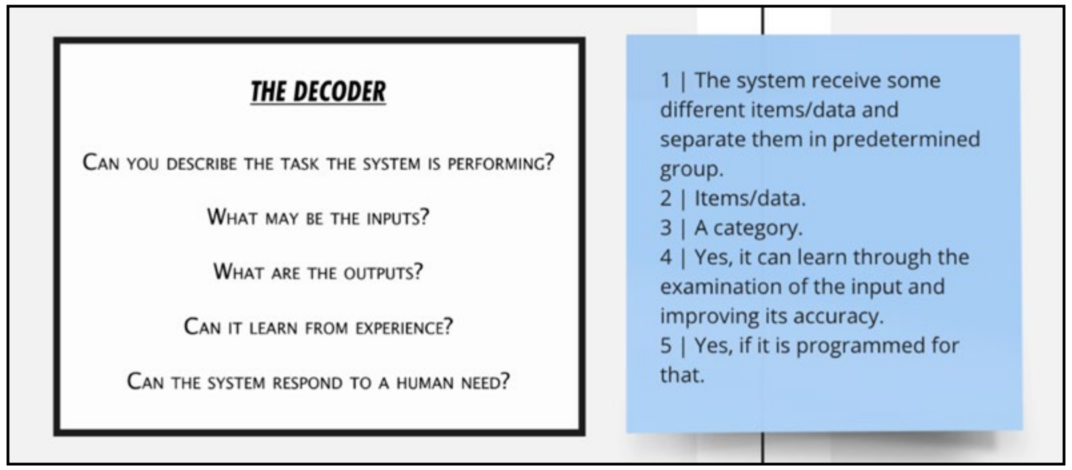


Fig. 4.13 | Example of analytic use of the Decoder to check one's understanding.

straightforward than the presentations because the participants had the chance to better articulate the ingredients of their ideas while not always they followed the same structure to talk about them in the given two-minute slots. Of course, these tools could not check the correctness of the inserted contents (as Tables 4.2 and 4.3 prove), but they fostered a helpful mental process, which was reflected in almost all the presentations.

Based on the specific evaluation that 14 (out of 17) students gave to these tools at the end of the workshop (Fig. 4.14), it can be seen that they have been quite positively perceived. Nine participants found the *Decoder* moderately helpful in understanding ML tasks, and two very much. At the same time, according to four respondents, the *Encoder* was very useful in framing the concept and moderately so for six others.

From the collected comments, however, it emerges that some improvements should

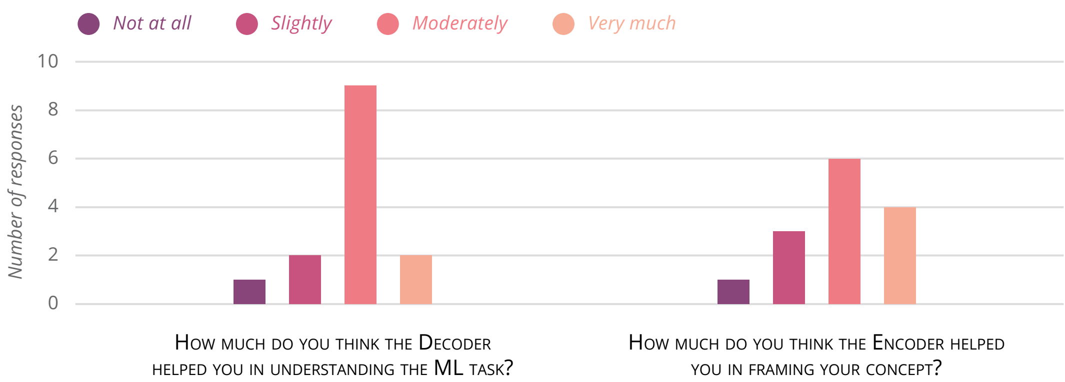


Fig. 4.14 | Assessment of Encoder and Decoder.

be applied to make both the *Decoder* and *Encoder* more explicit and self-standing tools. Indeed, they imply concepts like "learning from experience" that might be unclear if one missed or did not focus on the introductive presentation. Possibly, a more proper introduction to the tools could also improve their perceived usefulness.

4.2.2.4 ML Tasks to enable concept generation

To understand whether ML tasks could effectively enable design students with little or no prior knowledge of the subject matter to envision ML-infused solutions consistent with the technology capabilities, it was useful to establish whether the delivered concepts could really benefit from the integration of ML. According to the independent evaluation that the two researchers conducted (Tables 4.2 and 4.3), all the envisioned systems reported values between 3 and 4 (4 was the most positive grade on a 4-point Likert scale), indicating that all the groups succeeded in developing a consistent ML solution with the sole support of ML tasks. To this end, the *ML Suitability Matrix* (described in section 3.3.1) could have been useful. However, it would have required more in-depth explanations and considerations that were not possible in the given timeframe. Interestingly, though, the basic understanding of what ML systems can do, provided by ML tasks, was enough for design students used to work with digital technologies to design accordingly.

Contrary to expectations, the introduction of clustering, generation, and action selection to respond to Brief 2 led to the development of solutions uniquely implementable thanks to ML capabilities (all were marked with the highest score in terms of ML solution by both judges). Although the researchers believed that these tasks could be more complex for non-expert design students, they proved so manageable that the consistency with the ML tasks themselves was impeccable in almost all cases. The two exceptions, CH02 and LW02, reported a lower evaluation because they were not detailed enough to assess with certainty that the selection of clustering and generation aligned with what the groups thought. In the first case, the idea was so articulated that clustering could certainly be part of it to find correlations between people's time, energy, interests, museum places, and visit modalities. However, other systems should be added to suggest a personalized visit journey. In the second, a generation system could create advertising based on the detected trends. It is less clear, instead, how clustering should be implemented. It was assumed that the photos taken by the virtual tour visitors had to be mapped to find patterns of virtual visits to propose to new users. If so, clustering could be consistent, but without full certainty about what was meant, it could not get a full score.

A slightly opposite situation manifested for developing interventions integrating classification, regression, or sequence prediction. In response to Brief 1, the need for ML solutions was less neat, and often a combination with traditional programs was needed, but the participants did not highlight that. It is the case for CH01, which might need some automation to classify places according to locals' evaluations, comments on social media, etc., but could use traditional programming to suggest a place based on this classification. Similarly, the solution provided by KI01 to adjust the environmental conditions on a train and release commuters' stress included non-ML features. For instance, people could be directly asked how they felt to let the system infer the optimal condition based on scientific studies and prescriptive rules.

Further issues emerged concerning the consistency of the selected ML tasks and the envisioned ideas. Sequence prediction, in particular, was subject to misunderstandings. Again, CH01 proposed a sequence prediction system to suggest commuters new places to discover according to their emotional qualities. However, as they focused on the locals' characterization of these qualities and no possible

historical or sequential information could be retrieved, sequence prediction was of no use, and classification could be a better choice to create this new categorization. Even more patent is the misunderstanding of PE01, which uses regression to suggest content for the entertainment of lone commuters. It is against the fundamental premise that regression can output only numbers. On the contrary, it could use sequence prediction if it had access to historic users' data or a combination of classification and some predetermined rules to provide a recommendation.
In general, the extreme synthesis of ML capabilities in single ML tasks could not help to address multifaceted problems, for which more articulated solutions would require a combination of multiple ML and non-ML systems. It is especially valid for tasks like classification, regression, and sequence prediction that have very narrow capabilities and necessarily need to be complemented. Thus, as the participants did not have the means to address such complex issues, unclear situations emerged. KI01 concept is emblematic in this sense, as it is visible in the researchers' evaluation of ML task consistency. They selected the opposite values of the Likert scale, and both motivations are reasonable, depending on the interpretation. The proposed solution is challenging from its intent: to personalize environmental conditions on public transports that are not for individual use, but the actual feasibility of the idea was not the object of evaluation. To do so, the design students proposed integrating a sequence prediction system that could use physiological data to modify room temperature and scent to reduce stress levels. As presented, no sequential data can be identified for the system in use. Thus, it would not be possible to use sequence prediction. However, if one thinks about the training of the system itself, correlations between the modification of physiological data and environmental conditions would need sequential information to be inferred.
Overall, then, ML tasks can support the envisioning of consistent ML solutions. They might not be enough to address very articulated and complex ones to address real-world applications, but this could go beyond the scope of an introductory level.

4.2.2.5 Evaluation of forms and languages

In order to understand which kind of form and language worked best to enable design students to understand and apply ML tasks in practice, multiple strategies have been put in place, as this was the core of the experimentation. The participants were asked to assess the ML Pills they used during the design activities (after both the self-learning and application parts), as well as once the workshop ended and the complete overview of the possibilities was presented. 14 out of 17 participants responded to this questionnaire, and the ML Pills they used are illustrated in Fig. 4.15. Additional information comes from the researchers' observation and evaluation of the results.
The effectiveness of the ML Pills should be measured by their ability to facilitate understanding and support the design process. To achieve the first, students' evaluation is the main reference (Fig. 4.16). What emerges is that the example-based ML Pills were the most appreciated on all fronts, with Practical Examples resulting in the most successful format. This was also highlighted in the conclusive assessment (Fig. 4.17) and the comments, as they made it *"easier to understand how the tasks can work in a real environment."* During the reviews, instead, issues about the clarity of

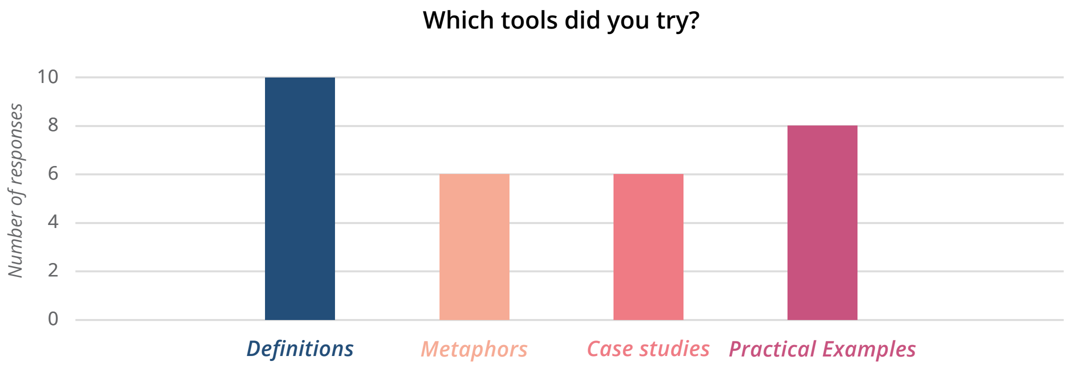


Fig. 4.15 | ML Pills used by the respondents to the questionnaires.

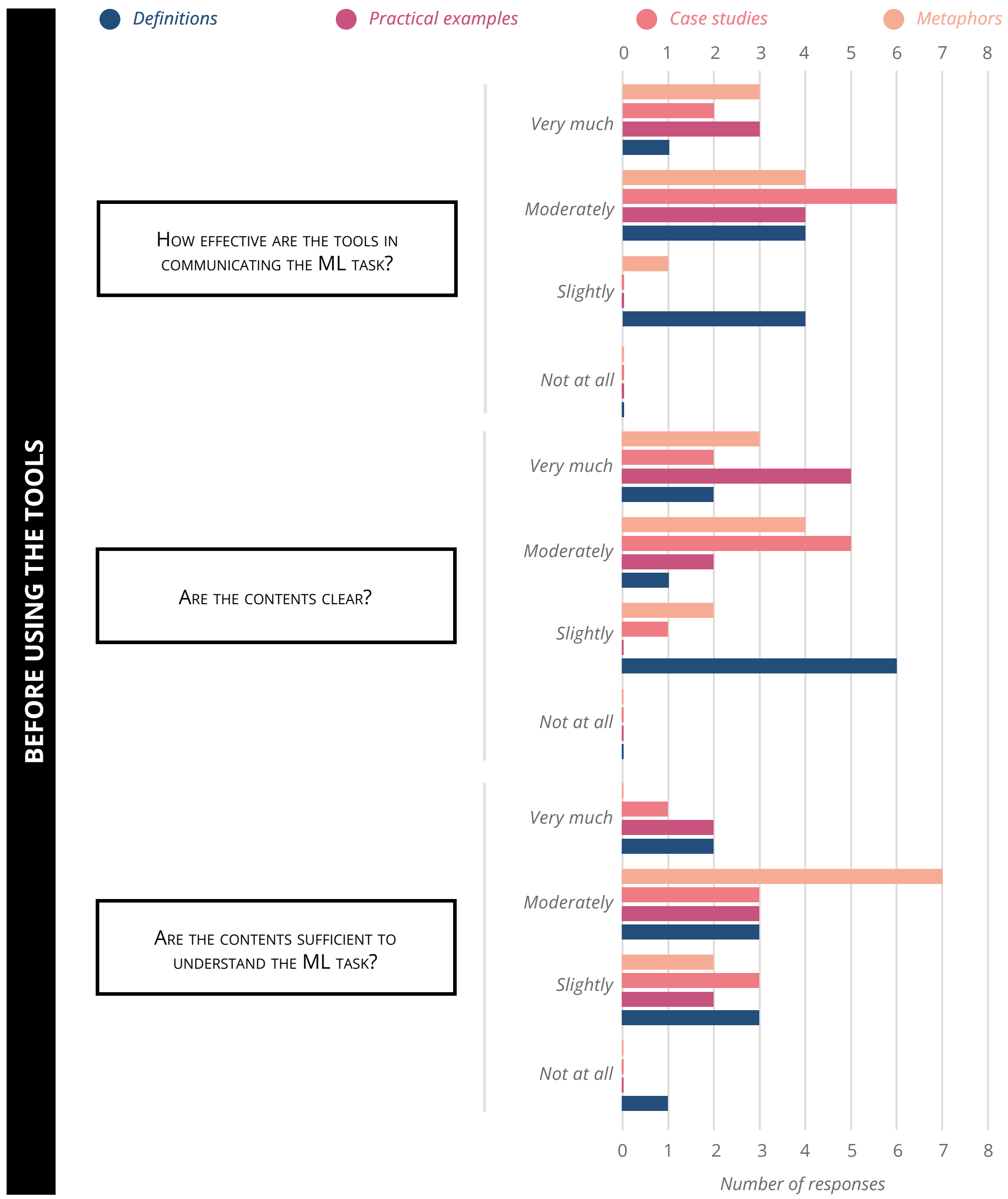


Fig. 4.16 | Students' assessment of the ML Pills before applying them to the concept generation.

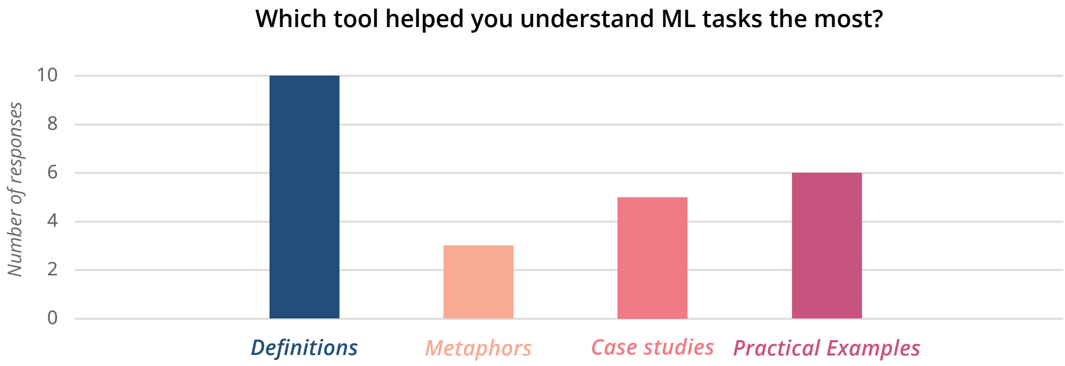


Fig. 4.17 | Identification of the most useful ML Pills according to the respondents.

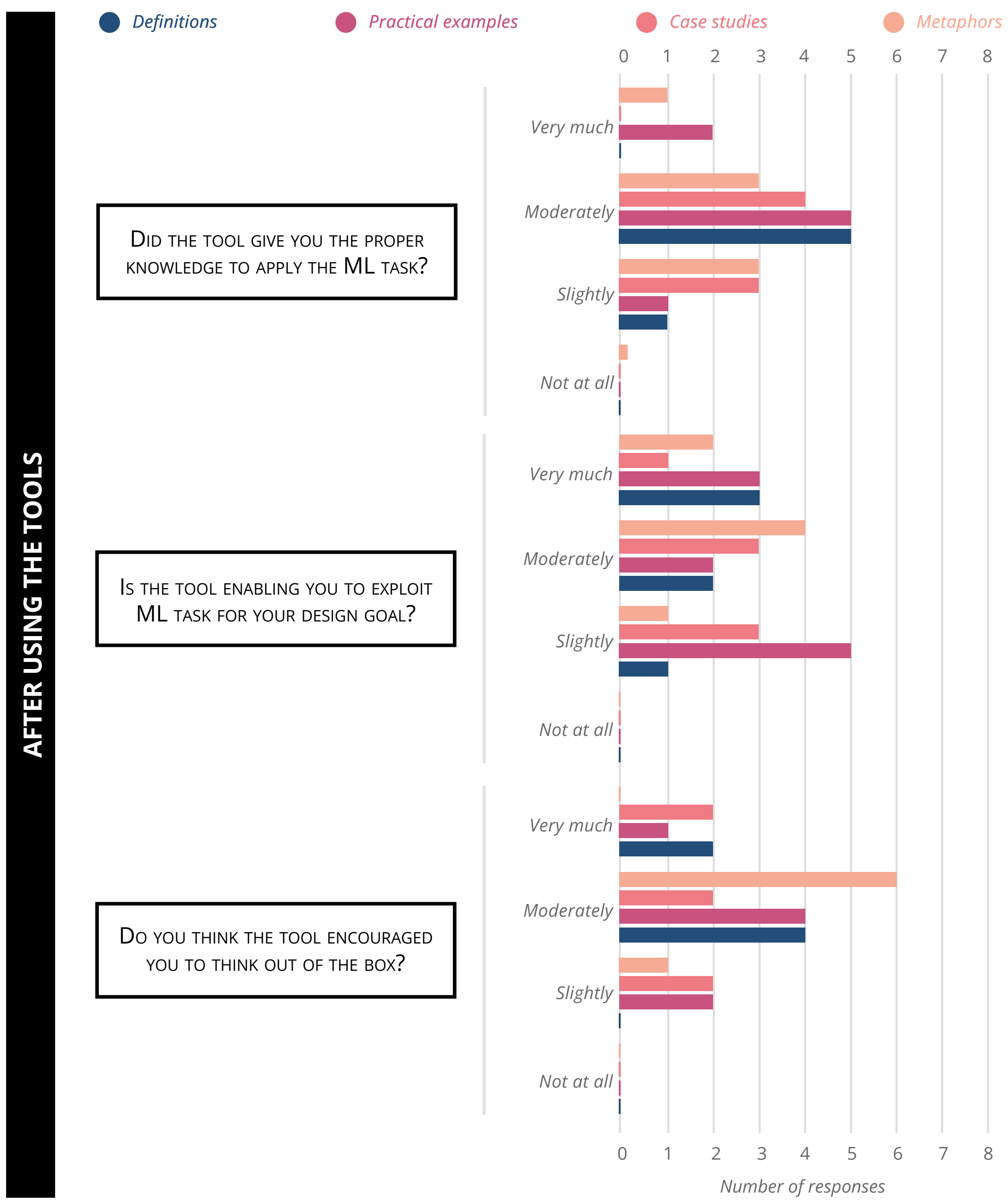


Fig. 4.18 | Students' assessment of the ML Pills after applying them to the concept generation.

the explanations contained in the Metaphors were brought out by two groups. A group dealing with Definitions stated, *"We don't know what to do! Before* [with practical examples], *it was clear."* The same comparison was made by another group that emphasized the clarity of Practical Examples in responding to Brief 2, in contrast to the first design activity, where they dealt with Definitions. In general, the lower ratings of the tools on the contents' sufficiency suggests that more information might be helpful.

To assess how ML Pills could support a consistent development of ML-infused ideas, both the participants' (Fig. 4.18) and the researchers' evaluations (Tab. 4.4) can be useful for triangulating data. Again, Practical Examples were the best in providing proper knowledge to apply the ML task and enabling its exploitation to achieve the design goal, as confirmed in the final questionnaire (Fig.4.19). Surprisingly, the second-best format for the same questions was Definitions, also considered the best tool to encourage out-of-the-box thinking, followed by Practical Examples. These results did not reflect the consistent application of ML tasks as assessed by the researchers. Indeed, the ML Pills producing the most coherent ideas with the ML task selected were the Metaphors, closely followed by Case Studies. Practical Examples reported the lowest average score. Regarding fostered creativity, instead, Practical Examples had the greatest impact on the originality of the outcomes, followed by Metaphors and Definitions, in contrast with the students' previous evaluations. In the end, though, the format that struck the most to enhance creativity was the Metaphor (Fig. 4.20).

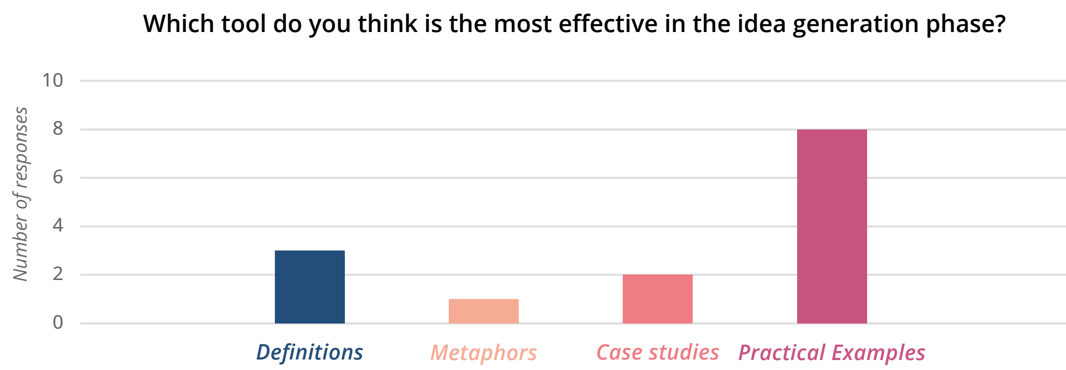


Fig. 4.19 | Students' responses about the most effective tool for idea generation.

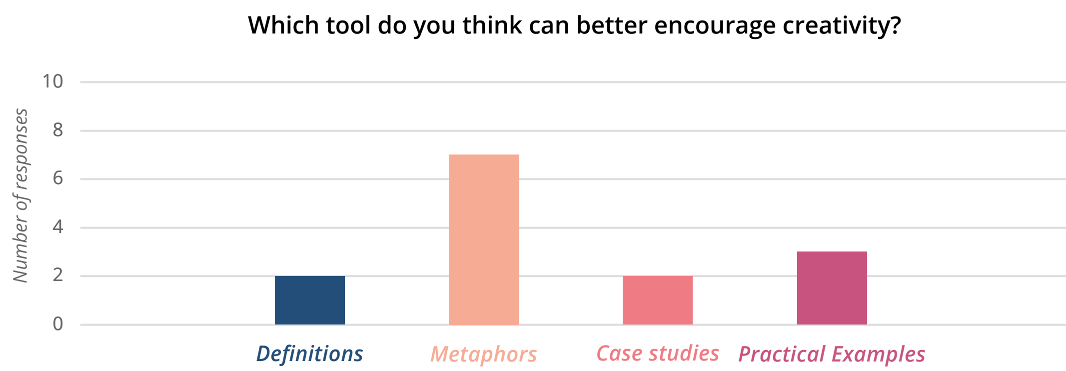


Fig. 4.20 | Students' responses about the most effective tool to encourage creativity.

Code	ML Pill	ML task	ML Task Consistency R1	ML Task Consistency R2	Average consistency	Originality R1	Originality R2	Average originality
LE01	*Case studies*	Regression + Sequence prediction	3	3	3,38	3	2	2,38
TI01	*Case studies*	Sequence prediction	4	3		1	1	
EY02	*Case studies*	Generation	4	4		3	4	
LW02	*Case studies*	Clustering	3	3		3	2	
CH01	*Definitions*	Sequence prediction	2	2	3,25	4	4	2,63
ZR01	*Definitions*	Sequence prediction	3	3		2	2	
LE02	*Definitions*	Clustering	4	4		2	3	
PE02	*Definitions*	Clustering	4	4		2	2	
KI01	*Metaphors*	Sequence prediction	1	4	3,50	3	4	3,00
LW01	*Metaphors*	Classification	4	3		3	2	
TI02	*Metaphors*	Clustering + Generation	4	4		3	3	
ZR02	*Metaphors*	Generation	4	4		3	3	
EY01	*Practical examples*	Regression	3	3	2,75	3	3	3,13
PE01	*Practical examples*	Regression	1	1		3	3	
CH02	*Practical examples*	Clustering	3	3		4	3	
KI02	*Practical examples*	Generation	4	4		3	3	

Tab. 4.4 | Researchers' evaluations of the developed concepts.

Concerning the design activity, several responses underlined the effectiveness of Practical Examples and Metaphors because they helped students ***visualize*** ML capabilities. This verb was indeed recurrent in the comments that wanted to emphasize the positive qualities of a ML Pill. This is possibly why Definitions portraying written information to be elaborated by the recipients were not as effective. From feedback, Practical Examples were particularly appreciated because of the immediacy to fuel and translate ML capabilities in the design process. Metaphors were defined as *"more poetic,"* and the identified value lay in enhancing a personal abstraction and interpretation process that could lead to disruptive ideas. The few comments on Case Studies highlighted their clarity, while the *"crucial"* role of Definitions was recognized but with the potential for being more effective if complemented with examples. Overall, the synthesis in Fig. 4.21 and the general preference expressed in the final questionnaire (Fig. 4.22) leave no doubts about the designers' tendency to make sense of the world and work by association in their reflective practice (Schön, 1983). But, as expected, a proper translation should merge different forms and languages. It immediately emerged as a necessity when the groups understood that there were multiple modalities for introducing ML capabilities, and it was confirmed by the responses to the direct conclusive questions (Fig. 4.23). The only respondent who stated that one tool was enough indicated Case Studies as the best option, but the majority deemed the combination of more ML Pills as a preferable strategy.

	Question	1st place	2nd place	3rd place	4th place
BEFORE USING THE TOOLS	*How effective are the tools in communicating the ML task?*	*Practical examples* 3,43	*Case studies* 3,25	*Metaphors* 3,11	*Definitions* 2,66
	Are the contents clear?	*Practical examples* 3,43	*Case studies* 3,25	*Metaphors* 3,11	*Definitions* 2,66
	Are the contents sufficient to understand the ML task?	*Practical examples* 3,43	*Metaphors* 3,11	*Case studies* 3,25	*Definitions* 2,66
AFTER USING THE TOOLS	*Did the tool give you the proper knowledge to apply the ML task?*	*Practical examples* 3,43	*Definitions* 2,66	*Metaphors* 3,11	*Case studies* 3,25
	Is the tool enabling you to exploit ML task for your design goal?	*Practical examples* 3,43	*Definitions* 2,66	*Metaphors* 3,11	*Case studies* 3,25
	Do you think the tool encouraged you to think out of the box?	*Definitions* 2,66	*Practical examples* 3,43	*Metaphors* 3,11	*Case studies* 3,25

Fig. 4.21 | Synthesis of the evaluations to determine the preferred tools.

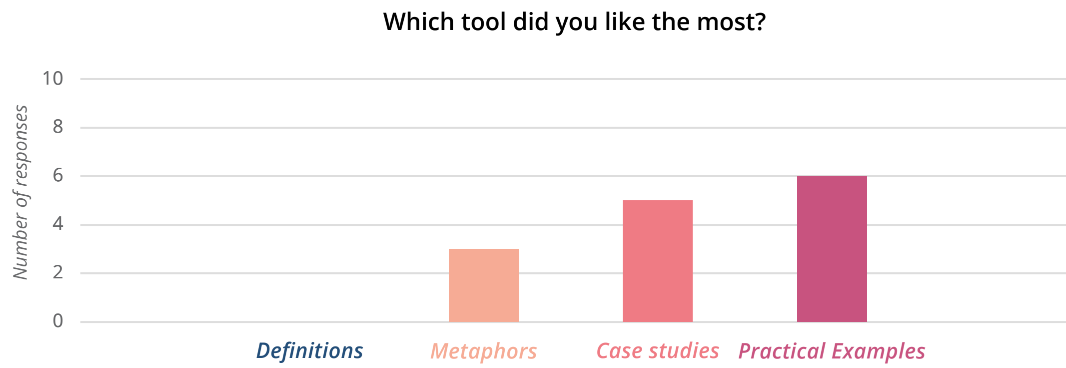


Fig. 4.22 | Participants' explicit preference expression.

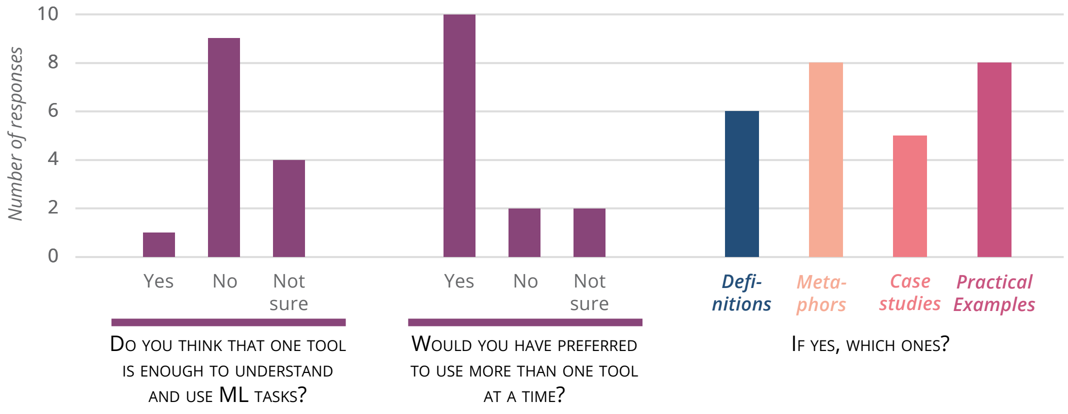


Fig. 4.23 | Preferences for combining ML Pills.

4.2.2.6 Design approach and value for humans

One of the main concerns of launching a design activity mainly based on and explicitly requiring the integration of ML was to convey a technology-driven approach that could lead to replicating engineers and computer scientists' paths. To avoid that and to encourage designers to maintain their human-centered approach, the design briefs required that the envisioned intervention aimed to add quality to people's experiences and bring originality. This was not intended in terms of radical innovation but as the demonstration of a different mindset to frame the problem with respect to current ML-infused solutions.
To assess these aspects, the researchers independently evaluated the ideas on a 4-point Likert scale (1: Not at all; 2: Slightly; 3: Moderately; 4: Very much), and a voting session among peers required the participants to distribute 15 votes among the concepts of their colleagues' proposals. The results of both assessments are presented in Tables 4.2 and 4.3.
The interventions ideated to respond to the first brief could be distinguished into two perspectives about achieving valuable solutions: one focused on practical and utilitarian needs, the other on emotional aspects. The latter reached more successful results in the peer evaluation session both in terms of happiness and originality. According to the researchers, the nature of the former coincided with quite low originality scores, as any ML programmer could have conceived them. However, the ratings showed that almost all the ideas added some value to people's experiences, with some doubts for LW01, which explanation and relevance were not very clear.
The second brief was perceived as more human-centered by the participants, and it reflected in both evaluations. Though, originality suffered because some solutions were recommendation systems.
Interestingly, among the 14 students who answered the final questionnaire (Fig. 4.24), none felt to have followed a technology-driven path. The majority (9) stated that their design experience was human-centered, whereas the remaining five indicated both. Indeed, three modalities to approach the design process emerged: in one case, it was a data-driven strategy, reflecting on the available or collectible data to envision solutions. A second approach was about transposing the ideas expressed in the tools (especially those portraying examples) directly into the brief context. In contrast, the most common included a more traditional brainstorming of possibilities based on the

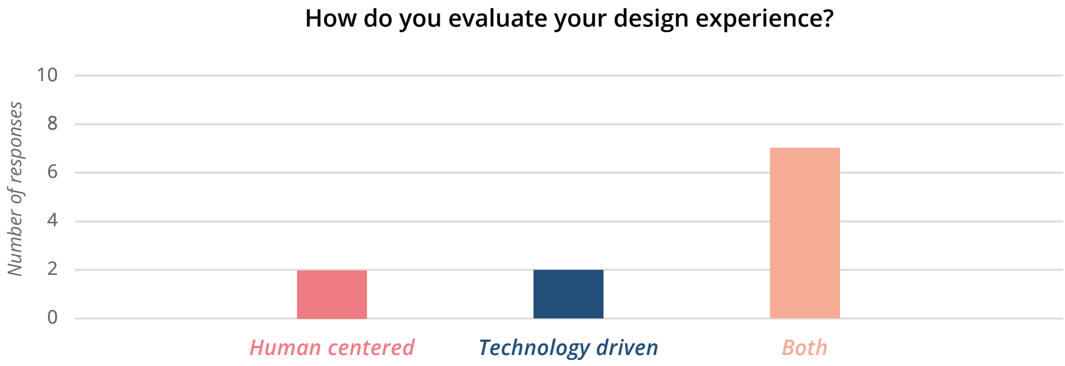


Fig. 4.24 | Participants' assessment of the design process followed.

issues emerging from the context and people's needs and a subsequent reflection on how to make ML fit and expand the selected idea.
Therefore, even if no prescriptive indications about the design process were given, the preferred way to approach the design activity remained (in almost all cases) the human-centered one to which they were used. However, despite the good intentions of the proposals, some manifested features that concern from an ethical standpoint. Just to cite an example, LE02 required a lot of personal and sensitive data for the system to create a community without thinking about the possible implications.
This result might be conditioned by the lack of time and the fact that the focus was not on ethical issues. Anyways, more explicit reasoning on ethical matters could also prevent cases of dubious morality.

4.2.2.7 Overall workshop experience

From their feedback, a predominantly positive perception characterized the participants' didactic experience. This clearly reflected the changed perspective on the self-assessment of ML knowledge level and the capability to include it in projects (Fig. 4.25).
Among the few comments received, most were enthusiastic, like *"Amazing work, well done!"* or *"I really enjoyed this workshop. Thank you!"* and others appreciated the workshop atmosphere.

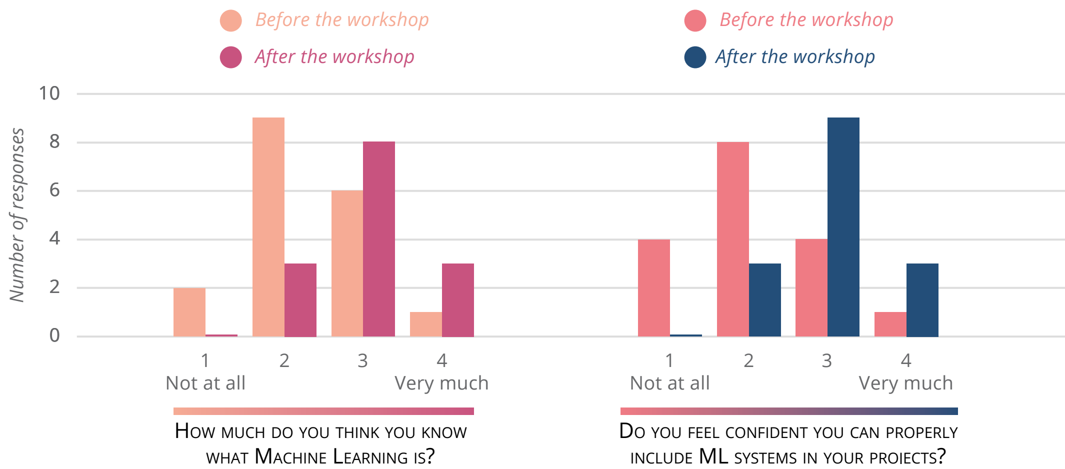


Fig. 4.25 | Self-perceived ML knowledge by students before and after the workshop.

However, also some constructive comments were received. One suggested having bigger groups to exchange more ideas and foster creativity. Indeed, it is consistent with the organizers' view. However, for the specific research objectives (to evaluate the possible forms and languages for the translation), it was preferred to opt for a greater number of groups having different experiences with the tools. More introduction to the design activities and more specific challenges have also been requested, but again contrasted with the research objectives. A better understanding of ML limitations and cost was also required to have a clear view of the implications of these systems. These are all valuable indications that can be implemented in further iterations of the experimentation, focusing on the learning activity. Finally, one of the relevant assumptions of the thesis was brought to light, suggesting collaboration with engineers.

4.2.3 Discussion

As a first experimentation of translating ML knowledge to design students, the outcomes were promising, and the selection of the presented contents seemed appropriate. The synthesis of ML systems in the essential characteristics defining agents was naturally apprehended and proved useful for the participants to present their ideas concisely. As well, independently from the ML Pills received, grasping ML capabilities from ML tasks seemed to enable students to envision new solutions by intuitively applying them in their design process. Most groups were able to extract the essential qualities of the ML capabilities from the ML Pills and transpose them in practice to support their ideas, and only a couple of them proposed solutions very similar to the examples provided.

With regard to the focus of the experimental educational activity, understanding the preferred forms to communicate ML knowledge to design students, clear indications can be drawn. The workshop reinforced the **importance of examples**, which was already evident in the case studies analysis of outreach strategies. Even though it might be difficult or uncertain to infer ML capabilities only through examples, Practical Examples pills have been pointed out as the tools that favored understanding the most, immediately followed by Case Studies. The background of the participants could have influenced their preference for practical applications. In fact, because of their educational path in the Digital and Interaction Design MSc, they were used to prototyping and "getting their hands dirty" with digital technologies. It is natural that this would be a more familiar language for them.

Case Studies were also a familiar tool to visualize the possibilities offered by ML, and Metaphors effectively worked as the kind of designerly abstractions that (Yang, 2018) suggests developing to facilitate designers' comprehension and exploration of ML. Instead, regardless of the efforts to translate technical notions into a more appropriate language and with synthetic communication strategies, Definitions were the most challenging form to grasp and operationalize ML capabilities. The participants found them too abstract, and the lack of visual or referential components made them very difficult to deal with.

Combining the features of the different ML Pills and enabling design students to decide how to navigate the contents, as one of them remarked, might be the most successful solution, but acknowledging the primary importance of examples and references to make sense of ML capabilities.

Moreover, despite setting originality as an objective, allowing students to experiment without expecting innovative solutions is more appropriate for them to start building their practical references. As (Antonelli, 2018) said, it is normal that, at the beginning, the design outputs are not perfect, and even monstrosities arise. Indeed, she adds, *"Every time a new technology is introduced, there is a moment of drunkenness because everybody experiments as they try to do their best. Then sobriety sets in, a mature baseline evolves, and people develop a critical sense".* Thus, once ML capabilities will be consolidated as design tools, designers should be able to focus on the core issues of the challenges and express their innovative potential.

Finally, even though there was not enough time and it was premature to focus on ethical issues before understanding how to introduce ML to design students, reflections about the limitations and impacts of this technology could have helped outline a more comprehensive picture of the potentialities of ML systems. This was perceivable in some concepts where personal data were indiscriminately used as inputs to not equally balanced outputs, and it was also reported in some comments. Moreover, ML systems were clearly intended as technical tools, and nobody enlarged the perspective to a broader social dimension. Again, this was not required nor encouraged in the short time available, but the results were very individualistic solutions that could easily fail in social contexts. Indeed, the participants generally addressed the design briefs with a human-centered approach, but being explicitly aware of the consequences and larger impacts of their decisions could dramatically increase the benefits they can bring.

4.3 ML Agents. The ultimate synthesis for knowledge transfer

Based on the *ML Pills for Designers* workshop results, the translation of the ML tasks has been reframed. To prepare the stage for wider sociotechnical considerations, the **connection with the definition of ML systems as agents** was made even more explicit and introduced in the title. Directly inspired by the worldwide reference AI textbook (Russell & Norvig, 2020), it demonstrates its relevance as it is at the basis of both the definition of AI given by the European Commission (High-Level Expert Group on Artificial Intelligence, 2019) – the most comprehensive in the author's view – and their framing as sociotechnical systems (van de Poel, 2020).

4.3.1 Structure and layout

ML Agents (Fig. 4.26) are a **transfer tool for foundational ML knowledge**. They synthesize and portray ML tasks, as categorized in the *ML Designerly Taxonomy*, to communicate ML capabilities. While maintaining technical correctness and clarity on their computer-based nature, the analogy with human agents helps designers understand how they can be part of larger systems. In a sense, they are a combination of the Definitions, Case Studies, and Practical Examples ML Pills. Although inspiring, fostering creativity, and possibly innovation by enabling personal abstraction, additional metaphors have not been implemented in *ML Agents* to focus on a more accurate knowledge transfer and less ambiguous

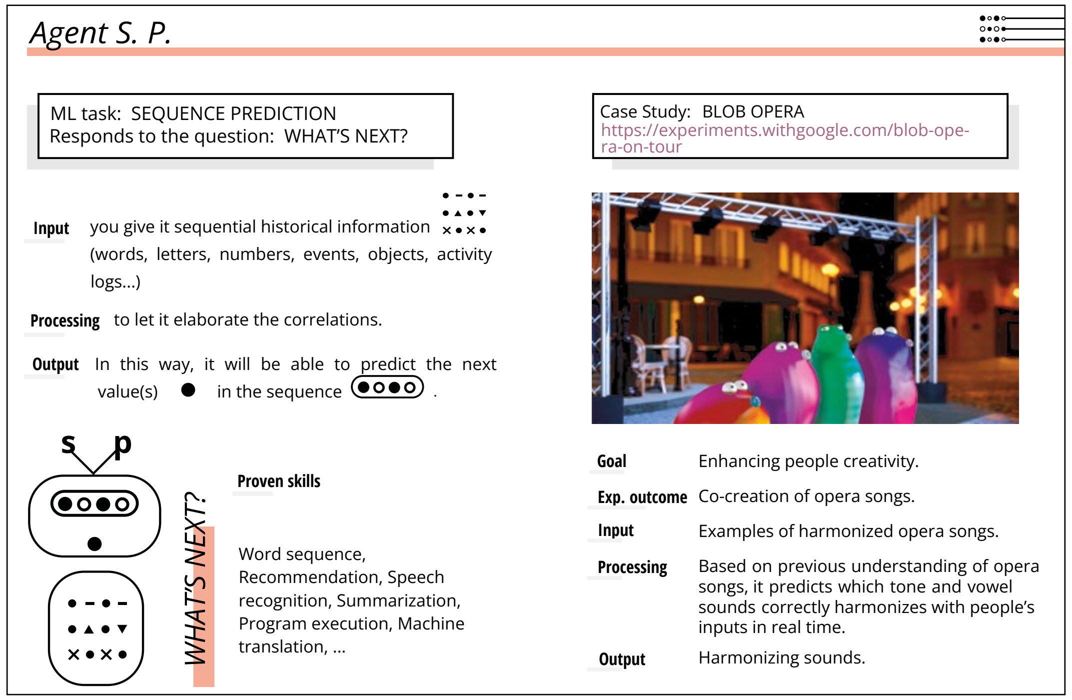


Fig. 4.26 | ML Agent example: Agent S.P. (Sequence Prediction).

message. Somehow, the *ML Agents* are already metaphorical figures themselves as they bring ML capabilities to life to remind of their agency while maintaining their artificial nature. Thus, for an introductory level, adding another layer of meaning was avoided not to cause confusion. However, for advanced steps of awareness (possibly subsequent to a broad understanding of the matter at hand), they can be exploited to look for ways to improve or disrupt current solutions.

In their current format, a further simplified definition depicts each ML agent following a clearly displayed input–processing–output structure. The reframed definitions explicitly describe the roles that human developers (addressed as an active part) and artificial agents play in producing the output. Moreover, they integrate graphic elements that ultimately generate visually appealing characters. Input and output symbols are included in the head of the robot-like figures, while what is in the body gives hints about their internal processing.

Like in the Definition ML Pills and following the example of Holbrook and Lovejoy's (2017) effective communication, an easy-to-remember question synthesizes the operating principle of each ML Agent. Here, the synthetic statements have all been homologated to the interrogative form to better portray the task beneath each agent. In addition, examples of applied capabilities from the corresponding dimension of the *ML Designerly Taxonomy* are provided in the form of skills of the ML Agent.

Moreover, a framed case study exemplifies the agent's capability and can also be personally explored to make sense of how the ML Agent can work and how it is built. This is halfway between a Case Study and a Practical Example ML Pill. Though, the structure of the explanation was modified to underline the role of people in the definition of a ML system. Its goal and expected outcome outline its basic design requirements. Then, the core particles characterizing an agent

(input-processing-output) are described. This also facilitates the interpretation of the definition as it creates an immediate connection.

The case studies have been selected according to their explicit manifestation of the ML task they stand for, their interactive nature, and – when possible – the availability of information about how they have been built. Of course, they can be easily complemented by more common (even if less explicit) examples from our daily lives, and they can be modified as needed to meet the requirements of specific contexts or briefs.

4.3.2 Strengths and limits

Taking advantage of the practical experience and insights collected during the *ML Pills for Designers workshop*, the *ML Agents* combine the qualities of different forms and languages for the translation, giving the freedom to choose how to approach the understanding of ML capabilities according to personal preferences and attitudes. Even though practical experiences proved to be very effective in relation to designers' way of learning, this might be the result of the specific background of the workshop participants. Willing to be as inclusive as possible in terms of designers' specializations, the *ML Agents* are not limiting the translation to hands-on learning. They try to maximize visual inputs, minimize textual cognitive load, and amount of information to meet the way of making sense of ML capabilities as most prominently emerged during the workshop and from secondary research (Yang, 2018; Dove & Fayard, 2020; Zdanowska & Taylor, 2022). An additional strength lies in the attention to avoid the sociotechnical blindness described by Johnson and Verdicchio (2017) by including people's role in determining how *ML Agents* work.

The *ML Agents* have been conceived as a printed booklet to support students' design activities and to support oral explanations. However, with further digital support, other strategies could be implemented to increase the potentialities of the visualizations. Indeed, despite the efforts to combine symbols and definitions to finally generate a character, animations like in R2D3's Visual Introduction to ML (n.d.) could provide a more immediate comprehension of how *ML Agents* work. Additionally, depending on the time at disposal, not all the parts of the *ML Agents* might be explorable, limiting the effectiveness of the multi-format translation. Though, they are a very flexible tool which contents and support can easily be adapted to include more varied examples, to make sense of everyday life products and services, and even to introduce more practical activities and possibly prototyping supports. For this, transferring *ML Agents* to a digital platform would be ideal, and it is planned as a future iteration.

4.4 Ethical concerns for a comprehensive awareness of ML

The first practical experimentation of the translation highlighted that, when designing for ML systems, relying on a human-centered approach, as we have been educated, is not enough. For richer and more valuable solutions, **a systemic perspective is necessary**. It would keep into account artificial artifacts and their impacts, as well as human agents, their needs, and the values to respect or promote for a more flourishing life.

4.4.1 Designing with and for values

Any action or artifact that people produce is not neutral, but it influences or impacts others to some extent. The issue is even more relevant when it comes to novel and still uncertain fields of experimentation. As von Schomberg (2013) noticed, *"techno-scientific applications can remain ethically problematic, even in cases where scientists and engineers have the best possible intentions and users have no conscious intention to misuse or abuse."*

Drawing from Responsible Research Innovation (RRI), some suggestions can be derived to design (ethically) acceptable, sustainable, and societally desirable artifacts. These include **(i) having grand challenges and the right impacts** to direct the design process, **(ii) anticipating positive and negative impacts** and **(iii) assessing the performance of the technology**, and **(iv) embedding values** (von Schomberg, 2013). If some of these indications are openly part of the design culture and even gave birth to subdisciplines like design futures, others are more tacit practices. However, incorporating values and moral considerations in the design of technology and any kind of product, service, or experience can *"shape the space of action of future users"* (van den Hoven, 2013). It can determine people's agency, nudge their behavior (Thaler & Sunstein, 2009), and define the affordances and constraints of the artifacts themselves.

Being strictly interrelated with the design domain, this could naturally be part of designers' practice. In particular, Van Allen (2018) attributed them *"an important positionality and value system that need a place at the table in creating new directions for AI itself,"* as these systems have *"built-in tendencies to promote or demote the realization of particular values"* (Brey, 2012 in Umbrello, 2020).

Then, according to Van den Hoven (2013), designers should learn and develop the ability to incorporate values into artifacts by reasoning about them not only in a conscious way but also explicitly. And – I would add – this kind of reflections should be shared with the entire design team. Embedding values into sociotechnical systems needs to happen from the very early phases of the design process. It has to respect two conditions: (i) the use or interaction with an artifact should achieve or contribute to a value because (ii) the system has been expressly designed for that value (van de Poel, 2020). The author underlines two elements: the intentionality of the practice and the fact that not all the components of the sociotechnical system should promote or preserve the intended value. Still, it ultimately has to be reached.

Within the field of ethics and technology, in the 1990s, Batya Friedman (2019) developed a methodology to advocate human principles when planning technology and support this process: Value Sensitive Design (VSD). Not being linked to any specific technology, it has found new relevance with AI and ML applications, as they bring new ethical challenges. For this purpose, Umbrello and Van de Poel (2021) proposed a value-sensitive design process for AI. In their adaptation of the original approach, they consider the entire life cycle of AI systems to be able to monitor the evolution of the algorithms over time and early detect possible unintended consequences. The process is articulated into four iterative phases. Similarly to any design process, it begins with a *context analysis*, which should include contextual variables such as societal challenges, existing technology, and systems, as well as the values and needs of the actors of the sociotechnical system.

The second crucial phase is *value identification*. For this task, the original framework proposed a tripartite strategy that included **conceptual investigations** to identify the values and possible trade-offs of direct and indirect stakeholders in theory, **empirical investigations** to expand and fine-tune them through practical research, and **technical investigations** aimed to point out issues and possibilities that specific technological applications might raise.

Umbrello and van de Poel, instead, introduce another important specification by distinguishing between promoted and respected values. They stress that it is not enough to avoid harm, but it is important to actually try to contribute to socially desirable objectives. Therefore, they suggest an explicit orientation toward positive impacts, like the Sustainable Development Goals of the United Nations. Then, values expressly identified as relevant for AI systems, like the *Ethics Guidelines for Trustworthy AI* (High-Level Expert Group on Artificial Intelligence, 2019), and those emerging from the context analysis should be considered.

The third phase involves the formulation of design requirements based on identified values (phase 2) and contextual analysis (phase 1) . To conclude, a prototyping phase should test if the system meets the design requirements and values are actually incorporated.

The concepts underlying value-sensitive design and RRI are not alien to the design discipline but, often, these reasonings occur implicitly and based on designers' sensibility. Instead, to build an educational method addressing the complexity of dealing with ML systems and current challenges, it is essential to make ethical dimensions explicit (Frascara, 2020) and provide design students with the theory beyond the problem space. This would avoid issues being treated superficially and just based on their own experience (Weil & Mayfield, 2020). In this sense, (Redström, 2020) recommends a much more theoretical design education as a way to be better prepared to anticipate how the future might unfold. This is why the research developed to comprehend how to structure these ethical considerations as a resource to enhance the envisioning of responsible and meaningful ML applications.

4.4.2 Method. Mapping ethical guidelines for a systematization

To introduce a value-driven approach to the design of ML-infused solutions, a fundamental step consists of understanding which values are uniquely relevant to their development and materialization. On this matter, Umbrello and van de Poel (2021) suggested the *Ethics Guidelines for Trustworthy AI* (High-Level Expert Group on Artificial Intelligence, 2019) as a reference. These are very comprehensive and well-articulated. However, plenty of ethical guidelines related to the design of AI and ML systems are flourishing across public and private sectors. The *AI Ethics Guidelines Global Inventory* (Algorithmic Watch, 2020), the largest online repository, collects 167 guidelines. It was taken as a starting point for a systematic exploratory content analysis to identify which values are significant for designing AI and ML systems and the related issues. Indeed, this activity aimed to determine the limitations, risks, implications, impacts, possibilities to overcome them, and opportunities for improvement that could be connected to the promotion or preservation of values in order to depict the correlations between them. These connections could be systematized in a framework to support design students envisioning responsible ML-infused solutions.

The documents in the *AI Ethics Guidelines Global Inventory* are categorized by sector or actor (academia, civil society, government, industry association, intergovernmental organization, international organization, private sector, professional association, religious institution, other), type (binding agreement, voluntary commitment, recommendation), and location.

Some parameters have been defined to select which guidelines should be included in the study. First, only the documents available in English, directly addressing AI (for instance, those focusing on robotics were discarded), and listing at least one value that AI or ML systems should embody were considered. Additionally, guidelines only portraying values (with no further explanations), instructions at a systemic level (e.g., international policies or governmental strategies), or design best practices at a technical level (in the domain of data and computer scientists and engineers) were excluded from the analysis. In fact, the first presented only redundant information, while the latter were not pertinent to the design scope. To complement this set of documents, particularly relevant and non-represented papers illustrating principles and indications for designing AI and ML systems and design ethics tools to integrate with suitable values for the UX of ML-infused artifacts were also included. The final selection counted 60 documents: academia (6), civil society (23), intergovernmental organizations (1), international organizations (1), private companies (15), professional associations (2), literature (3), design ethics tools (9). All the retained documents are depicted in Tab. 4.5.

The first cycle of the content analysis combined different *exploratory coding* methods, and it was followed by two additional iterations (Saldaña, 2009). *Hypothesis and provisional coding* guided the approach to the analysis by defining a predetermined list of codes to start screening the contents. Of course, the codes could and have evolved throughout the study to adapt to the actually available contents and, for instance, to comprehend all the detected nuances of values.

Provisional coding anticipates categories based on previously framed knowledge. In this case, the thorough argumentation of the *Ethics Guidelines for Trustworthy AI* (High-Level Expert Group on Artificial Intelligence, 2019) represented a useful baseline for organizing the investigation. Hence, a distinction between principles (as foundational elements for a system of belief) and values (as requirements or smaller components of principles) was introduced. The initial set of predetermined codes included the four ethical principles *(respect for human autonomy, prevention of harm, fairness, and explicability)* and the seven requirements indicated in the European guidelines: *(1) human agency and oversight, (2) technical robustness and safety, (3) privacy and data governance, (4) transparency, (5) diversity, non-discrimination and fairness, (6) environmental and societal well-being, and (7) accountability* (High-Level Expert Group on Artificial Intelligence, 2019).

Promotion of flourishing was also added to the list of principles, in line with the explicit orientation toward good advocated by (Umbrello & van de Poel, 2021), and the articulation of bioethical principles that include beneficence (Floridi et al., 2018). In antithesis to the *prevention of harm*, it suggests a more proactive attitude towards life, well-being, growth, progress, and prosperity.

Name	Issuer	Sector	Year	Location	Type
Vienna Manifesto on Digital Humanism	*Faculty of Informatics, TU Wien*	Academia	2019	Austria	Volun. commitment
Trustworthy Use of Artificial Intelligence	*Fraunhofer Institute for Intelligent Analysis and Information Systems IAIS*	Academia	2019	Germany	Recommendation
The Japanese Society for Artificial Intelligence Ethical Guidelines	*Japanese Society for AI*	Academia	2017	Japan	Volun. commitment
Understanding artificial intelligence ethics and safety	*The Alan Turing Institute*	Academia	2019	UK	
A Framework for the Ethical use of advanced Data Science Methods in the Humanitarian Sector	*The Humanitarian Data Science and Ethics Group*	Academia	2020	European Union	Recommendation
Montreal Declaration for Responsible AI	*Université de Montréal*	Academia	2018	Canada	Volun. commitment
The Toronto Declaration	*Amnesty International*	Civil society	2018	UK	Recommendation
AI4People's Ethical Framework for a Good AI Society	*Atomium - EISMD (AI4Poeple)*	Civil society	2018	European Union	Recommendation
Algo.Rules	*Bertelsmann Stiftung / iRights.Lab*	Civil society	2019	Germany	Recommendation
Digital Decisions	*Center for Democracy & technology (CDT)*	Civil society		US	Recommendation
Governing Artificial Intelligence. Upholding Human Rights & Dignity	*Data & Society*	Civil society	2018	US	Recommendation
Data Ethics Principles	*DataEthics.eu*	Civil society	2017	Denmark	Recommendation
Principles for Accountable Algorithms and a Social Impact Statement for Algorithms	*FAT/ML*	Civil society		International	Recommendation
Asilomar AI Principles	*Future of Life Institute*	Civil society	2017	US	Volun. commitment
Holberton Turing Oath	*No organization*	Civil society		International	Volun. commitment
Open AI Charter	*Open AI*	Civil society	2018	US	Volun. commitment
Privacy and Freedom of Expression In the Age of Artificial Intelligence	*Privacy International & Article 19*	Civil society	2018	UK	Recommendation
Principles for the Governance of AI	*Science, Law, and Society (SLS) Initiative*	Civil society	2017	US	Recommendation
The Good Technology Standard	*The Good Technology Collective*	Civil society	2018	International	Recommendation
The Responsible Machine Learning Principles	*The Institute for Ethical and Machine Learning*	Civil society		UK	Recommendation
Artificial Intelligence and Machine Learning Policy Paper	*Internet Society*	Civil society	2017	US	Recommendation
Civil Rights Principles for the Era of Big Data	*The Leadership Conference on Civil and Human Rights*	Civil society	2014	US	Recommendation
Data Ethics Canvas	*The Open Data Institute*	Civil society	2019	UK	Recommendation
Universal Guidelines for Artificial Intelligence	*The Public Voice*	Civil society	2018	International	Recommendation
Preliminary study on the Ethics of Artificial Intelligence	*UNESCO*	Civil society	2019	France	Recommendation
Top 10 Principles for Ethical Artificial Intelligence	*UNI Global Union*	Civil society	2017	International	Recommendation
Principles for Responsible AI	*Women leading in AI*	Civil society	2019	International	Recommendation
White Paper: How to Prevent Discriminatory Outcomes in Machine Learning	*World Economic Forum*	Civil society	2018	International	Recommendation
Principles for responsible stewardship of trustworthy AI	*G20*	Intergovern-mental org.	2019	International	Volun. commitment
Recommendation of the Council on Artificial Intelligence	*OECD*	International organization	2019	International	Recommendation
Universal Principles of Data Ethics	*Accenture*	Private sector	2016	US	Recommendation

Name	Issuer	Sector	Year	Location	Type
Saftey and Ethics	*Deep Mind*	Private sector		US	Volun. commitment
People & AI Partnership Guidebook	*Google*	Private sector		US	Recommendation
IBM's Principles for Trust and Transparency	*IBM*	Private sector	2018	US	Volun. commitment
Everyday Ethics for Artificial Intelligence	*IBM*	Private sector		US	Recommendation
The Future Computed – AI and its role in society	*Microsoft*	Private sector	2019	US	Recommendation
Five guiding principles for responsible use of AI in healthcare and healthy living	*Philips*	Private sector	2020	Netherlands	Recommendation
A practical guide to Responsible AI	*Price Waterhouse Coopers UK*	Private sector	2019	UK	Recommendation
he Ethics of Code: Developing AI for Business with Five Core Principles	*Sage*	Private sector	2017	US	Volun. commitment
Sony Group AI Ethics Guidelines	*Sony Group*	Private sector	2019	Japan	Volun. commitment
Principos / Principles	*Telefonica*	Private sector	2018	Spain	Binding agreement
Telia Company Guiding Principles on trusted AI ethics	*Telia Company*	Private sector		Sweden	Volun. commitment
Unity's six guiding AI principles	*Unity*	Private sector	2018	US	Volun. commitment
7 Principles of Designing Good AI Products	*UX Studio Team*	Private sector	2018	Hungary	Recommendation
Vodafone AI Framework	*Vodafone Group*	Private sector	2019	UK	Volun. commitment
Ethically Aligned Design	*IEEE*	Professional association	2019	International	Recommendation
Responsible AI: Global Policy Framework	*ITechLaw*	Professional association	2019	US	Recommendation
How to Design AI for Social Good: Seven Essential Factors	*Floridi, L., Cowls, J., King, T. C., Taddeo, M.*	Literature	2020	Science and Engineering Ethics, 26(3), 1771–1796.	Journal article
Human-Centered Machine Learning.	*Holbrook, J., Lovejoy, J.*	Literature	2017	Medium	Web page
How AI can be a force for good.	*Taddeo, M., Floridi, L.*	Literature	2018	Science, 361, 751–752	Journal article
Tarot Cards of Tech	*Artefact*	Design tools	2017	US	Card deck + web
Ethical Explorer – Field Guide	*Omidyar Network*	Design tools	2020		Booklet + web
Ethical Explorer – Risk Zones	*Omidyar Network*	Design tools	2020		Card deck + web
Ethics for Designers – Moral Agent	*Jet Gispen*	Design tools	2017	TU Delft	Card deck
Ethics for Designers – Moral Value Map	*Jet Gispen*	Design tools	2017	TU Delft	Map
Intelligence Augmentation Design Toolkit	*Futurice*	Design tools			Booklet, canvases, maps, cards
AI Ethics Cards	*IDEO*	Design tools	2019		Card deck
AI Blindspot	*Assembly program*	Design tools	2019	MIT Media Lab	Card deck + web

Tab. 4.5 | List of documents analyzed for the content analysis of ethical guidelines.

Hypothesis coding, instead, is based on the researcher's assumptions about what is expected to be found. In this case, the objective was to identify values, limitations, risks, implications, impacts, possibilities, and opportunities, so these were included in the initial codes to direct the analysis. To better specify these broad categories, an additional level of *initial coding* was implemented by using analytic memos to closely examine the nuances of data, and a second cycle of *focused coding* organized them into categories based on the detected differences and similarities.

Inevitably, the study shares the typical limitations of qualitative methods, like the subjectivity of the selection and analysis. The language limitation surely skewed the dataset, and the fast pace at which guidelines are emerging may have resulted in missed documents, even though it was not detected in a later check for its comprehensiveness. Additionally, the kind of analyzed materials can be categorized as grey literature. However, using the AI Ethics Guidelines Global Inventory as a primary reference and a systematic methodology should enhance the validity of the research.

4.4.3 Outlining the results of the systematic content analysis

Identified principles	Coded values (n. documents / n. groups)
	Trustworthiness / Trust (23/8)
Promotion of FLOURISHING (20/8)	Beneficence (7/4), **Well-being** (22/7), Democracy (23/7), Societal/Social beneficence (9/5), **Sustainability** (24/8), Environmental beneficence (8/5)
Prevention of HARM (21/6)	Non-maleficence (2/2), Data protection (17/7), Access to data (1/1), Data quality (1/1), Data governance (5/4), Privacy (46/8), Accuracy (25/7), Security (39/7), Safety (44/8), Resilience (9/6), Robustness (23/7), Consistency (24/8), **Reliability** (29/7)
Attention to FAIRNESS	Human rights (36/8), Human dignity (14/6), Universal design (3/3), **Accessibility** (26/6), Fairness (44/8), **Representativeness** (24/8), Inclusivity (31/8), Impartiality / Neutrality / (avoid) Bias (48/8), Equality (34/8), Non-discrimination (48/8), Diversity (2/2), **Justice** (24/8)
Increase of INTELLIGIBILITY	Accountability (37/8), Responsibility (52/8), Auditability (17/4), **(Meaningful) transparency** (46/7), Communication (34/7), Traceability (8/5), **Explainability** (39/8), Explicability (17/6), Interpretability (15/6), Measurability (9/6), Openness (9/4), Intelligibility (26/7), Reproducibility (9/4)
Respect of HUMAN AUTONOMY	**Human agency** (27/7), Oversight (18/7), **Freedom** (26/7), Human autonomy (15/7), Human control (11/5)

Tab. 4.6 | Principles and values identified from the content analysis. Highlighted are the most representative and comprehensive values.

Very soon, the analysis demonstrated that the predetermined list of codes (from both *hypothesis and provisional coding* methods) was unsuitable for the content.

Concerning the categorization of principles and values, a first attempt tried to identify both levels, differentiating when the terms were used as overarching principles or not. However, often this distinction was not present in the guidelines, so the principles that could have a counterpart within the values were quantified as such. At the same time, *promotion of flourishing* and *prevention of harm* were traced in the documents, as they have a slightly different meaning than the corresponding non-maleficence and beneficence.

A multilevel organization of the codes also proved flawed. With the systematization of the European guidelines as a reference, the initial idea was to add a layer of more

RISKS (what threatens the realization of the value?)

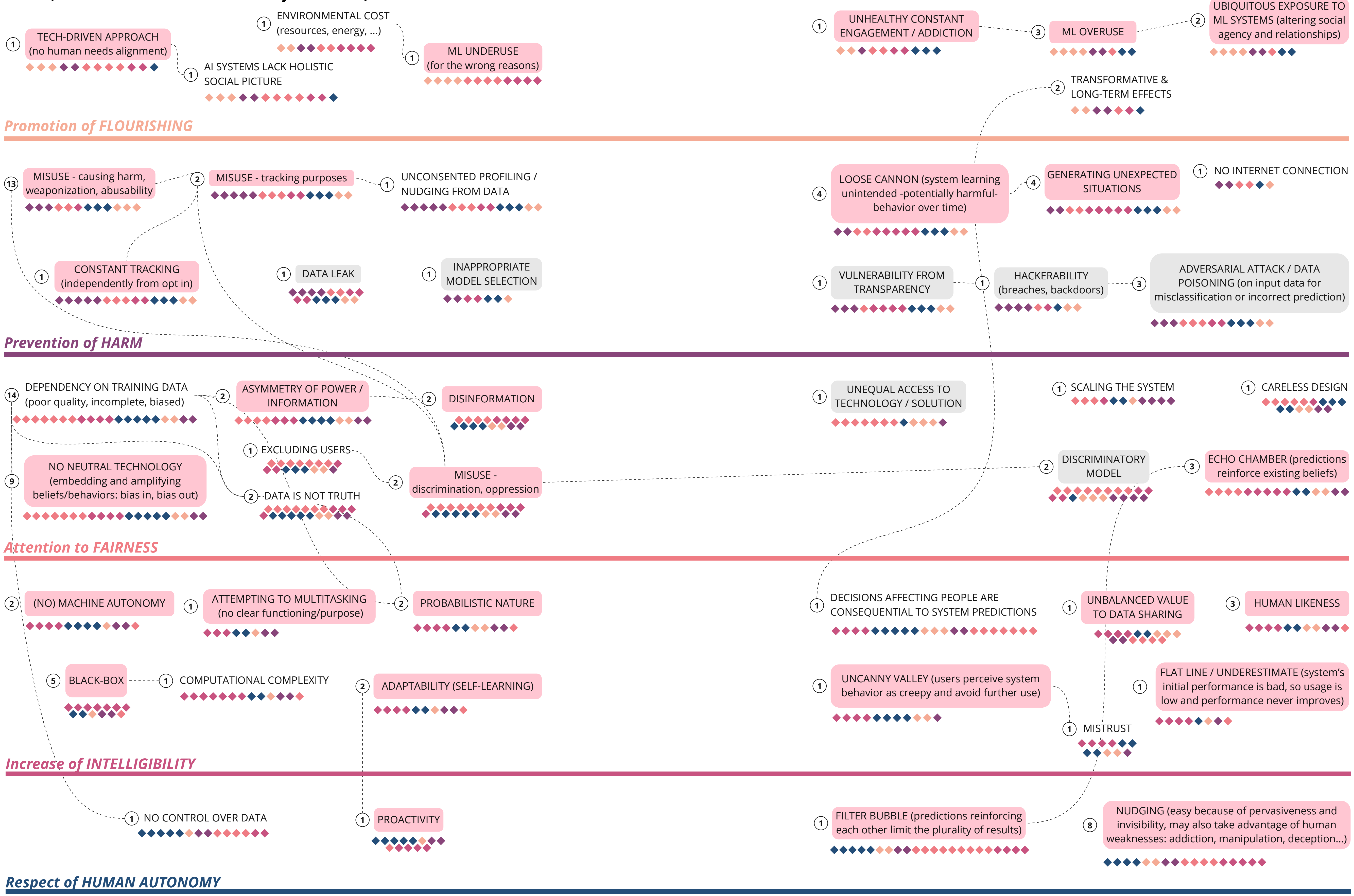


Fig. 4.27 | Risks resulting from the content analysis. Highlighted in pink are the risks retained more significant for designers to develop responsible ML solutions, and in grey are those that ML experts can tackle. Risks are associated with a number of values, represented by the diamonds (that correspond to the principle indicated by the color), and connected to similar ones.

POSSIBILITIES (what can be done to pursue the value?)

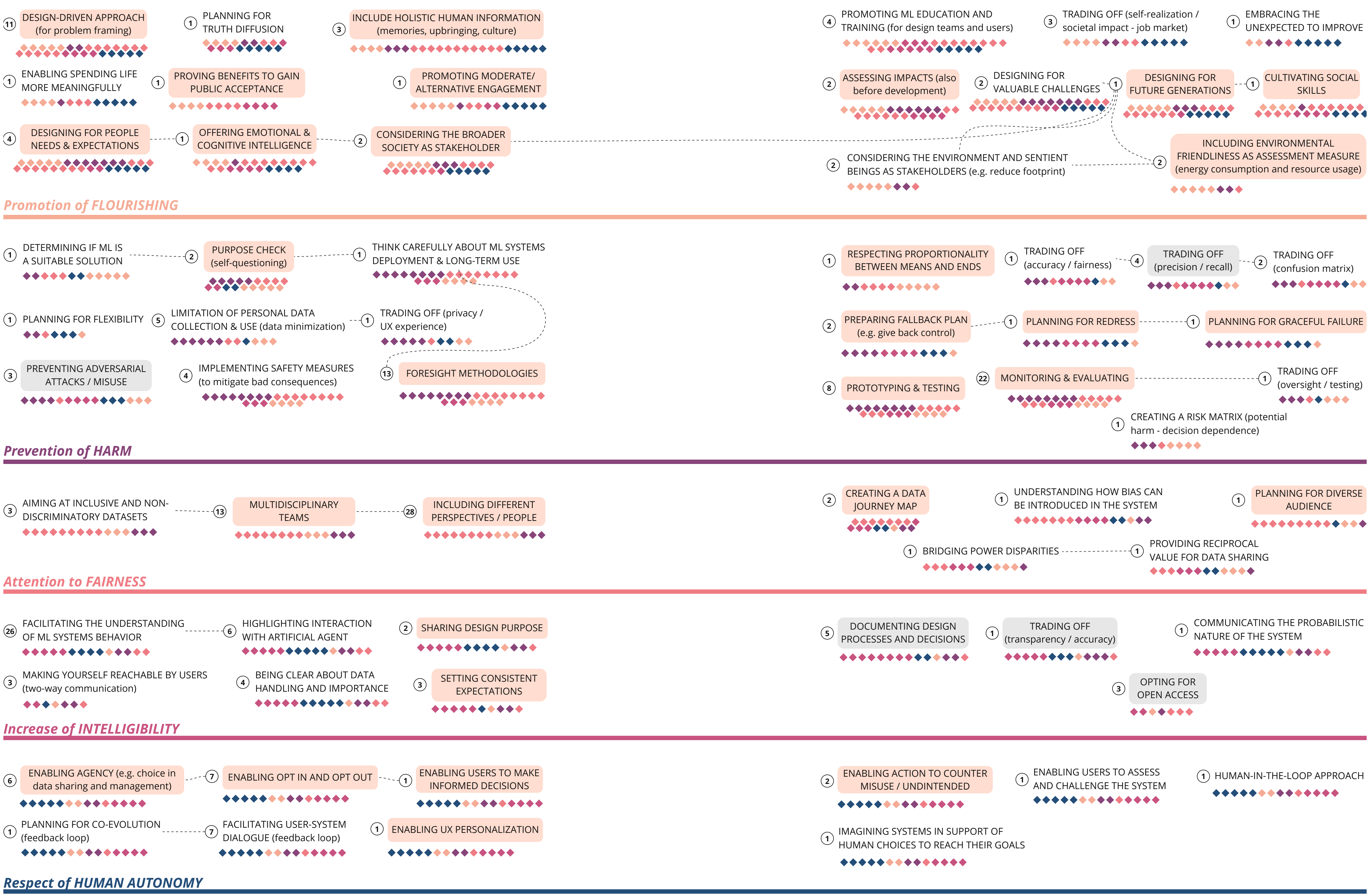


Fig. 4.28 | Possibilities resulting from the content analysis. Highlighted in orange are the possibilities that correspond to common design practices, and in grey are those that ML experts can tackle. Possibilities are associated to a number of values, represented by the diamonds (that correspond to the principle indicated by the color), and connected to similar ones.

grained values beneath the predetermined ones. Nonetheless, building exclusive relationships was very difficult, as they are very interconnected.
Finally, a list of five principles and 50 values was compiled. Tab. 4.6 portrays a broad categorization of the principles and their most strongly related values. The term *intelligibility* replaced *explicability* in the principles to include the possibility to grasp or understand something not only through a formal explanation but also intuitively. Additionally, as it was recurring in many documents, *trustworthiness* was noted but kept in an overarching position to indicate how it embraces all the principles. The values, instead, are not framed in any hierarchy but the most representative and comprehensive ones have been highlighted in the table.
Categorizing the issues related to the design of AI and ML systems also proved problematic. Indeed, a fine-grained distinction between limitations, risks, implications, impacts, possibilities and opportunities was not unambiguously identifiable. For this reason, *risks* and *possibilities* were assumed as the main themes to be specified through the *initial and focus coding* activities. These were more effective because they respectively encompassed all the threats that can hinder the realization of values and the suggestions to pursue, promote, or preserve them.
In total, 47 risks and 67 possibilities were identified. However, the attempt to build a theoretical systematization, clearly highlighting which *risks* and *possibilities* related to effectively embedding specific values in ML applications, failed. Indeed, some risks could be opportunities as well as limitations and several concealed multiple interpretations and could connect with as many values. *Misuse*, for instance, was declined in three different ways to show which values it could threaten: *privacy* if intended for tracking purposes, *security* if meant to cause harm, or *non-discrimination* if tied to prejudice, intolerance, or oppression. Additionally, despite excluding very technical and high-level systemic guidelines, some risks could not be directly addressed by designers as they entail more specialized competences and roles.
Similar issues occurred in identifying relationships between *values* and *possibilities* to promote or preserve them, as the same remedies could be beneficial for different problems. Moreover, most of the suggestions were very generic and could not provide significant guidance.
The plurality of links for both *risks* and *possibilities* is portrayed in Figures 27 and 28.

4.4.4 Elaborating findings. When rigid frameworks are not suitable.

Overall, the qualitative content analysis of the ethical guidelines to design AI and ML systems portrayed the complexity of the issue. While a vast range of values and principles can be drawn, it is difficult to clearly distinguish what are AI and ML limitations, implications, impacts, risks, possibilities, and opportunities from current theoretical argumentations. Indeed, these concepts are so much interrelated that they can easily overlap. For instance, having a probabilistic nature that feeds on data provided by human agents is one of the core characteristics of ML systems. Hence, having biased datasets is a plausible risk that represents an inherent limitation of this technology. In turn, it would implicate erroneous decisions and could impact people in many ways, from prohibiting access to social services or medical care to trivial wrong film recommendations. At the same time, biased datasets can be an opportunity to reflect on social prejudices and unfair constructs or practices, but also give the possibility to implement co-design strategies to overcome biases. Thus, a synthetic and discrete definition of each facet could not be identified and the preliminary ambition to create a framework to connect these aspects and finally get a comprehensive picture of causal relationships between them could not be fulfilled.
However, some interesting insights could be inferred from this inquiry. Firstly, although the subject of the analysis were guidelines, which one might expect to provide practical directions for the design and development of AI and ML systems, the closer they were to the field of ethics, the more general and high-level were the suggested solutions. Overall, no prescriptive indications could be found in such an uncertain context, and the value of ethical considerations lied in the reflective process they could trigger. Additionally, when not strictly linked to the development process at the model or algorithmic level, the provided indications to prevent, overcome, or limit any undesired impact highly echoed well-established design practices. From the most explicit application of a design-driven approach to properly frame problems and of participatory, systemic, and human-centered principles to interaction or UX solutions (like planning for graceful failure or facilitating user-system dialogue). Most of the outlined possibilities to promote or preserve values would be no novelty for design students. Therefore, while ML-related *values* and *risks* could offer relevant and actionable insights to the educational experience, the same cannot be stated for the collected design possibilities, which instead risk limiting creativity. On the contrary, letting design students autonomously reflect-in-action (Schön, 1983) might be a more powerful driver for learning-by-doing and stimulating a potentially radical process of reinterpretation and innovation (Norman & Verganti, 2014; Antonelli, 2018; Yang, 2020).

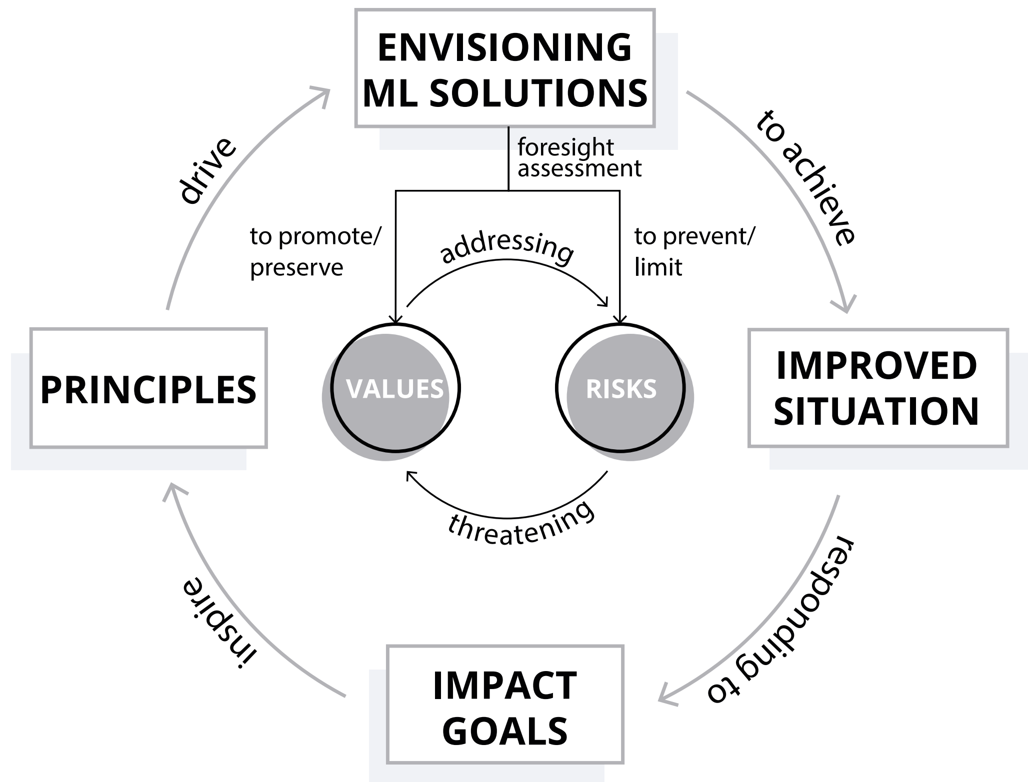


Fig. 4.29 | Responsible Cycle for ML Design

For this reason, instead of a rigid theoretical systematization portraying the relationships between ML limitations, implications, and suggested solutions, a **procedural framework highlighting the interconnections between the foundational elements for a responsible ML design process** was developed (Fig. 4.29). It synthesizes the research findings on ethical concerns, contributing to frame ML knowledge for transfer.

The *Responsible Cycle for ML Design* was structured in a circular form. This underlines its iterative nature and implies that an intuition on any of the four major components could be an entry point to start developing an idea, to be progressively refined by clarifying all of them. Therefore, having the basic design requirement (Simon, 1996) to *improve a situation*, one should define which grand challenges or *impact* might drive the design process (von Schomberg, 2013), and which *principles* are to be pursued (van den Hoven, 2013; Umbrello & van de Poel, 2021). Then, *envisioning a responsible ML solution* implies anticipating and assessing which *risks* need to be prevented or limited by identifying and embedding relevant *values* that must be promoted or preserved.

4.5 Foundational assumptions for the educational method

Summing up the knowledge and experience acquired in the early stages of research, the reflecting phase related to RQ2 represented a milestone. Different facets have been explored to understand ***how to frame ML knowledge for transfer***, from a designerly systematization to preferable forms and languages and the depiction of key ethical concerns. However, the research activities and findings can be synthesized in a few foundational assumptions that steered the development of the educational models.

Specifically, enabling the envisioning of meaningful ML-infused solutions translates into three main requirements. The first one is intrinsic to design education, and it resides in the capability to properly **frame the challenges** to face. Any conceived idea should be in line with **relevant human needs**. But, as it constitutes an essential prerequisite expected from design students' skills, the research does not implement any specific strategy to foster it.

The other two requirements, instead, represent the core subject matter of the doctoral inquiry and entail that for ML-infused solutions to be meaningful, they should at least be (i) **consistent** with the possibilities offered by the technology and (ii) **responsible** with respect to people and their present and future ecosystems. These qualities are respectively tied to knowledge derived from the fields of ML and ethics and delineate a multidisciplinary approach. Even if highly interdependent for a successful result, the theoretical constructs and the tools to introduce these concepts in design education can form two different research strands to be tested separately and addressed as needed.

The *ML Designerly Taxonomy* (section 4.1.2) expressly aims (i) to achieve **technological consistency** by matching *technical capabilities* and *human values*, as suggested by Yang (2020), and extending them to *system processing modalities* to build the premises for designers-ML experts communication. It bridges designerly and technical perspectives on the design of ML systems, which are at the extremes of a multi-level systematization and converge on the core capabilities ML systems currently have. Indeed, **ML tasks** (classification, regression, sequence prediction, generation, clustering, and action selection) **and their basic explanation** (in the form of input-processing-output and accompanied by examples) have been assumed as the **minimal and focal element for introducing ML to design students and enabling them to envision solutions integrating this technology,** as synthesized in the *ML Agents*. No further technical information (learning paradigms, methods, nor algorithms) or practical experiences are provided to introduce ML to design students. Just a few demystifying concepts complementing the explanation (i.e., ML systems as agents, their learning process, complexity, and autonomy) could be enough to foster consistent solutions.

(ii) To help designers conceive **responsible ML solutions**, instead, two essential concepts need to be included in the translation. The first is about reminding that ML systems are special kinds of **sociotechnical systems** (van de Poel, 2020). This implies that, to design them, one should keep in consideration not only the artificial agents and technical artifacts but also the people who decide how ML is developed, those that interact with it and those that are impacted by it. Because of this, a systemic and holistic approach should be encouraged.

The second issue underlines the importance of **intentionally considering and embedding values** in artifacts early in the design process (van den Hoven, 2013). To do so, a *Responsible Cycle for ML Design* (Fig. 4.29) was outlined and its core elements was identified through a systematic analysis of the main AI ethical guidelines currently available and collected in (Algorithmic Watch, 2020). The inquiry suggested a further indication for the development of an educational model for designers. The guidelines mainly portray principles and values to overcome or limit unexpected or undesirable implications of AI and ML systems. These can serve as useful references for designers' foresight and assessment of envisioned applications integrating this technology. However, when available, the possible remedies to prevent or limit negative impacts highly resonate with typical human-centered design approaches. For this reason, while ethics can be a fruitful source for unfolding critical insights about the design of ML-infused artifacts, developing valuable solutions should be designers' responsibility. Thus, the translation should **frame essential ethical concerns and incentivize reflection-in-action, to educate thoughtful professionals.**

This box contains insightful points of reflection from experts' interviews. Following the foundational assumptions previously stated, they revolve around the concepts of (i) ML capabilities as focal points in the translation and (ii) design students' exploration of ethical concerns.

(i) Focusing on what ML can do as a translation strategy for designers is a position shared by all experts but with prospects for improvement. Indeed, according to Jodi Forlizzi, design students *"need a different sense of how machine learning works and more about the capabilities,"* as experienced in the course she and John Zimmerman offer at Carnegie Mellon, *The design of AI products and services,* in which they *"want designers to think more abstractly about the AI and machine learning. So instead of teaching them about algorithms and types of learning, we teach about the space of opportunity and the capabilities."* *What happens with this approach is clearly explained by Peter Krogh, who says that it is necessary to make "somebody unfamiliar with the field to get over that threshold and say: «Oh, yeah, that's what machine learning systems do! »"*

Another position on which all agree is:

DESIGNERS SHOULD NOT REPLACE ML EXPERTS

"I don't think designers need to be made into machine learning specialists" (Forlizzi), *"not necessarily* [they are required] *to get good about it, or to be able to take on the kind of work assignments that software engineers would be able to do"* (Johan Redström).
In fact, as a professional designer, *"as long as you're in a decently sized organization, you're probably not going to be expected to know how to code any of it, but you do need to understand how it functions"* (John Sharp), *"you don't have to understand so much about the mathematics of it"* (Redström).
Put simply, designers *"should not become machine learning experts because we're basically not good at it."* (Krogh). As Zimmerman said, *"going to the coding level is like suddenly I'm doing material science, and designers don't need to do material science to be really good."* And he reinforces this position with another vivid parallel that really captures the reason why programming skills are not essential to designers: *"I don't think most architects work directly with concrete, but they're very effective at using concrete. They don't necessarily have to pour it. But they do need a sense of its capabilities. And I think too often learning machine learning is about learning the mechanisms. You learn how it does something. You don't learn what it can do"* (Zimmerman).
In the end, the important point is *"not that designers need to be able to operate autonomously from engineers, but they need to be equals in the conversation."* (Sharp)

Indeed, within the field of ML, a clear

AREA FOR DESIGNERS' INTERVENTION

can be identified. As design is an applied discipline, it seems natural (and needed) for it to focus on how ML systems could and should be applied in the world. Therefore, the central area of the ***ML Designerly Taxonomy***, where ML capabilities lie, has been generally recognized as a suitable concern for designers and a starting point from which design can actually bring a positive contribution. Even if Forlizzi would focus more on the level of *Applied Capabilities*, getting into a more granular dimension of the concrete possibilities that ML systems offer today, according to Krogh, the epicenter of *ML Tasks* is the *"battlefield"* where designers' expertise can effectively complement current development and deployment of ML systems towards understanding what problems are suitable to be addressed. In his words, it is a *"shared challenge between the machine algorithm designers and classical designers,"* and it would be interesting to see it unfolding while cooperating with ML experts.

To complete the translation of ML capabilities to design students, then, it is important to let them

EMPATHIZE AND DEVELOP A CONSTRUCTIVELY CRITICAL ATTITUDE TOWARD TECHNOLOGISTS' PERSPECTIVES.

What has been highlighted by different interviewees is a further step in the comprehension of what ML can do, which resides in conveying the technical point of view more explicitly. Specifically, as Krogh said, *"we also need to educate designers to unpack and analyze what exactly is said in the technology field."* Without going into the details of programming, design students should be indirectly or directly (e.g., in a collaborative context) exposed to ML experts' mindset and beliefs in order to shape a deeper awareness of the problematics related to introducing ML systems in the world. Through careful listening, designers should mature an autonomous critical perspective that could help them read between the lines of the broad ML discourse and reframe its meanings and the problems it can respond to.

This reinforces the necessity (previously suggested) to demystify current myths that spread not only throughout society but also in the technical field itself.

ML EXPERTS' FALSE NARRATIVE

needs to be debunked. As Krogh put it, engineers and computer scientists *"have a particular way of studying their technology and a particular way of talking about this, which is: the future is always going to be better."* Typically, as Redström pointed out, ML experts are super-efficient at optimization, so once they pursue a direction, they will make incrementally better solutions. Yet, the direction itself is seldom questioned, and the underlying assumption that it is undoubtedly positive might lead to dangerous results. This is why a critical and creative approach that enables the identification of flaws or new possible paths is essential. It should be promoted in design education to avoid students continuing to believe the ML narrative like *"a super amazing human intelligence, while it is not even a toddler,"* as Zimmerman reported about most of those they have encountered.

Additionally, as it was not possible to investigate it within the constraints of the research, another indisputable point of improvement is the introduction of

MORE PRACTICAL EXPERIENCES TO ENHANCE UNDERSTANDING.

A learning-by-doing approach might be accentuated because ML Tasks as the sole expression of the capabilities of this technology may be too abstract, as pointed out by Marianne Graves Petersen. She stated, *"We need some understanding of the dynamics of the technology, and we need some way of exploring it actively, to experience how it works. So, we can't just have the high-level concepts or have them explained. We need some way of getting it, of working with a designerly way of doing the action and reflection and seeing what happens".* Analogously, Redström, talking about ML systems, underlined that *"it really helps to see them work, to try them. So, to your notion of general explanation in a familiar language, I may add practical experimentation. So that you actually get to see something. Usually, you don't get all of it, but you get a sense of what it is."*

Strictly related to the previous points, another way to get closer to a consistent understanding of ML capabilities and to practically debunk false narratives is to emphasize

WHAT ML CAN REASONABLY DO TO GIVE SPACE FOR REAL OPPORTUNITIES

Zimmerman strongly advocates that understanding what ML can do is only one part of the problem. The other is about what it can ***reasonably*** do, which is strongly tied to performance. Even if clickbait stories of exceptional outcomes can occasionally emerge, ML systems usually do not perform very well.
To clarify this, he effectively quoted the metaphor by Cassie Kozyrkov (chief decision scientist at Google), who defined AI and ML systems as an island full of drunk people. They can do a lot of work but of arguable quality, which results in not totally reliable outcomes.
ML experts are perfectly aware of the fact that current models can be almost perfectly accurate where no one needs them, but not where they might be helpful. Yet, as depicted by Krogh, possibly their trust in a better future finally pictures a distorted story. Therefore, to grasp ML capabilities, Zimmerman said that it is essential to lower expectations in order to *"find a ton of things you can do,"* and we need to *"sensitize students to recognize where the performance of a mediocre model is valuable."*

(ii) Analogously, fostering responsibility in the design process has predictably found a consensus among the interviewees, and it raised reflections that emphasized some aspects and traced the path to possible refinements.
The first point is the undiscussed

IMPORTANCE OF GUIDANCE

Whether a responsible orientation of design education is a consolidated practice (like in the case of Parsons' MFA Design and Technology program) or something to explore more thoroughly, the need to devote attention to how design students are led to deal with ethics-related matters is frequently reported.
On the one hand, *"conversations about technology and ethics and larger questions of equity, inclusion, social justice, and the environment are not consistently on students' radars"* (Sharp). So, addressing these issues is a first step toward making them understand how they are relevant to their role and to engage them in such discourse. Otherwise, he added, letting students lose, with no proper guidance, well-selected examples, and opportunities for feedback and critique, some *"are likely to go off and replicate patterns they see out in the world,"* the same patterns that need to be broken.
On the other hand, the intentionality in including ethical issues in design education is a felt requisite. Redström suggested that *"we need to support our students more in the ethics domain,"* possibly with dedicated courses, because it is often implied within teaching or, as Zimmerman noticed, *"design researchers naively assume they are ethicists,"* with inadequate results.

For these very reasons,

DESIGNERS SHOULD NOT WORK ALONE

Cooperation with other professionals can clarify what the actual problems are or might be, both from a theoretical and moral point of view (ethicists) and a practical one (ML experts and providers). These include issues of technical and financial feasibility – deepened in Chapter 6. As Graves Petersen put it, to make quality connections between values and ML systems design, *"There's a big bridge to cross. And if designers are not dressed up for going a little bit out on the wobbly bridge together with the engineers, it becomes detached."* To get to responsible solutions, Zimmerman underlined that designers *"have to work in partnerships with other expertise and not think of this as something designers do in isolation. They are part of a collective voice in defining this."*
Hence, providing design students with values and risks to enhance reflections can be a good starting point, *"but then we need to have it conceptualized as a kind of high-level process that can be discussed"* (Graves Petersen). In Redström's opinion, it implies helping designers *"structure thinking in a way that allows them to communicate it."* He added that if they work in interdisciplinary teams, it is not enough for them to

IN EXPERTS' WORDS

propose solutions, *"They also need to explain why it should be in this or that way or why this is better than the other."*

GUIDELINES VS EMPATHY

For the purpose of educating design students to think and talk about ethical implications, as assumed in this research, they do not need to be told what they can or cannot do (Redström), meaning that guidelines are often not useful and too high-level, not allowing designers to ask the right questions or challenge the system (Graves Petersen).
Instead, leveraging on the fact that *"designers are empathy experts"* (Zimmerman) is helpful in this process, and it leads back to the *"history we have in dealing with sort of ethics in the applied way of seeing the ethical issues in a one-to-one interaction"* (Redström).

Of course, there are

MULTIPLE FACTORS TO KEEP IN MIND

and defining ML systems as sociotechnical systems was recognized as relevant in this context. According to Redström, it helps to understand how complicated these systems can be as they extend to a network of artifacts, people, decisions, and relationships. This complex configuration also implies that *"much of what's ethically problematic about them is unintended. So, there are a lot of things going on that we didn't mean to happen. They just happened because things are interacting in ways that we couldn't foresee"* (Redström). Indeed, to simplify this concept and to make it more operationalizable, Forlizzi broke it down into *"sociotechnical systems at build time and then sociotechnical systems at runtime."* She proposed to clearly distinguish between these two moments to convey different kinds of considerations. In particular, during the design process, reflections on technology and people may focus on their mutual interactions and the feedback loops. Still, also the problems that can happen when the systems are in execution need to be thought of. Therefore, unpacking the concept of sociotechnical systems might help to have it addressed in all its different nuances.
As depicted in the definition of the basic assumptions of the research, another relevant factor for the suggested reflective foresight activity is that it takes place

EARLY IN THE PROCESS TO SUPPORT PROBLEM SOLUTION SELECTION.

As both Zimmerman and Redström pointed out, ethical considerations are usually addressed only once what to make has been defined instead of being drivers for deciding what should be pursued among the countless possibilities that one might come up with. If reflections on risks and values are encouraged soon in the process, when preliminary decisions are made, it is often still possible to deliver value without incurring predictable risks. Indeed, there are so many other things that can be done to respond to a problem, and *"the process of problem solution selection is where designers can have a ton of impact because we are also the world's experts on how to ideate and broadly explore a solution space"* (Zimmerman).
Additionally, Graves Petersen underlined that if designers were more aware of all the design choices that characterize the development of ML systems at a conceptual level, they could also have a big role in *"some of these still technical but high-level decisions that have a big impact on how* [they are] *actually being responsibly done [...] how they are being used, how they influence people's lives"*, and so on.

Overall, the effort of anticipating risks and embedding values into a project can be synthesized in

ML STORYTELLING

as Krogh brilliantly illustrated in a vivid argumentation. Summing up his major points, he suggested that *"part of getting designers closer to AI is saying "how do you tell the story?". Indeed, before providing solutions, "we may, as designers, have an intention, an idea, something that we would like to signify,"* and *"there's something about the type of stories that we, as designers, signify which could help us articulate the values that are in the systems,"* but we cannot control how it is understood. The least we can do is be aware of the story we are telling through our decisions and be conscious architects of it, echoing (Thaler & Sunstein, 2009). As Krogh highlighted, much of developing ML systems is not about providing solutions but telling stories, like defining selection criteria, classifying, filtering things out, and having algorithms reflecting biases. So, a constant question we could ask ourselves during the design process is, *"What kind of story would emerge at the other end?"* (Krogh).
He envisioned different ways for designers to be ML storytellers. One is about exaggerating the stories that could arise from risks and biases. Another is about taking the distances *"from the idea that designers do good,"* accepting it and declaring it explicitly by acknowledging that *"This was the best possible solution we came up with. We know there are downsides. It's not perfect. Sorry, guys."* And finally, *"fairy tales are also stories that help us navigate life because they tell the general story or a very emotionally strong story that enables us to behave in a particular way in society afterward. They should be moralizing, they're emotionally attaching, they're good advice,"* and they can be used as a reference to embed and talk about values in ML systems. However one might decide to build them, in the end, *"these stories are all helping us to grasp something,"* including how to be responsible in the design of ML systems, and they should be fostered during the design process.

TO SUM UP

The chapter addresses RQ2: ***How to frame ML knowledge for transfer?*** For this, the investigation builds a bridge across the disciplines involved.

- A ML Designerly Taxonomy is proposed based on the insights gained in the previous stages of the research. It is a synthetic theoretical construction that bridges ML knowledge and its application in the real world, connecting human-centered and technical perspectives, thus, designers and ML programmers. The link between *human values* and *system processing modalities* is constructed in 5 different levels (Conceptual Knowledge, to understand the potentialities of ML systems in relation to human capabilities; Designerly Knowledge, to identify the concrete opportunities designers have to exploit ML capabilities; Operational Knowledge, related to current applications of ML systems; Technical Knowledge, including the technical principles and processes underneath; and Operative Knowledge, implying the understanding or mastery of ML systems' functioning, which can extend to the fine-tuning of ML systems' models). In turn, they take the form of dimensions (Intents, Design actions, Applied capabilities, ML tasks, Learning paradigms, and Learning methods) that are articulated into features, the main elements with which designers and ML developers can identify ML systems.
- To assess (i) the preferable forms and languages for framing ML knowledge for design students, (ii) the effectiveness of the *Decoder* and *Encoder* tools (section 3.3.1) to make sense of and communicate a ML-infused solution, and (iii) the appropriateness of the taxonomy categorization of ML tasks to show the possibilities and stimulate the envisioning of ML-based solutions, the *ML Pills for Designers* workshop was organized and produced promising outcomes. It could be observed that:
- The presented contents seemed appropriate.
    - The synthesis of ML systems as agents (described by input, task, and output) was naturally apprehended and proved useful for the participants to present their ideas concisely.
    - ML capabilities depicted by ML tasks enabled students to envision new solutions by intuitively applying them in their design process.
    - Among the four formats individually proposed to communicate ML tasks, the primary importance of examples (both in the form of *Practical Examples* and *Case Studies*) was reinforced; *Metaphors* worked well as abstractions to facilitate designers' comprehension and exploration of ML; while *Definitions* were the most challenging to grasp because they lacked visual or referential components, and their language was not designer-friendly enough.
    - A combination of different formats of communication might be the most successful solution.
    - Generally, the participants addressed the design briefs with a human-centered approach, but the outputs were mostly individualistic solutions that could easily fail in social contexts. Thus, even if premature in the experimental context of the workshop, explicit ethical reflections would prove beneficial.
- *ML Agents* have been developed to transfer foundational ML knowledge based on the workshop results. Clarifying their computer-based nature, they identify the main capabilities of ML depicted in the ML Tasks dimension of the *ML Designerly Taxonomy* (classification, regression, sequence prediction, generation, clustering, and action selection) as agents (with an input-processing-output structure). To explain them, they combine more simplified and illustrated definitions, with case studies as explorable examples.
- Looking for ethical support as a missing element for envisioning meaningful ML systems, the research identified principles from responsible research and innovation (RRI) and value-sensitive design as relevant references. They include the definition of "right impacts" to pursue, foresight, assessment, and the intentional *integration of values as key moments in a responsible design process.*

- To facilitate design students' concept development of ML-infused solutions, the inquiry aimed at identifying ML-related values, risks, and possibilities to overcome or limit them through a systematic content analysis of 61 ethical guidelines for AI. The results include the following:
    - A list of five principles (*promotion of flourishing, prevention of harm, attention to fairness, increase of intelligibility, and respect of human autonomy*) and fifty values was compiled. Of these, ten have been identified as the most representative and comprehensive (*accessibility, agency, explainability, freedom, justice/non-discrimination, reliability, representativeness, sustainability, (meaningful) transparency, and well-being*)
    - 47 *risks* and 67 *possibilities* were identified. Though, it was not feasible to construct a theoretical systematization that clearly highlights how they relate to each other in combination with specific *values* because they are so much interrelated that they can easily overlap.
    - A procedural framework highlighting the interconnections between the foundational elements for a responsible ML design process was instead developed to synthesize the findings on ethical concerns contributing to frame ML knowledge for transfer. The *Responsible Cycle for ML Design* shows that to achieve an *improved situation*. One should define an *impact goal* to be reached through *principles* that lead to the *envisioning of a responsible solution*. This implies anticipating and assessing which *risks* need to be prevented or limited by identifying and embedding relevant *values* that must be promoted or preserved.
- At this stage, the foundational assumptions of the research were inferred. Specifically, for ML-infused solutions to be meaningful, they should at least be (i) **consistent** with the possibilities offered by the technology and (ii) **responsible** with respect to people and their present and future ecosystems.
    - The achievement of technological consistency can be favored by the *ML Designerly Taxonomy*, and, in particular, ML tasks (as depicted by *ML Agents*) have been assumed as the minimal and focal element for introducing ML to design students and enabling them to envision solutions integrating this technology.
    - To help designers conceive responsible ML solutions, the translation should clarify that ML systems are a **special kind of sociotechnical systems**, the design process should intentionally consider and early embed values (as in the *Responsible Cycle for ML Design*), and just ethical concerns should be provided to incentivize students' reflection-in-action (indeed, the general solutions provided by ethical guidelines are not relevant nor instructive).

5. CONCRETIZING HYPOTHESES IN MODELS AND TOOLS TO SUPPORT DESIGN EDUCATION

By understanding and engaging with new technology in a designerly way, design researchers can envision new forms and new purposes for the technology through the creation of sensitizing concepts, and help initiate a wave of design innovations.

(Yang et al., 2018)

C H A P T E R F I V E

The chapter focuses on ***how to operationalize theoretical constructs into models and tools to be implemented and tested in educational contexts?*** (RQ3), identifying the key features of ML (for) design didactic experiences.

An initial experimentation aimed to merge the three disciplinary perspectives investigated (ML, design, and ethics) in a single educational activity to assess how design students might respond to a multi-layered ML design knowledge (5.1). An *Introductory Game to ML Responsible Design* was developed and submitted to four Ph.D. students in design from Politecnico di Milano with no or little knowledge of the topic to gather insightful feedback. While the approach was smoothly assimilated, useful indications for developing tools to materialize the knowledge transfer emerged. **Flexibility, modularity, learners' agency, and freedom to explore the contents** were then implemented in the construction of the *Concept Building Blocks* tool to support the envisioning of a ML-infused system consistent with the technological capabilities and the *VALUable by Design Expansion*, adding a responsible orientation (5.2).

The versatility of the educational approach and an adapted version of the *Concept Building Blocks* were tested in the workshop *Superpowered Museums* (5.3). Held in a challenging context with interior design students not used to working with technology, it needed some content adaptations but proved effective in its scope to enable them to envision ML-enhanced spatial experiences.

Summing up, useful insights were drawn from these experimentations (5.4) to build an educational method that could respond to various educational needs and contexts, as described in the next chapter.

5

5.1 A holistic approach

In the action research spirit that characterizes the investigation, the theoretical elaborations produced in the previous two cycles (in response to RQ1 and RQ2) required a practical translation for their assessment. A major issue to test before moving on with the development of the educational models was the suitability, for design students, of a concise holistic educational activity encompassing the foundational concepts identified for each of the disciplines involved: ML, ethics, and design. A synthetic sheet (pages 158 - 159) and a detailed description follow.

5.1.1 Method. Synthesizing the different disciplinary perspectives.

5.1.1.1 General framing and ILOs

The *Introductory Game to ML Responsible Design* was developed to assess how design students might respond to being exposed to a **multi-layered ML design knowledge**, with no ambition to be a comprehensive tool. To the best of the author's knowledge, this was the first empirical attempt to merge the three disciplinary perspectives on ML systems development.
Due to its experimental nature, the primary purpose of the activity was to gather useful feedback and insights on the effectiveness of the approach, which was achieved through observation and a final semi-structured interview for in-depth discussion with the participants. Thus, it targeted four Ph.D. students in design from Politecnico di Milano. Their closeness to the intended audience for the research, analytical skills, didactic experience, and consciousness about design theories and methods made them relevant resources for insightful discussions and peer-evaluation of the activity. Additionally, the limited number of participants ensured the researcher an in-depth look at their experience of the activity and reflections in real time. Her intervention, though, was limited to the situations in which it was expressly requested for clarification.
The experimentation articulated as a physical- or digital-based engaging experience, tested in two sessions of about one to two hours. Two students with no prior knowledge of the subject matter attended the in-presence one (Fig. 5.1). The second involved Ph.D. students with research areas more or less related to ML as a technology. A Miro board (Fig. 5.2) supported this session to simulate real-world collaboration best.

Introductory Game to ML Responsible Design

Contextual information

WHAT	1 to 2h playful experience (cooperative board game).
WHEN	14 and 16 July 2021.
WHERE	Politecnico di Milano, Design Department and online.
WHO	The author.
STUDENTS INVOLVED	4 Ph.D. students in design from Politecnico di Milano (divided in pairs). Voluntary participation.

Research Rationale

RQ UNDER INVESTIGATION	***RQ3: How to operationalize theoretical constructs into models and tools to be implemented and tested in educational contexts?***
RESEARCH OBJECTIVE(S)	(i) Empirically attempt to merge three disciplinary perspectives (AI, design, ethics) to envision ML systems. (ii) Assess how design students might respond to being exposed to a multi-layered ML design knowledge.
TARGET AUDIENCE RELEVANCE	Ph.D. students in design are close to the intended audience for the research and have more developed analytical skills, didactic experience, and consciousness about design theories and methods, essential for insightful discussions and a peer-evaluation of the activity.

Methodological framing

EXPLORATION STRATEGY	The experimentation consisted of **two sessions** (one physical and one digital-based to test both forms and iterate the activity). Two students with no prior knowledge of the subject matter attended the in-presence one. The other two, with research areas more or less related to ML as a technology, participated online with the support of a Miro board. The limited number of participants per session allowed the researcher to **unobtrusively observe** the players interacting with the educational materials and discussing between themselves. Further insights were then collected **through in-depth semi-structured interviews** with the Ph.D. fellows.
DATA COLLECTION	• Observation • Semi-structured interviews after the experience
RESEARCHER'S ROLE	Unobtrusive observer (expressly assisting the activity to collect insights and requested to intervene for clarifications only when unavoidably necessary).

Structure of the educational activity

ILOs	***Knowledge***	• Familiarize with **basic ML knowledge.** • Get in touch with possible **ML-related risks and values**.
	Skills	• Pursue a value-driven approach while handling ML as a design material. • Set fruitful discussions in team collaboration.
	Values	• Understand how to be responsible designers in the contemporary situation. • Identify ML as an asset to face big challenges.
EXPECTED IMPACT(S)		• Trigger **reflections and discussions about building awareness on ML** from an interdisciplinary perspective. • Pave the way towards **introducing an explicit value-driven approach** to envisioning ML-infused solutions in design education.
CONTENTS		• Value-sensitive design principles • ML capabilities • ML-related risks • ML-related values • Requirements for responsible research innovation (RRI)
TOOLS	***Knowledge transfer & design activities***	Board game materials
OUTPUT		Concept of a ML-enabled solution to address a Sustainable Development Goal challenge.

Findings

KEY INSIGHTS	(i) A simulated design process proved successful to integrate and transfer knowledge about ML and ethics. (ii) The holistic approach was effortlessly assimilated by the testers, who comfortably navigated across disciplines, confirming that design educational background is functional to deal with complex systemic issues.
ISSUES FOR FUTURE INVESTIGATION	• Enough agency and freedom are essential for learners to explore the contents, form their own idea on the subject matter, and discuss about it to make decisions. • Tools should be precise, straightforward, and need further simplification. • They should engage, support reflection, facilitate decision-making, and expand design thinking without prescribing design actions. • Physical materials are more effective boundary objects to trigger discussion.

Fig. 5.1 | In-presence playtest.

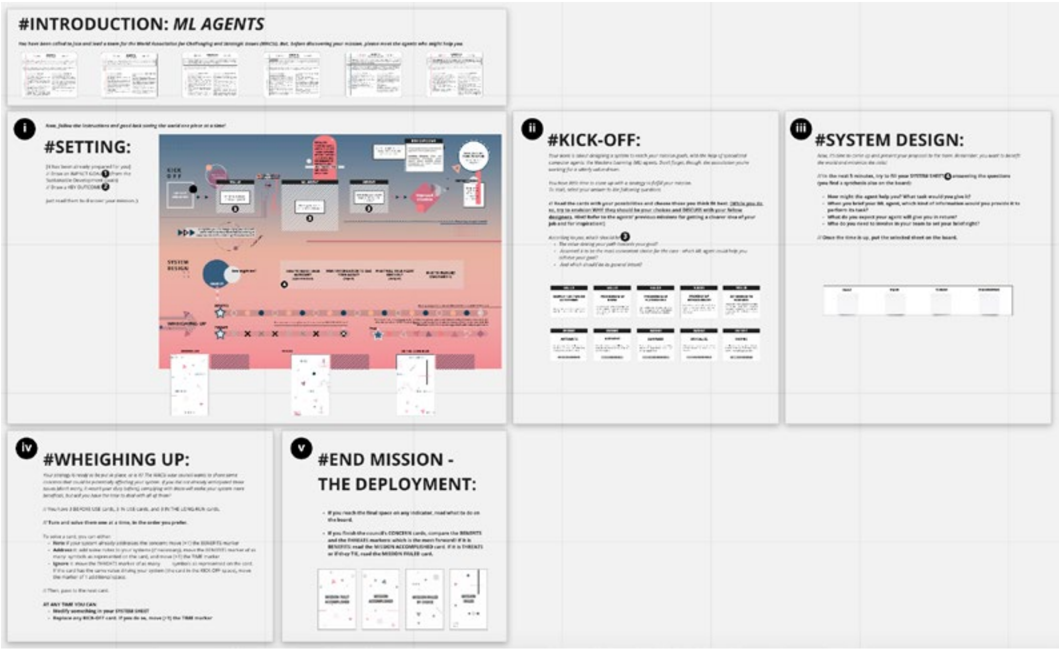


Fig. 5.2 | Overview of the digital Miro board supporting the online activity.

As learning outcomes, the design student representatives were intended to be able to **(i) familiarize** themselves **with basic ML knowledge**, including related risks and values, **(ii) pursue a value-driven collaborative approach** while handling ML as a design material, and **(iii) understand how to be responsible designers** in the contemporary situation and exploit ML as an asset to face big challenges. The overall expected impact was to trigger reflections and discussions about building awareness on the ML topic from an interdisciplinary perspective and paving the way towards introducing an explicit value-driven approach to envisioning ML-infused solutions in design education.

5.1.1.2 Outlining the cooperative board game

The didactic activity was imagined as a possible introduction to a more extensive educational path. It was developed as a cooperative board game to simulate teamwork in design environments. In this way, players could gently learn new content, follow procedural information, freely express themselves, and make mistakes in a practical but safe context (Huizinga, 1938).
Once again, the ideation phase of the design process frames the educational experience. Indeed, considering how the different disciplines could contribute to each of the five stages of the design thinking process (Fig. 5.3) – as illustrated by Dam and Siang (2021) – ideation would be the most fruitful.

	EMPATHISE	DEFINE	IDEATE	PROTOTYPE	TEST
DESIGN	*researching space for intervention*	*framing the problem*	*envisioning possible solutions*	*setting up tests and anticipa-ting outcomes*	*monitoring and assessing impacts*
ETHICS	*influencing the research*	*driving towards a meaningful purpose*	*fostering acceptability, sustainability and desirability*	*orienting choices*	*defining parameters and trade-offs*
ML	/	/	*opening up new possibilities*	*setting technical constraints*	*supporting the evaluation system*

Fig. 5.3 | Author's analysis of disciplinary contributions at each stage of the design thinking process.

Then, it sets the background of the whole game, divided into four sections.

Phase 1: SET-UP. At the service of a fictional *World Association for Challenging and Strategic Issues* (WACSI), the players were designers called to accomplish a mission. It is defined by a long-term *impact goal* (identified by a specific target of one of the UN's SDGs) and a *key outcome*, narrowing down the problem as a near-term and observable change or behavior to promote and including information about the context and target audience. According to their specific goal, the players had to responsibly design a system enhanced by ML and aimed to improve the current situation.

Phase 2: KICK-OFF. According to value-sensitive design suggestions (section 4.4.1), a reflection about what values to embed consciously in a system needs to happen early in the process. For this, the *Principle* cards (Fig. 5.4) were primarily provided to select the one that most prominently should drive the design process. These coincide with the principles identified in the content analysis of AI ethics guidelines described in 4.4.3.

PRINCIPLE	PRINCIPLE	PRINCIPLE	PRINCIPLE	PRINCIPLE
PROMOTION OF FLOURISHING	**PREVENTION OF HARM**	**RESPECT FOR HUMAN AUTONOMY**	**INCREASE OF INTELLIGIBILITY**	**ATTENTION TO FAIRNESS**
Life, well-being, growth, progress and prosperity of ecosystems should be fostered and nurtured	No material or mental damage has to be inflicted to people and their ecosystems, nor existing ones have to be worsen	People must be free to make their own decisions and to take control	Immediacy and understandability have to be guaranteed, whether with a proper explanation or intuitively	People and their ecosystems need to receive just and impartial treatment, respecting a balanced proportionality between means and ends

Fig. 5.4 | Principle cards provided in the game.

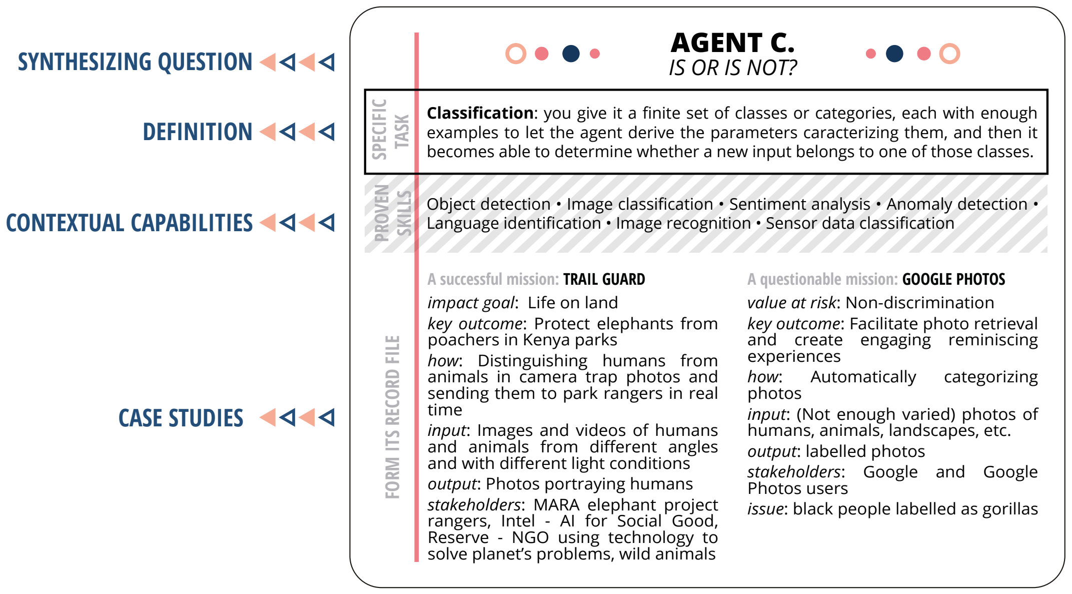


Fig. 5.5 | Example of a ML Agents' sheet.

Subsequently, the focus shifted to the technical side, with the core capabilities of ML systems presented in a synthetic version of *ML Agents* (4.3). Here the visualizations were not included because of the limited space (Fig. 5.5), leaving room for two different case studies that strengthen the ethical value that ML applications may have. In particular, one successfully tackled a worthy challenge, while the other described a questionable case to set good and bad examples for each ML capability.
Finally, the designerly perspective entered the equation. The players were supplied with four general *Intents* to beneficially impact human life (Fig. 5.6), as portrayed in the *Conceptual Level* of the *ML Designerly Taxonomy* (4.1.2). One should be chosen to respond to the question: *"How would your system improve the current situation?"*
Hence, in this earliest phase of the design process, elements from all three disciplines were represented in the form of *Principles*, *ML Agents,* and *Intents* to steer the concept development toward a meaningful solution. Explicitly discussing and defining these foundational elements, which could be freely modified in further iterations, should also help maintain the team aligned on the same premises.
Of course, even in this case, the idea construction was constrained and limited to integrating one ML agent, which does not indicate that ML is a certain solution nor that complex tasks can be tackled with a single ML capability. However, the didactic purposes and the short, light, and playful format required this simplification to favor immediacy.

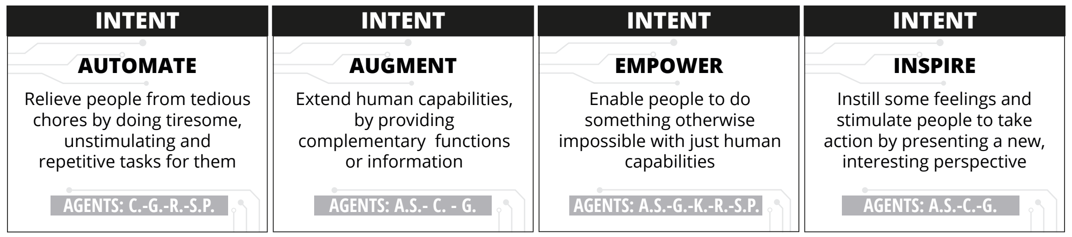


Fig. 5.6 | Intent cards provided in the game.

Phase 3: SYSTEM DESIGN. Given the overarching goal and having determined the essential elements to kick off the design process, the backbone and the boundaries of the system to design should be clear. Thus, to outline the structure of a concept, a *System Sheet* had to be compiled specifying the ultimate synthesis of *ML Agents* (task, inputs, outputs), and the stakeholders that need to be involved, a consolidated practice in service and participatory design. Assuming that previous discussions would also imply an early conception of an idea, a time limit of five minutes was provided for this activity to avoid overthinking. Ultimately, to prevent possible difficulties in systematizing the idea, examples were prepared for the players to compare or adopt once the time was up. After the comparison, final adjustments could be made before the closing phase.

Phase 4: WEIGHING UP. In the last active part of the game, the players had to ponder on their idea and try to refine it in a foresight process. To do so, they were required to address the concerns raised by the fictional WACSI wise council. *Concern Cards* (Fig. 5.7) represented these and included possible risks related to ML systems, values, and options designers might have to avoid such hindrances. The cards also highlighted (with different colors and text) to which discipline the concern most prominently belonged, and they were an elaboration of the results of the content analysis of the AI ethical guidelines (4.4.3).Another factor to consider in this phase was time. Despite being important to anticipate and prevent risks, not every concern could be addressed before deploying a system as this would represent endless work. To simulate a design process in which foresight and assessment were integrated, the *Concern Cards* were divided into three decks according to when the concern might arise: *before use*, while

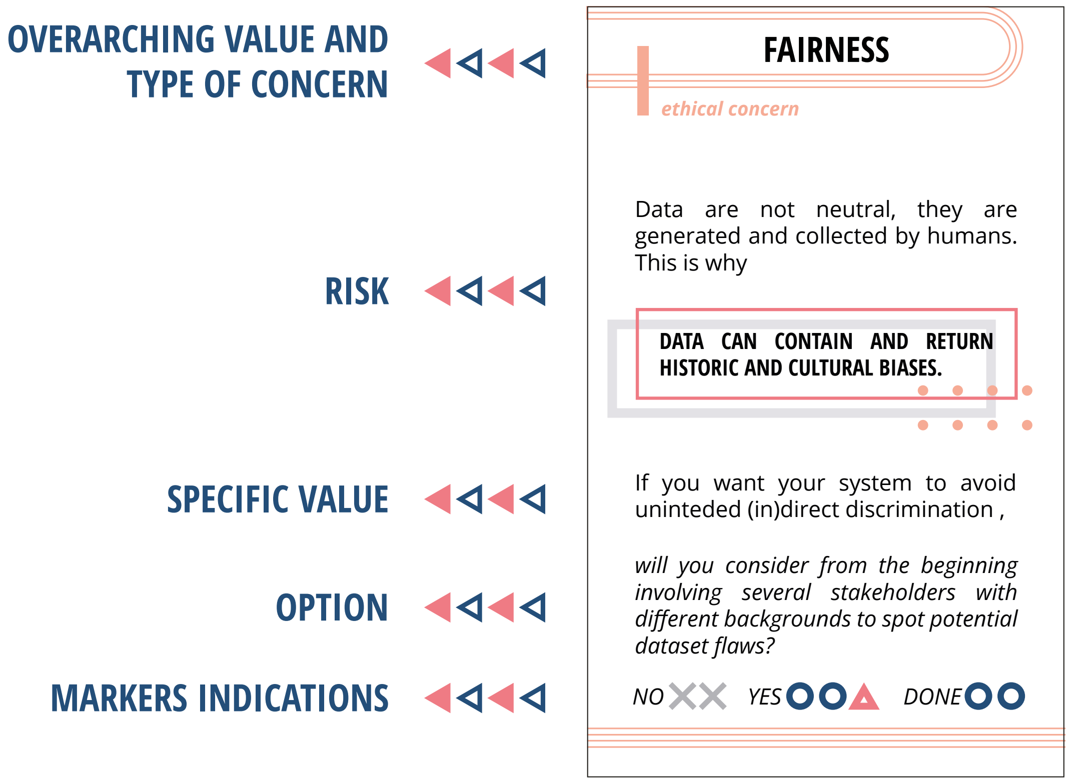


Fig. 5.7 | Representation and description of a concern card.

the model is being developed; *in use*, if it emerges once deployed; and in the *long run*, if it develops over time.
Three random cards were provided for each deck. Addressing them (in any order) through discussion and identifying possible solutions was of uttermost importance to increase the beneficial impacts of the system. However, because of time constraints, the players needed to decide which concerns to address and which to disregard. To accomplish the former, they could add notes to their system sheet or replace any of the *Value, ML Agent*, or *Intent Cards* (optional, as it would cost additional time). This caused *Time* and *Benefits* indicators to advance. Ignoring a concern could happen if it was deemed irrelevant or not preferable to comply with, which could expose the system as a potential *threat*, making only the related indicator move. Of course, the time indicator did not have enough space to fulfill all the concerns cards.
A responsible and thoughtful approach was awarded in this attempt to balance the good or harm that the envisioned system might cause. If the idea already responded to a concern card, with no further action, only the *Benefits* marker would move.
Instructions about the movements of the markers were provided on the cards and varied according to their potential impact.

MISSION END. End-game conditions included solving all the *Concern Cards* or reaching the final space of any indicator (using all the available time or attaining the maximum benefits or threats). The level of positive or harmful impact achieved by the designed system determined how the mission was completed. Four different epilogues were available (Fig. 5.8), though they all converged in presenting the fundamental requirements for responsible innovation that students could guard for future reference: being (ethically) acceptable, sustainable, and socially desirable (von Schomberg, 2013).

Ultimately, the game left open space for further considerations. In fact, whether a system satisfied these requisites – regardless of the mission's outcome – was up to the players to understand and, hopefully, discuss.

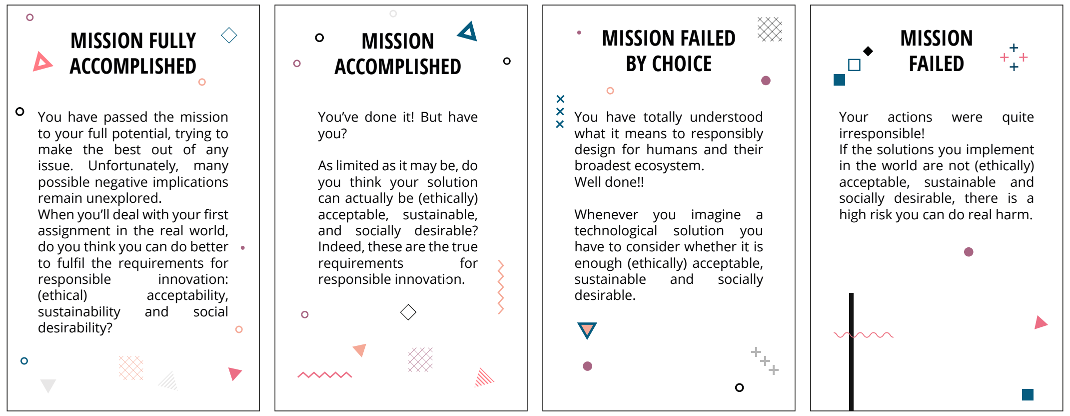


Fig. 5.8 | Available Mission End cards.

5.1.2 *Results of an introductory game to ML responsible design*

Focusing on understanding how the activity affects the didactic purpose, the playtests' results are presented here.
Although the first in-presence session inspired few changes for the second (mainly to the visual organization of the boards and the modalities to provide some game materials), the structure of the tests was the same. For both experimentations, the researcher predetermined the *impact goal* and *key outcome* to facilitate the kick-off and reduce time. SDG #16 – *Peace, justice, and strong institutions* – was selected as a long-term goal, while the more specific yet general brief was to prevent deadly attacks orchestrated by organized crime in public spaces. Examples of datasets as possible inputs for the system (e.g., video and wiretapping recordings, police reports, expert profiling models, propaganda on social networks, direct testimonies) and backup sample system sheets complemented the setting.
Both the concepts responding to the goal consistently exploited ML capabilities and included ethical reflections from the beginning. An overview of the *System Sheets* is illustrated in Fig. 5.9.

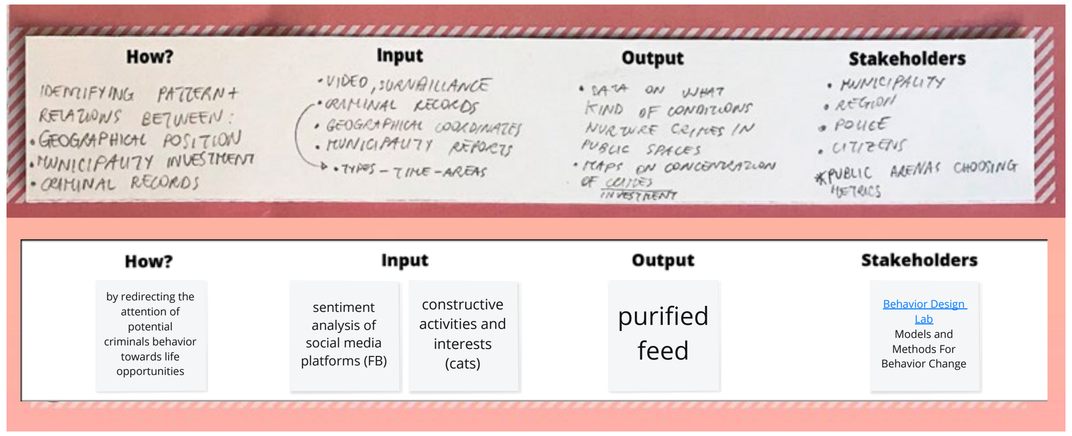


Fig. 5.9 | Outlined concepts from the in-presence (above) and online (below) playtests.

During the first playtest, a detector of the conditions that nurture crime (given context and municipality investments) and including different stakeholders for a plural perspective was identified as a preventive solution to avoid criminal attacks. In the second, a recommendation system was proposed to deter criminals from malicious activities and drive them towards more constructive ones.

5.1.2.1 Limits of the translation. Identifying the criticalities of the experience

The experimentations have been very informative in spotting the weaknesses of the instructional framework.
A first critical point was to properly convey the value of ML in facing relevant challenges for society. Selecting a Sustainable Development Goal as the ultimate intended impact is not sufficient. As it was not explicitly expressed, in the first session, the high level of the system ambition left the players neutral, as if it was a problem like any other, and the same happened in the second playtest despite it being remarked. This might result from a natural predisposition of designers to deal with any kind of

challenge or might have been influenced by the fact that, as Ph.D. students in the Department of Design at Politecnico di Milano, the participants were already used to dealing with such high goals. In any case, *ML Agents* were comfortably managed to address the mission, regardless of the scale of the problem.

The way they were presented, instead, highlighted some difficulties. As emerged in the *ML Pills for Designers* workshop, the combination of *Definitions* and *Case Studies* proved effective, and the level was appropriate. However, the first impact with the technical presentation was a little disorienting for the players. Even though the *ML Agents sheets* contained essential information, the players had to process a lot of content at the same time, which required time and effort.

Two modalities were tested to reduce the impact of the cognitive load. In the first session the participants had to go through them during the *kick-off* phase, when they had to select one, dramatically breaking the activity's rhythm. In the other, the players were suggested to read the sheets before starting the game. In both cases, though, it remained a burdensome task. Moreover, as counterevidence, visualizations (left out of the provided tools) could have enhanced assimilation and memory, as one of the participants suggested.

A possibility could have been to organize a collective introduction, aided by visual supports, to facilitate the knowledge transfer, using *ML Agents sheets* just as reminders.

Another limit regarded the tone of the *Concern Cards*, which language was directly derived from the analyzed guidelines. To facilitate their fruition, they needed simplification and, possibly, a more explicit structure of the contents.

Overall, the game experience was balanced with the capabilities and background knowledge of the testers. The two-pages rules booklet predominantly guided the players, while the contents on the board were quite ignored.

Some uncertainty emerged only in the *system design* phase of the second playtest, and it required the researcher's intervention as a facilitator. In this regard, despite trying to change the organization of the contents on the board to limit external intervention and augment the visibility of important definitions for the participants to play, expert assistance was always sought.

5.1.2.2 Challenges. Unexpected outcomes and space for improvement

As acknowledged in the previous workshop experimentation, the simplification of identifying only one ML capability to address a given problem does not reflect a complete view of reality. For sure, it is helpful to limit the time and complexity of introductory education. Still, these systems are rarely linear and often necessitate integrating multiple capabilities to achieve seemingly simple goals. Concerning this, some issues emerged during the second session's *system design* phase. Aiming to build a recommendation system to orient potential criminals towards positive activities instead of perpetrating in noxious environments, the players identified sequence prediction (processing historical information) as a means to detect which contents to provide. As shown in Fig.5.10, they wanted to feed the system with twofold sequential data: sentiment analysis of posts on social media and videos or ads of constructive activities capable of instilling positive addition towards constructive subjects. Though, sentiment analysis (identifying hate or violence in written posts to understand

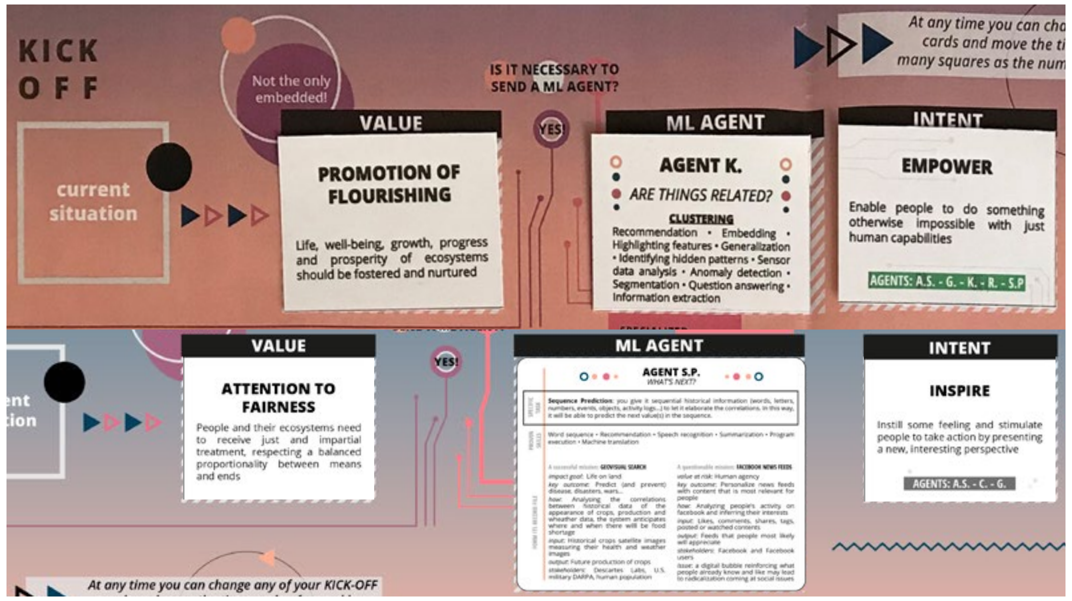


Fig. 5.10 | Selected elements for the kick-off phase in the in-presence (above) and online (below) playtests.

who the system should target) was the product of a classification system, and the overlapping of multiple ML capabilities created a short circuit. Here, the knowledge transfer needs to find a good balance in eliciting solutions that are true to reality or immediately understandable.

Another trade-off concerned the framing of the brief that the players had to address. The choice was challenging because the recipients would have no previous expertise in dealing with the design of ML systems. On the one hand, identifying a *key outcome* that expressly implied a ML solution could have been too restrictive and risked encouraging a technology-driven approach. On the other, generating a very broad brief, leaving space for several possibilities, could lead to non-ML solutions. In the end, considerable freedom was left with the predetermined *key outcome,* but some facilitating materials were provided in case difficulties emerged. They included possible input datasets and sample *system sheets*. Surprisingly, during both experiences, the participants had no trouble delivering a system idea quickly without paying attention to any of the helping materials. On the contrary, in the second playtest, the sample sheets were not at all included, and it caused no problem.

Once again, the case studies from the *ML Agents* sheets played a central role in properly understanding the requirements and as inspiration. Though, whether the systemic vision of Ph.D. students might have affected this successful accomplishment has to be understood.

Eventually, some challenges unfolded during the *weighing-up* phase. None of the participants considered making changes to the *Value*, *ML Agent*, or *Intent* initially selected, diminishing the iterative nature of the process. In contrast, complying with a *Concern Card* was very easy and defined as prescriptive during the first session, while ignoring it caused negative feelings. As this could compromise the goal of fostering thoughtful reflections, the *Concern Cards* could be left more open to stimulate students to further explore and modify their projects.

5.1.2.3 Strengths. Where the translation succeeds

The best result of the educational activity has been the unsolicited exclamation *"I learned something today!"* pronounced by one of the Ph.D. students at the end of the game. In fact, despite having been exposed to a substantial cognitive load, all the participants proved and confirmed that they were able to finely assimilate the provided knowledge across the three disciplinary areas. They seemed at ease with the given tasks and tools, and these effectively provided the essential equipment to envision a ML-infused solution.
Moreover, the synthesis of *Values* and *Intents* projected toward beneficial outcomes encountered a very positive response, and they were at the basis of the systems' foundations (Fig. 5.10). In fact, both groups were very proactive in addressing the problem to prevent criminal attacks. The testers of the first session selected *promotion of flourishing* as a driving value to help people appropriate their territory and avoid crimes and discarded *prevention of harm* as they felt it was limiting. The others chose *inspire* as *Intent* of the system to encourage potential criminals to pursue constructive interests, addressing the very root of the problem.
Aiming at encouraging discussion, the combination of *Values*, *ML Agents,* and *Intents* was very effective as, already in the *kick-off* phase, several ideas were brainstormed. The *system design* process flowed smoothly, and the players needed no assistance for defining even the constituting elements, being able to produce coherent and original results.
Additionally, from the very beginning, ethical concerns were explicitly discussed. For example, the players who proposed the positively addicting recommendation system to deter criminals from malicious activities were perfectly aware of their system's insidious and manipulative nature. However, they considered this solution more ethically acceptable compared to capillary and undifferentiated surveillance. What emerged is that, even without a formal ethics education, the Ph.D. students manifested a responsible approach to design, possibly because they already had a value-driven mindset. Being early-stage researchers, their sensitivity may be increased; still, these kinds of considerations are at the basis of human-centered reasoning.

5.1.3 Discussion. Freedom and agency as fundamental principles

The testers effortlessly assimilated the multidisciplinary approach, and design as educational background proved to be functional to this end. If the previous practical experimentation showed that Digital and Interaction design students finely dealt with ML-related knowledge, this one expanded the boundaries of the participants' educational backgrounds to less technology-related ones. More importantly, it demonstrated how design Ph.D. students could comfortably navigate across disciplines. Although the study has limits in terms of representativeness, it confirmed Findeli's assumptions on Ethics, Aesthetics, and Design (1994). In his opinion, **ethical deliberations do not differ from any other decision-making process**, making the empathic designer already equipped to tackle such issues. He further reinforces the hypothesis that designers can take on multidisciplinary challenges by indicating intuition as the basis for a systemic apprehension of a complex reality, a skill rooted in design education and practice.
Another crucial point that emerged from the educational experiences, in both the physical and digital environments, was the high potential of the materials and the assignments to effectively foster communication. In fact, the initial selection of the foundational elements to address the problem, the definition of the system characteristics, and the final weighing are all collaborative activities. The *Mission End cards* also encouraged a retrospective discussion on the developed idea and the entire experience. The exchange of personal perspectives fueled by ethical concerns and ML comprehension is of uttermost importance in the light of a meaningful humanity-centered design as they convey critical thinking, which can lead to thorough foresight and assessment activities and, eventually, radical innovation.
For this reason, a valuable indication for the operationalization of the theoretical research assumptions (including the *ML Design Taxonomy* and the *Responsible Cycle for ML Design*) emerged from this educational activity. Learners must have enough agency and freedom to explore the contents, form their own idea on the subject matter, and discuss about it to make decisions. In this respect, the *Introductory Game to ML Responsible Design* has wide room for improvement. The attention to support design students in the complex task at hand turned out to be too prescriptive, especially in the *weighing* part, where, instead, critical reflections to uncover new interpretations of what might be significant for people are sought. If a **practical approach to deal with ML knowledge is undoubtedly suitable for an applied discipline like design**, a good **compromise between guidance and unconstrained active involvement of the recipients needs to be reached.**

5.2 A need for flexibility and modularity: the Concept Building Blocks tool

Building on the experience of the *Introductory Game to ML Responsible Design* and aiming to **increase freedom and agency** for design students learning how to handle ML, a further iteration in the operationalization of the theoretical materials for the translation was necessary.
On this matter, according to Stolterman's (2008) view about what (interaction) designers appreciate, tools should be precise and straightforward. While – without prescribing design actions – frameworks, concepts, or even philosophical ideas and approaches should be presented in open-ended ways that engage, intrigue, support reflection, facilitate decision-making and expand design thinking. Thus, following this inspiration, a new tool has been developed to synthesize all the lessons learned so far.

5.2.1 A versatile approach for non-expert introduction.

From the premises stated above, the effort has been directed toward constructing a flexible and modular tool. Avoiding prescriptiveness, in fact, learners should not feel trapped in a theoretical, or rather rhetorical, cage. On the contrary, they should be **thrilled to explore new knowledge at their own pace and critically put in place their own interpretation**, which is the trigger for radical innovation (Norman & Verganti, 2014).

Additionally, flexibility would allow the translation of ML knowledge to adapt to different contexts and serve different design purposes, as the potential applications of ML can range over multiple fields. Modularity could also benefit in several ways. It should be possible to break down the tool into different parts that could convey specific pieces of knowledge independently and in relation to the most diverse topics. Each component might make sense on its own, and multiple uses can be imagined according to the task at hand. Indeed, as an operationalization of a theoretical structure, it should be open to personal interpretations but with a solid organizational reference.
For designers, this kind of fruition is particularly achievable in an applied setting, where knowledge is not consumed for the sake of knowing but for understanding how it can serve practical purposes. Indeed, many essential competencies for design are tacit and can be learned only by exposing students to solving problems in a designerly way (Weil & Mayfield, 2020).
According to the foundational assumptions derived from the research (section 4.5), the tool aims to support the development of both a consistent and responsible ML solution. In the name of flexibility and modularity, though, these requirements can be separated and employed individually. Here, the core part related to building a design-driven consistent application of ML capabilities is described.

5.2.1.1 Purpose

The *Concept Building Blocks* tool – CBB henceforth – (Fig. 5.11) is a quite literal representation of the *ML Designerly Taxonomy*. It is conceived as a versatile tool that can serve multiple purposes. On the one side, it aims to **convey the pieces of knowledge** framed in the theoretical construct, highlighting what ML systems can actually do and how their capabilities may relate to broader design objectives. Of course, this can include identifying, analyzing, and making sense of existing ML solutions. On the other, it might also be an **inspirational tool**, suggesting to ML systems' designers (both from a technical or non-technical background) what

INTENT DECK	Augment (1) Automate (2)	Augment (1) Empower (3)	Augment (1) Empower (3) Inspire (4)	Automate (2) Empower (3) Augment (1) Inspire (4)	Empower (3) Automate (2) Inspire (4)	Automate (2) Empower (3)
DESIGNERLY DECK	Organize (9) Recognize (6) Detect (5) Analyse (8) Select (7) Specify (10)	Plan (13) Specify (10) Forecast (12) Calculate (11)	Suggest (15) Personalize (14) Summarize (16) Forecast (12)	Edit (17) Reconstruct (18) Create (19) Interact (with) (20)	Personalize (14) Distinguish (21) Summarize (16) Match (22) Research (23) Suggest (15) Analyse (8) Organize (9)	Plan (13) Make decisions (about) (25) Suggest (15) Personalize (14) Co-evolve (with) (24) Interact (with) (20) Analyse (8) Create (19)
CAPABILITIES DECK	Object detection (29) Image classification (26) Sentiment analysis (30) Anomaly detection (32) Language identification (28) Image/Face recognition (27) Sensor data analysis (34) Speech/Voice recognition (31) NLP (33)	Numerical value prediction (35)	NLP (33) Word sequence (37) Recommenda-tion (38) Summarization (36)	NLP (33) Image restoration (41) Image transformation (42) Scene reconstruction (43) Text generation (45) Image generation (40) Sound generation (44) Content generation (39)	NLP (33) Embedding (words, sensor streams...) (46) Identifying hidden patterns (49) Highlight features (48) Generalization (47) Anomaly detection (32) Segmentation (52) Sensor data analysis (34) Information retrieval (50) Question an-swering (50) Recommenda-tion (38)	Translation to/from speech (33) Task planning (58) Motion planning (54) Movement control (55) Preference learning (57) Game playing (53) Optimization (56)
INPUT DECK	Activity logs (65) Audio content (66) Numerical properties (67) Sensor logs (68) Visual contents (69) Written contents (70)	Activity logs (65) Audio content (66) Numerical properties (67) Sensor logs (68) Visual contents (69) Written contents (70)	Activity logs (65) Audio content (66) Numerical properties (67) Sensor logs (68) Visual contents (69) Written contents (70) Historic data (71)	Audio content (66) Visual contents (69) Written contents (70)	Activity logs (65) Audio content (66) Numerical properties (67) Sensor logs (68) Visual contents (69) Written contents (70)	Audio content (66) Visual contents (69) Written contents (70) Environmental information (72)
ML AGENTS DECK	CLASSIFICA-TION (59) Is or is not?	REGRESSION (60) What number?	SEQUENCE PREDICTION (61) What's next?	GENERATION (62) Anything new?	CLUSTERING (63) Are things related?	ACTION SELECTION (64) What to do?
OUTPUT DECK	Typology (73) Anomaly (74) Quality (75) Sentiment (76) Subject (77) Topic (78)	Dimension (79) Quantity (80) Time (81) Value (82) Location (83) Preference (84)	Topic (78) Location (83) Event (85) Need (86) Action (87) Audio content (88) Behavior (89) Move (90) Suggestion (91) Trend (92) Visual content (93) Written content (94)	Subject (77) Audio content (88) Visual content (93) Written content (94) 3D model (95) Performance (96)	Anomaly (74) Quality (75) Sentiment (76) Subject (77) Topic (78) Preference (84) Trend (92) Affinity (97) Difference (98) Groups (99) Information (100) Relation (101) Response (102)	Sentiment (76) Subject (77) Preference (84) Action (87) Behavior (89) Move (90) Suggestion (91) Performance (96) Response (102) Decision (103)

Tab. 5.1 | List of cards and distribution according to the possible connections. In parentheses, the number printed on each card.

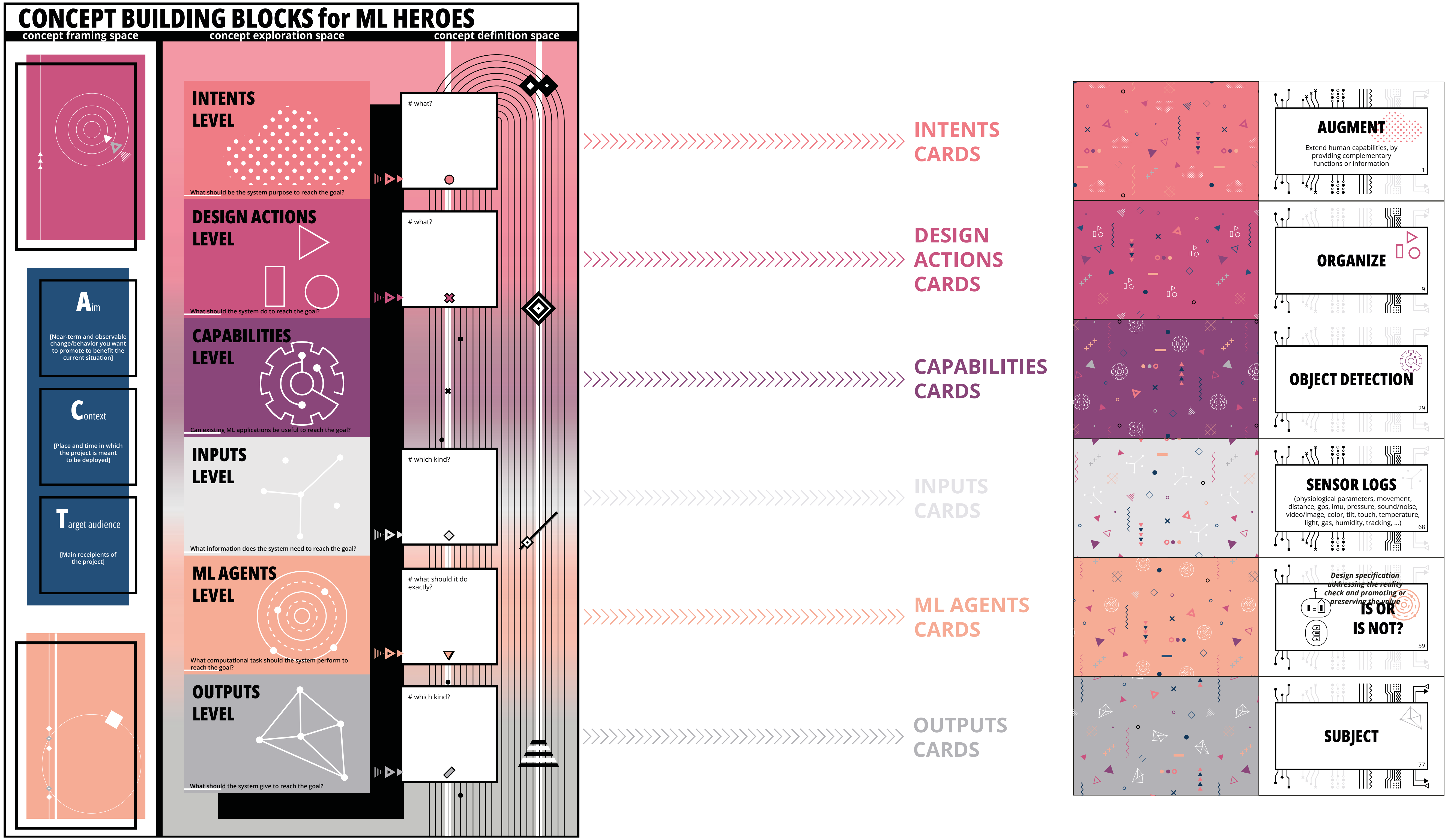


Fig. 5.11 | Concept Building Blocks board (on the left) and examples of cards (on the right).

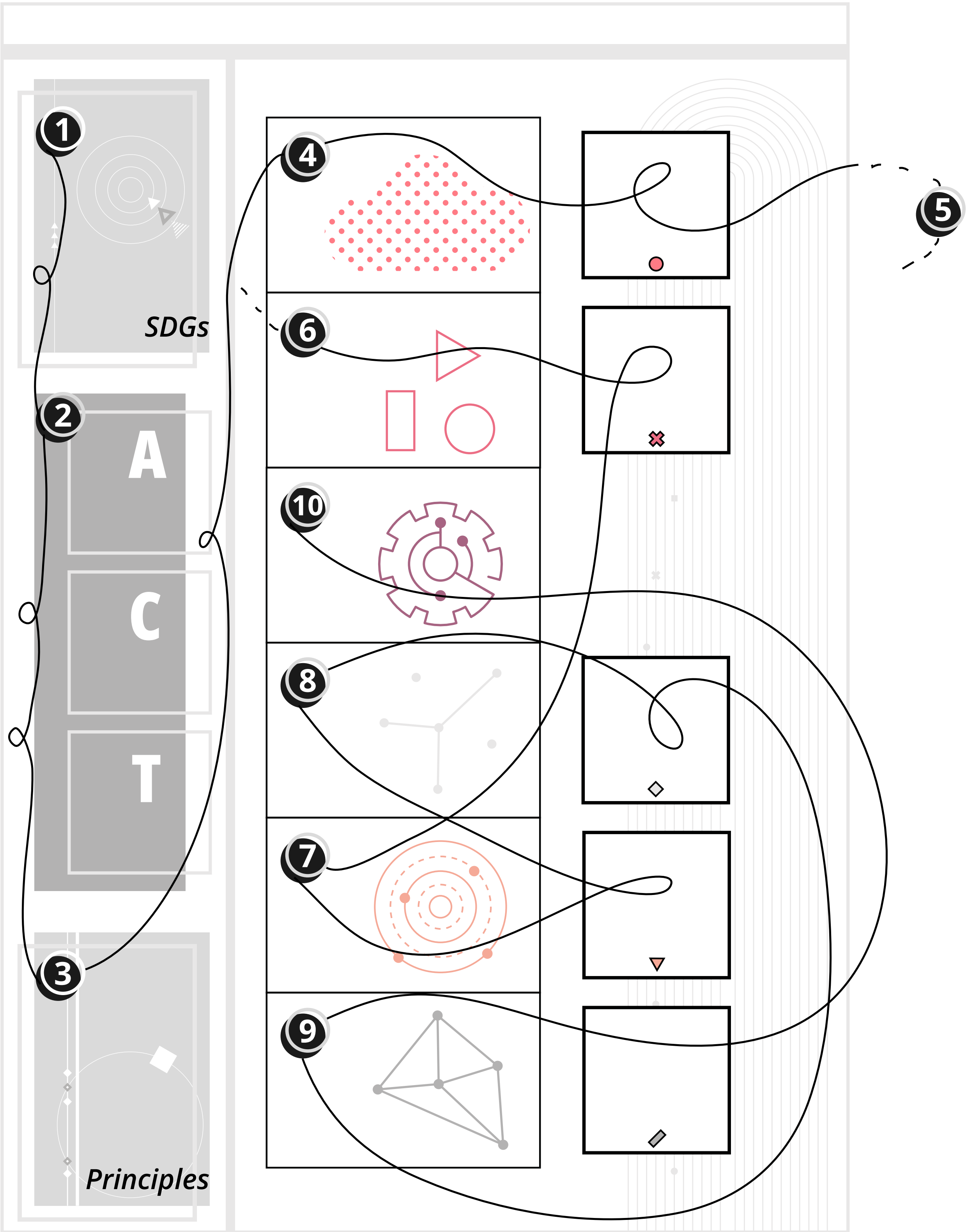


Fig. 5.12 | Example of Human-driven instructions as portrayed in the concept compass booklet.

Value-based

Do you want to promote or protect values in a particular context or while pursuing a goal?

1. **IMPACT GOALS SPACE |** (optional) Are you working to achieve a long-term, significant impact, like one of the United Nations' Sustainable Development Goals? You can choose one from the SDGs cards (or create one of your own) and place it on the board as shown in the picture.

2. **A.C.T. SPACE |** Outline your brief according to its basic elements (Aim, Context, and Target audience).

3. **PRINCIPLES SPACE |** Which of the overarching principles in the related deck may drive your system towards reaching your A.C.T.? (Note: it won't be the only principle embedded, just the main one). Place the selected Principle card as shown in the picture.

4. I**NTENTS LEVEL |** Can one of the Intent cards reflect your objective? If so, place it on the board and specify what your intent refers to (now or later in the concept definition).

5. ***Check* |** No Intent card reflects your actual intent? Maybe ML is not the best tool to reach your goal. You can use the ML Suitability Matrix to check. If ML is a consistent solution for your problem, go to the next step. If not, look for alternatives (this tool has no more use for you) or reframe your aim to be more aligned with ML capabilities.

6. **DESIGN ACTIONS LEVEL |** How could your system fulfill your intent? Consider: one ML agent can do just one specific thing. In general, they are great at replacing humans in burdensome and repetitive tasks, at saving time, energy, and resources, and they could also free people from the digital world. Specifically, what would you like it to do to achieve your design goal? Also specify the object of the design action (now or later in the concept definition).

kind of agent can be helpful for a given purpose or to which goal a ML system may aspire. However, while intrinsically including the previous, its **primary function is procedural, envisioned especially for a didactic context**. It gives hints about how the elements of the theoretical construct combine with each other but leaves freedom for interpretation and practical implementation. In fact, the tool has been developed to guide and support students in the concept generation of a meaningful product or service integrating a ML system. Although conceived primarily for design students, it may also be employed as educational support for computer scientists and engineers to foster a design-driven approach and expand their reflections beyond algorithms and coding. As well it could serve as a collaborative boundary tool to facilitate collaboration between designers and ML experts.

5.2.1.2 Description of the tool and its elements

In the attempt to be as much comprehensive as possible, the CBB tool does not offer a unique path that anyone must follow to design a ML system. Instead, it tries to **address the possible initial conditions** that one might encounter when starting a project, like having a predefined design brief, an available dataset, a specific system to implement, an artifact to redesign and improve, a moral behavior or value to advocate in society, or just a will to explore ML systems for good outcomes. The offered approaches, depicting these different entry points for the concept generation, are respectively *Problem-based, Data-driven, Technology-driven, Artifact-driven, Value-based, or Human-driven.* The order does not reflect their level of importance and, despite the names, all of them are based on a designerly attitude to addressing problems.

They are collected in a booklet, the *Concept Compass* (Fig. 5.12), to be consulted as an introduction to the tool. In a question format, it describes the elements characterizing each approach – so that one can recognize the closest to their own situation – and then proposes a non-binding step-by-step guide to the CBB.

The tool is also composed of a board supporting the concept framing, exploration, and definition and some decks of cards portraying the contents of the taxonomy with functional integrations.

The *concept framing space* on the board provides areas for the foundational elements of the concept. The central one is called *design A.C.T.,* and it entails the definition of the basic elements of problem framing: *Aim, Context,* and *Target audience*. The *concept exploration space* is based on the structure of the *ML Design Taxonomy*. It consists of different levels, spanning from more designerly aspects (top) to more technical ones (bottom), complemented with cards showing the related contents, which ultimately will depict the skeleton of the concept. (The list of the cards for each level is illustrated in Tab. 5.1.) From the upper part, they include *intents, design actions, capabilities, inputs, ML agents,* and *outputs* levels. Inputs and outputs are added to the theoretical construct and introduce practical and essential information to orientate the development of ML systems.

One rule governs the establishment of the structure. Each black or gray symbol on the cards (depending on whether it is highlighted or not) represents the possible connection with a particular capability of a ML system (*ML Agent*). When black, the symbols communicate that what is on the card is consistent with the related ML capability. Two cards can only be combined if at least one of the highlighted symbols

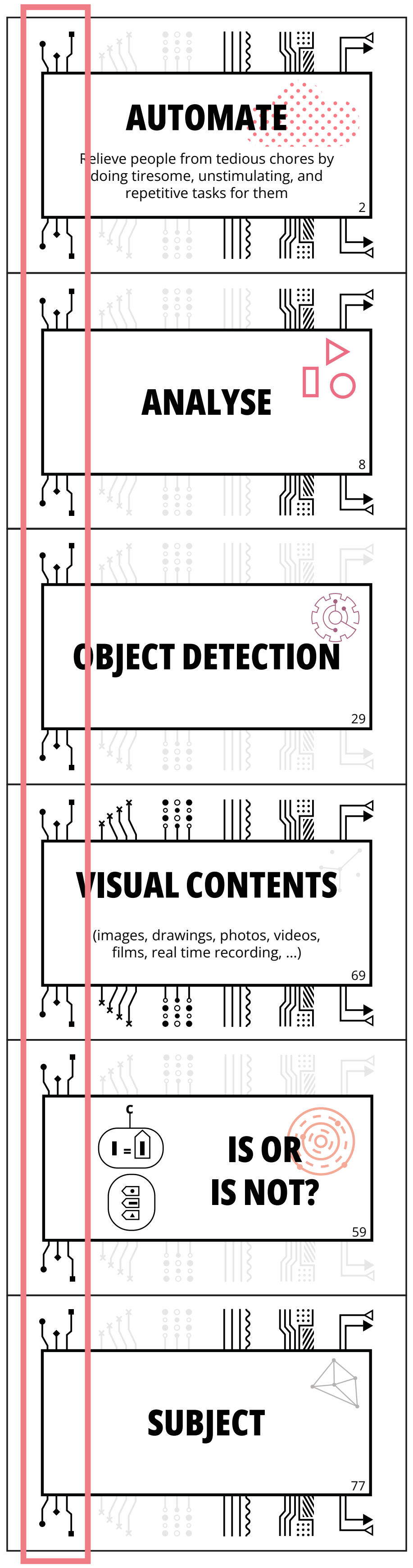


Fig. 5.13 | Cards correctly placed with a continuous line of matching symbols.

matches. To build a complete structure, compatible with the selected *ML Agent*, all the cards must possess the corresponding symbol, depicting a continuous line throughout all the levels (Fig. 5.13). This should help students individuate the relationships between *ML Agents* and the other constituents of the structure, also inspiring coherent potential solutions.

In case some connections or elements (from almost any level) are missing, the cards can be customized accordingly. (However, *ML agents* and *intents* cards are most likely fixed because at the basis of the systematization.)

The last part of the CBB is the *concept definition space,* which allows the personalization of the system concept. It comprises squared blank cards marked to match the different levels. These can be used to write details of one's idea, specifying the general suggestions provided by the *concept exploration cards*.

Finally, the design process suggested by the CBB can be complemented with the *ML Suitability Matrix* (section 3.3.1) to reflect on and assess whether it makes sense for the solution under development to be integrated with ML capabilities and if it is aligned with what can be valuable for people.

According to the approach selected with the *concept compass*, the parts of the main board can be addressed in a different order, even randomly, if one decides not to follow the given instructions.

5.2.2 When ethics come into play: a VALUable by Design expansion

While the basic format of the CBB tool facilitates the generation of designerly conceived concepts consistent with the technology at hand, an implementation was needed to integrate a responsible and value-driven approach to the design of ML systems.

This is the goal of the *VALUable by Design expansion* (VDE henceforth). Inspired by principles from responsible research innovation – RRI (von Schomberg, 2013) – and

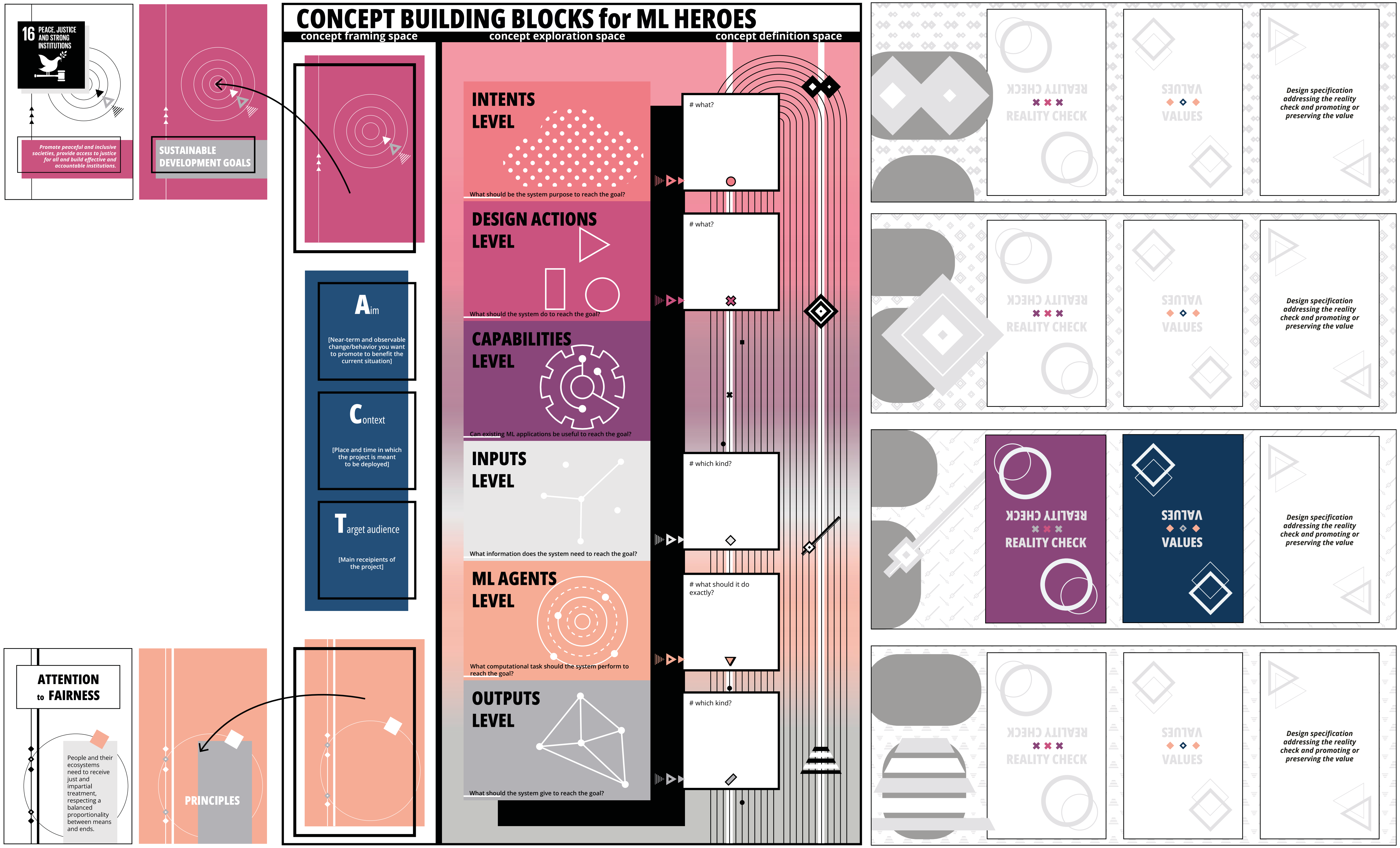


Fig. 5.14 | VALUable by Design Expansion complementing the Concept Building Blocks.

value sensitive design – VSD (van de Poel, 2020; van den Hoven, 2013), it merges the theoretical assumptions learned so far and stresses the importance of considering values and possible impacts to drive the early stages of the design process. Explicitly tackling those issues not only is in line with recommended ethical practices (van den Hoven, 2013), but it also encounters the suggestions of (Redström, 2020), who stresses the necessity to include theory in the design practice. It can help students increase their awareness of the positive and negative consequences that their decisions in the concept generation phase might have. In this case, additional guidance is of utmost relevance as students are invited to explore possibly unfamiliar areas.

Basically, the expansion uses the same mechanism as the original tool with a few modifications. First, the *concept compass* has no use, as only one path is suggested to promote the understanding of the approach. Then, students primarily encounter the *concept framing space*, which presents the significant addition of two card decks (Fig. 5.14). To direct the activity towards beneficial outcomes – the right impacts or grand challenges advocated by (von Schomberg, 2013) – a general objective from the United Nations' Sustainable Development Goals (SDGs) and a principle to drive the concept generation must be selected from the related decks. The SDGs have been chosen as overarching goals because of their broad nature and wide recognition, which implies a good availability of online resources for supplementary investigation, such as (AI for Good, n.d.). Instead, the suggested principles are those derived from the analysis of ethical guidelines on AI (section 4.4.3).

The next part of the tool, the *concept exploration space,* remains unvaried, with the only recommendation to complete it from top to bottom. Instead, the *concept definition space*, in addition to the detailing of the concept structure, presents further *concept specification boards*. Once the concept has begun to take shape and its coherence with the preselected principle has been checked, these boards, supported by *Reality Check* and *Value* cards, induce higher levels of reflection and encourage responsible additions to the concept. Both decks represent a synthesis of the content analysis of the ethics guidelines (4.4.3).

Ten values, more finely articulating the principles, emerged as the most recurrent and relevant for designing ML systems: *accessibility, agency, explainability, freedom, justice/non-discrimination, reliability, representativeness, sustainability, (meaningful) transparency,* and *well-being.* Instead, twenty-five *Reality Check* cards synthesize possible risks and implications affecting the implementation of ML systems that designers can anticipate and address. These are presented with designer-friendly terminology and questions helping to determine the consistency of the issue in relation to the ideated concept and encouraging the research of solutions without giving any hint. Based on the previous steps of the investigation, in fact, it is more educational and stimulating to let design students think about possibilities to counter potential risks without giving them the references that abound in ML-related ethical guidelines.

Summing up, students have to draw a *Reality Check* card. If coherent with the concept under development, they need to mitigate the identified risks by choosing a *Value* to promote or preserve and placing both cards on the board. To finally complete it, a blank card should be filled with a design specification that would keep into account

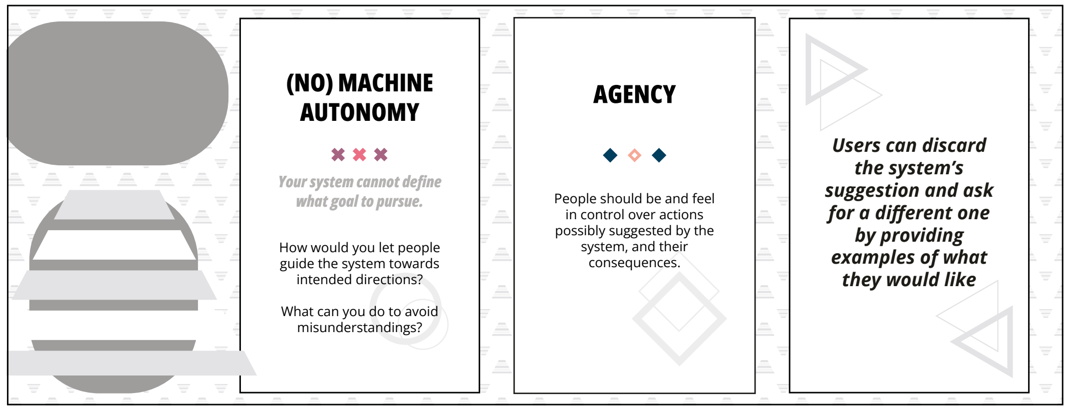


Fig. 5.15 | Example of a completed specification board.

the previous elements (Fig. 5.15). In total, four *specification boards* are provided, but the exercise could be limited or extended at will.

The one described is the main envisaged application for the VDE materials. However, in favor of flexibility and modularity, the *Reality Check* and *Value cards* may find further applications as independent foresight or testing tools during or after the design process of products and services that integrate ML or similar technologies.

5.3 First experimentations: Superpowered Museums

Once again, the synthesis reached in the planning phase and culminated in the CBB tool needed practical experimentation to assess whether it could be a viable solution to operationalize the previously developed theoretical constructs. Building on previous experiences with DID and Ph.D. students, and having collected some feedback about the preferable forms, languages, and requirements for the multidisciplinary transfer of ML knowledge, this time, the CBB tool had to be tested for being able to **(i) allow a broader understanding of the translation**, and **(ii) adapt to a well-characterized design education environment.** To do so, the *Superpowered Museums* workshop was organized. Also in this case, a synthetic overview (pages 182-183) and a detailed description are offered in the following.

5.3.1 Methods

5.3.1.1 Target audience and context

For the previously stated reasons, a challenging target audience was selected. Specifically, the workshop involved 38 students enrolled in the third year of the Interior Design bachelor's degree at Politecnico di Milano. For their background, the participants were not used to working with technology. They could have had no interest in the topic and less developed critical thinking than master or Ph.D. students. However, they surely had the right mindset, which was enough to test if the ML translation could reach any design student with some education on how the design process applies in practice. Of course, they were not expected to have previous knowledge of ML.

Superpowered Museums

Contextual information

WHAT	One-and-a-half-day workshop (12h).
WHEN	22-23 November 2021.
WHERE	Politecnico di Milano, Interior Design School, Final Synthesis Studio on *"Myths, heroes and superheroes: exhibition narratives and itineraries"* by professors Claudia Baldi, Raffaella Trocchianesi, Ilaria Bollati, and Paolo Maria Fumagalli.
WHO	The author.
STUDENTS INVOLVED	38 interior design students enrolled in the 3rd year of the bachelor program in Politecnico di Milano. Participation was mandatory as they were attending the Final Synthesis Studio that hosted the workshop.

Research Rationale

RQ UNDER INVESTIGATION	***RQ3: How to operationalize theoretical constructs into models and tools to be implemented and tested in educational contexts?***
RESEARCH OBJECTIVE(S)	(i) Assess the effectiveness of the CBB (in an adapted version) as operationalization of the ML Designerly Taxonomy. (ii) Understand if the CBB allows a broader understanding of the translation. (iii) Test the CBB and the approach's flexibility to adapt to a well-characterized design education environment. (iv) Prove the versatility of the contents to address different kinds of design students.
TARGET AUDIENCE RELEVANCE	Interior design students were selected as a challenging target audience. For their background, they were not used to working with technology. Of course, they were not expected to have previous knowledge of ML. They could have had no interest in the topic and less developed critical thinking than master or Ph.D. students. With these premises, they were relevant to test if the ML translation could reach any design student with some education on how the design process applies in practice.

Methodological framing

EXPLORATION STRATEGY	The workshop was adapted to the topic and context of the Final Synthesis studio by introducing specific contents to bridge ML systems and exhibition design for cultural experiences and by assisting to the student's in-itinere projects presentation the week before. In that occasion, a preliminary questionnaire, and a practical exploration of examples of ML systems (with no explanation) were launched for the students to complete before the workshop. It then developed with an initial theoretical introduction and then it mainly focused on a practical design activity, supported by the researcher, and in which the students worked in 10 groups of three or four people, as they were doing for the studio. Students were also asked to keep track of their process along the way. The design activity concluded with final presentations, a peer-review session, a conclusive discussion and a post-workshop questionnaire to complete the data collection.
DATA COLLECTION	• Questionnaires (pre, during, and post-workshop) • Observation • Visual and conceptual output produced by the participants • Students' presentations and final discussion (subject of a subsequent content analysis)
RESEARCHER'S ROLE	Participant observer (facilitating the educational activities while gathering feedback and insights).

Structure of the educational activity

ILOs	***Knowledge***	• Understand **basic ML capabilities** and infer their potentialities. • Familiarize with a responsible design approach.
	Skills	• Approach ML as an **interior design material.** • Recognize and introduce ML systems to solve **complex problems.** • Build design scenarios to **explore limits and potentialities** of envisioned concepts.
	Values	• Identify **ML as an asset** also for interior design. • Understand the importance of foresight in the early phases of the design process. • Understand how technology can shape environments and behaviors.
EXPECTED IMPACT(S)		**Build basic awareness** on ML to instill curiosity toward the topic.
CONTENTS		• Basic ingredients of ubiquitous computing and pervasive interaction • ML demystification • AI and ML definitions • Introductory elements of responsible design • Science Fiction Prototypes (SFPs) as foresight tools. • ML capabilities
TOOLS	***Knowledge transfer***	• Theoretical introduction • ML Agents (with practical examples previously explored by the students individually)
	Design activities	• CBB (with no capabilities level, not to influence the envisioning of futuristic solutions, and with two approaches: design-driven and technology-driven, each tested by 5 groups)
OUTPUT		ML-infused concept of a new (futuristic, unusual, unexpected, etc.) and more meaningful, stimulating, educational, or exciting visitors' experience of a space in an exhibition or museum context. Represented by a: • *movie poster* • *movie trailer*

Findings

KEY INSIGHTS	• Overall, the provided knowledge seemed mostly consistently assimilated and applied by the students, underlying the effectiveness of the theoretical systematization and tools. • The explanation of ML Agents seemed easily understandable and the practical experimentation with the case studies triggered reflections on the potential limits of ML systems. • Interior design students proved capable and at ease in facing the unusual challenge (compared to their educational background) and using a new material. • All the solutions responded to the brief, and ML did not cause the participants to be diverted from the objective of creating new experiences for entertainment in cultural institutions. • Students seemed very engaged in the design activity, and some reported they had fun in the process. • Implementing the necessary expedients, CBB can be a flexible means for the translation, as it can support even not obvious design challenges. • CBB enabled to create consistent connections and to reason "in blocks" as a way to simplify the process. Also appreciated were its capabilities to visualize, experiment, clarify, present, and create solidity in the process and stimulate effective and innovative ideas. • ML was always perceived as a tool in this process, not the end of the whole experience.
ISSUES FOR FUTURE INVESTIGATION	• The educational method should concentrate on the needs and predisposition of design students to foster their personal understanding and sense-making, leveraging on their domain knowledge to help them build confidence in dealing with an unfamiliar topic. • More explicit ethical reflections should be fostered as the envisioned solutions raised some concerns due to a peculiar correlation between the disappearance of the interface and an apparent user's loss of agency. • The strong interest in deepening technology-related understanding and skills demonstrated by the students should reflect in education. In their perspective, keeping technology separated from their educational path would be anachronistic. Hence the research could consistently address any kind of design specialization.

Participation in the workshop was not voluntary, as it was included in the formative program of a Final Synthesis Studio held by professors Claudia Baldi, Raffaella Trocchianesi, Ilaria Bollati, and Paolo Maria Fumagalli. The studio represented a suitable context for the experimentation because, in addition to the helpfulness of the faculty in hosting the workshop, it offered a predetermined design context in which ML knowledge had to be incorporated. It dealt with cultural enhancement with the purpose of developing new scenarios for museum fruition and new exhibition systems in the field of narrative museums, specifically focusing on the theme *"Myths, heroes and superheroes: exhibition narratives and itineraries."* Thus, the studio represented a chance to test the flexibility of the translation in an atypical environment for issues most commonly related to interaction, digital, or UX design.

5.3.1.2 General framing: procedure, structure, and contents

To bridge ML systems and exhibition design for cultural experiences in a twelve-hour, non-technical workshop, two strategies were adopted: the introduction of ubiquitous computing (Weiser, 1993, 1994) to connect technology and interior spaces and Science Fiction Prototypes – SFPs (Johnson, 2011; Keinonen, 2006) as a way to put theory in practice, without worrying about the feasibility of the solutions.
The experiment merged research and educational activities to enable design students to understand basic ML capabilities, consistently infer their potentialities, and introduce ML-infused solutions to enhance spatial experiences.
The preparation for the workshop started the week before. The researcher got acquainted with the projects that the students were developing for the Final Synthesis Studio to gather insights for effectively connecting the workshop activity with the course. Then, the participants were asked to complete two tasks before the workshop. They had to fill in a questionnaire, useful for the researcher to infer their prior knowledge and preconceptions regarding ML and the relationship between technology and physical spaces, and explore the case studies illustrating *ML Agents* (section 4.3), with no preliminary explanation.
The workshop was developed over two days. During the first one, a preliminary phase included the introduction of a few theoretical contents: basic ingredients of ubiquitous computing and pervasive interaction, truth and myths about ML, introductory elements of responsible design, and SFPs as anticipatory instruments for envisioning concepts. *ML Agents* were also presented to introduce ML capabilities and explain the connection with the examples given the week before.
The rest of the workshop focused on a practical design activity, for which the students worked in 10 groups of three or four people, as they were doing for the studio. Each group was asked to imagine a new (futuristic, unusual, unexpected, etc.) and drastically improved visitors' experience of a space in an exhibition or museum context to make it more meaningful, stimulating, educational, or exciting. The experience had to be enabled by (at least) one *ML Agent* and take the form of a superpower that they attributed to the space itself.
To support the concept generation, a variation of the CBB tool was provided (Fig. 5.16). Testing its modularity, in fact, the capabilities level – indicating current applications of ML systems – was excluded, not to overwhelm and influence the envisioning of new solutions with what already exists. Additionally, the *Concept Compass* was not

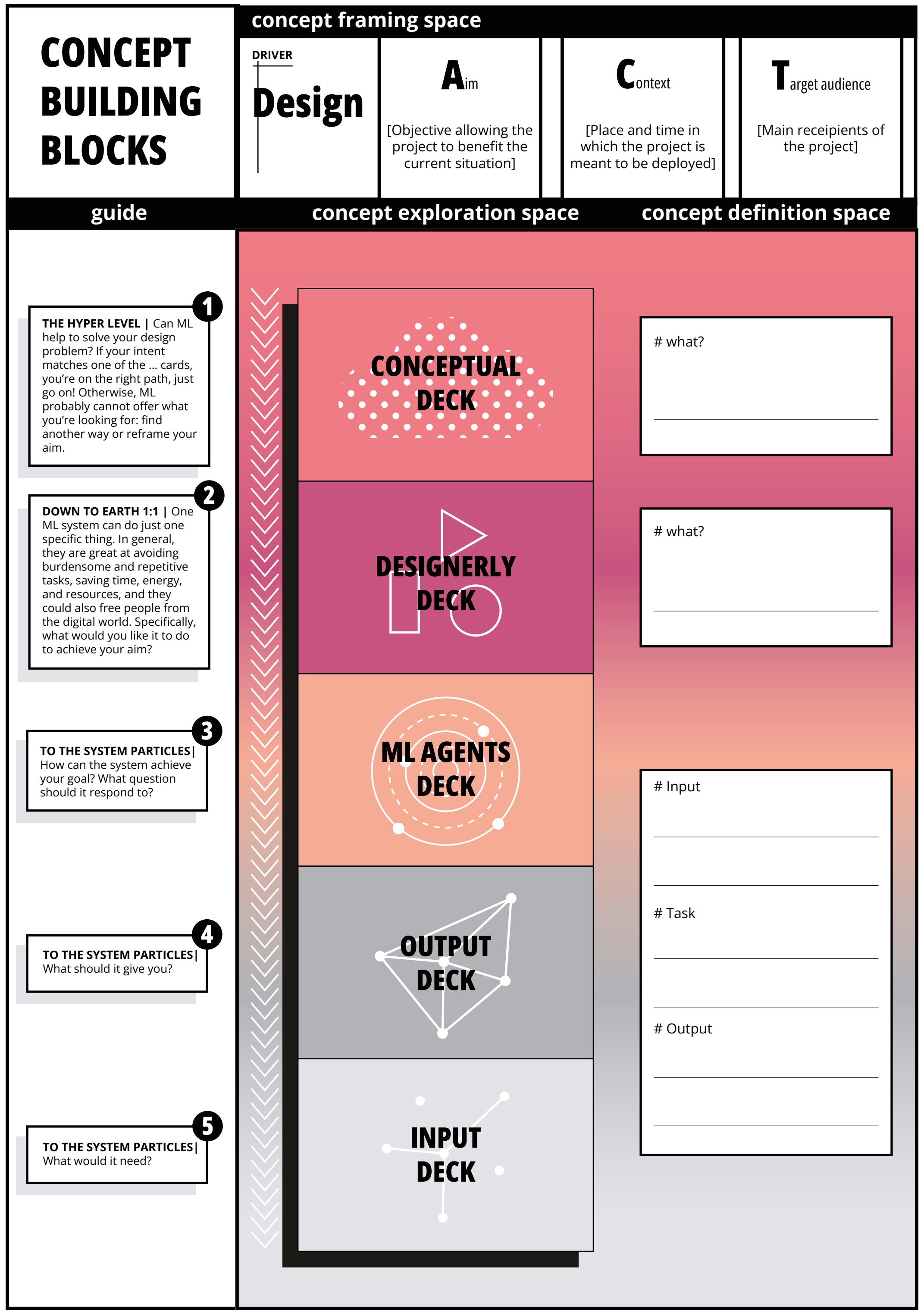


Fig. 5.16 | Adapted CBB board for the design-driven approach .

provided. Instead, in a *guiding space* on the board, the researcher proposed two paths that were equally distributed among the groups in relation to the projects they previously developed for the Final Synthesis Studio. Those already envisioning experiential exhibitions were associated with a design-driven path (corresponding to the problem-based one on the *Compass*) as they could frame the workshop concept as

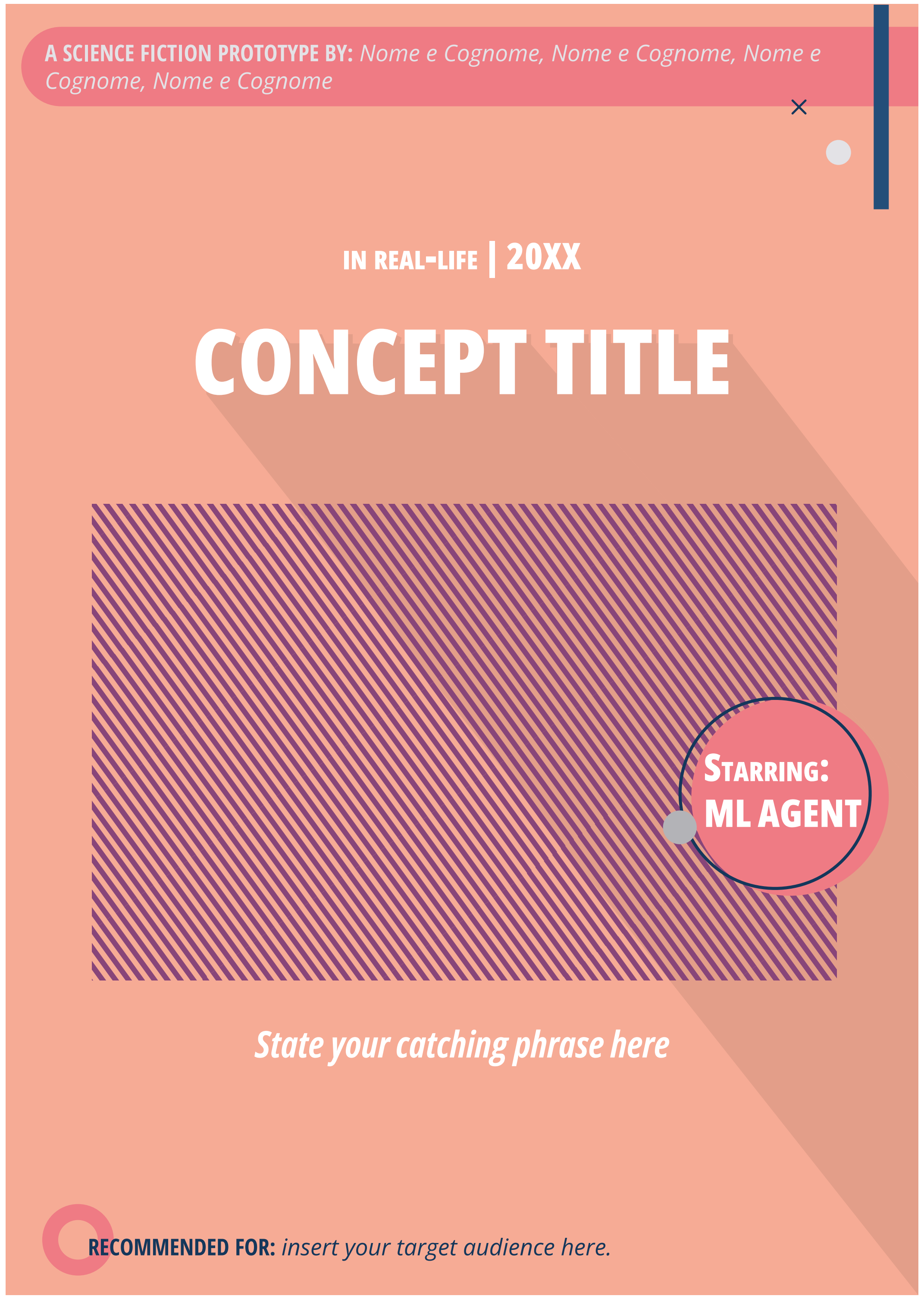


Fig. 5.17 | Predetermined poster layout including: concept's authors, foreseen time of deployment, employed technology (ML agent), target audience, and main features expressed by a title (including the superpower name), a catching phrase, and an evocative picture.

an extension of their previous work. The others were given a technology-driven path, using *ML Agents* to inspire potential spatial experiences.
Then, to maintain a parallel with the cinematic tradition of science fiction and to provide means to communicate their ideas, the participants had to synthesize their concepts in a pre-defined *movie poster* (Fig. 5.17) and to build a more compelling and layered narrative in a *movie trailer* represented with a storyboard. Both outputs, together with the completion of the CBB tool, had to be uploaded on a common Miro Board. To fulfill such activities, students were free to manage their own time, although they were suggested a time span of two and a half hours to deliver the first output and devote the entire afternoon (four hours) to complete the second one.
Day two was reserved for the *première*: groups' presentations of their *movie trailers* (concepts). The *critique*, a collective peer review of each idea, followed. Seven minutes per group were reserved for each phase. Ultimately, the workshop was wrapped up by the *end credits*: a post-workshop questionnaire was submitted, and a conclusive discussion session was triggered by few questions on the Miro board and live.

5.3.1.3 Methods of inquiry

In the experimentation, the researcher played the facilitator role and applied multiple data collection methods (Robson & McCartan, 2015).
Pre- and post-workshop questionnaires aimed to assess their prior knowledge and preconceptions about ML, their perceived relationship between digital and physical worlds, and their perspective on the possible role of technology in interior design. Additionally, combined with the concept development results, they aimed to assess the effectiveness of the knowledge transfer and the provided tools.
Then, throughout the workshop, insights and data were collected through the researcher's participant observation of the groups' activity, the requested visual materials uploaded by the participants on a Miro board, and the collective presentations (Figure 5.18) and discussions that were audio and video recorded for later transcription and analysis. (Students signed a photo/video recording consent form for research dissemination.)
The insights are derived by a content analysis articulated in a first cycle of *initial coding* breaking down the qualitative data to infer categories and themes in a second cycle of *focused and axial coding* (Saldaña, 2009).

Fig. 5.18 | Concept presentation moment .

#Gr.	Concept	Input-Task-Output	ML solution	ML Task Consistency	ML Impact	Envvironmental UX	Typology of experience	Interface	Value/cost	Personal Data Sharing
#01	**Flashforward – Trust your future self (year 2355) |** To facilitate flows in a museum experience, the director possesses all visitors' personal data so that, when they are faced with a crossroads, the environment projects a bright shadow anticipating one's probable route choices. This stimulates people to follow the data-driven way or not. Either way, they are pushed to move forward.	(i) Historical data (habits, behaviors, appearance of the visitor) - (t) **Sequence prediction** - (o) Event (future alter-ego projection to show the most probable route)	Very much	Somewhat	Prominent role	Natural interaction + Cognitive load	Spatial distortion (environmental augmentation)	None	Somewhat	High
#02	**The fear inductor – Megamind (year 2035) |** In a very introspective exhibition, visitors are challenged to face their own fears. An aseptic environment welcomes the subjects, whose body data are collected to capture their parameters in a resting situation. Then, the setting gradually materializes the visitors' deepest fears – by changing its physical characteristics or inserting new elements.	(i) Sensor logs (Images, videos, 3d models, and physiological reactions of the visitors) - (t) **Clustering** - (o) Sentiment (spatial and sound alteration)	Somewhat	Very much	Indirect activator	Natural interaction	Spatial distortion (adaptive environment)	None	Little	Medium
#03	**What's this? – Imagine, you can (year 3250) |** To enable future inhabitants of planet Earth to experience glimpses of the extinct human civilization, an exhibition allows visitors to test their imagination and put it on stage. Before entering a special room, they must wear a headset and mentally answer the questions: "How do you envision human society, and how do you think they used the objects in the exhibit?". A ML generation system detects their thoughts and elaborates the display of a room, including the freshly inputted personal visualizations and the ones of the previous visitors. The result is an unexpected and surreal environment.	(i) Intellectual suggestions (memories, stimuli, and visitors' ideas) - (t) **Generation** - (o) 3D model (3D immersive environment)	Somewhat	Very much	Prominent role	Natural interaction	Spatial distortion (adaptive environment)	None	Little	Low
#04	**Portable world – Explore the world behind your mask (year 2200) |** In front of a mirror, a classification system detects visitors' attitude and associates different kinds of masks with each person. A generation agent creates custom worlds they access through the mirror/portal. Here, people can meet others who share the same qualities. They can come back anytime through a pocket mirror provided at the end of the exhibition, with the risk of this extremely personalized reality becoming prevalent.	(i) Pictures and videos of 5 types of fictional characters in motion - (t) **Classification + Generation** - (o) The mirror/portal transports the visitor into a world based on the category characteristics, changing the appearance of the visitor	Very much	Very much	Indirect activator	Natural interaction	Spatial distortion (distortion) + Perceptive distortion (transfiguration)	Threshold	Little	Medium
#05	**Metamorphosis – Discovering a new sensory experience (year 2100) |** An exhibition on fantastic creatures enables visitors to immerse themselves in the beasts. A ML system derives the qualities of the fictional creature's sensory perception starting from real animals' ones. Then, these are reproduced and experienced by visitors through special suits and viewers.	(i) Pictures and videos of 5 types of fictional characters in motion - (t) **Generation** - (o) Performance (Reality perception from the fantastic animal's point of view)	Very much	Very much	Indirect activator	Sensorial perception	Perceptive distortion (transfiguration)	Concrete filter	Very much	None
#06	**Self-vision – Your bespoke space (year 2040) |** Through visors, a blank room becomes a totally personalized exhibition space. Visitors generate their alter ego – a fantastic creature determined by conscious choices and the detection of their behavior – to be guided and perceive the environment "in its shoes," while the displayed contents vary by person.	(i) Visual contents (exhibition information, paths, contents) - (t) **Generation** - (o) Visual content (path and contents of the exhibition)	Very much	Very much	Indirect activator	Sensorial perception	Perceptive distortion (transfiguration) + Spatial distortion (adaptive environment)	Concrete filter	Somewhat	Medium
#07	**Insight – The museum that knows you (year 2080) |** During the visit, an exhibition space collects the behaviors and reactions of the visitors. Then, a hyper-personalized VR experience with tailored content is generated to encourage return and interesting sharing moments among the visitors.	(i) Sensor logs (behavioral parameters) - (t) **Sequence prediction** - (o) Event (exhibition with personalized contents)	Very much	Somewhat	Indirect activator	Natural interaction	Spatial distortion (adaptive environment)	None	Somewhat	Medium
#08	**Drop 720 – A fall in other dimensions (year 2100) |** An exhibition on Greek mythological female figures promises a 720° experience. Visitors are scanned throughout the exhibition to collect their bodily characteristics, behaviors, and even psychological traits. Based on these data, a female figure is associated with the visitor, who is sucked into a timeless vortex and brought back to the woman's world. Here, (s)he can experience and learn episodes of the mythological character by directly interacting with her dimension and even acquiring her perceptions and capabilities (e.g., in Circe's shoes, one can turn men into pigs).	(i) Sensor logs (visitors' scanning + historical figure data) - (t) **Clustering** - (o) Knowledge acquired through active experience	Somewhat	Very much	Indirect activator	Sensorial perception	Spatial distortion (displacement) + Empowering	Threshold	Little	High
#09	**Synderesis – The self-awareness of good and evil (year 2230) |** Set in a Milan where people have lost the consciousness of their being, the Synderesis project aims at restoring self-awareness by reactivating the distinction between good and evil. Visitors' past experiences are extracted from their memory at the exhibition entrance. As they walk through the main corridor, the most emotionally related to a personal conception of good and evil are projected to one's central nervous system as visual and auditory perceptions that overlap the physical reality. After the experience, all citizens are awakened from their indeterminacy.	(i) Sensor logs (past experiences, memories, sensations) - (t) **Clustering** - (o) Relation (images and audio grouped according to the visitor's conception of good and evil)	Very much	Very much	No physical realm	Sensorial perception + Cognitive load	Perceptive distortion (cognitive level)	None	Little	High
#10	An exhibition invites people seeking extreme sensations to "wear" superpowers. In the room of invisibility, visitors' bodily appearance is substituted by 3D projections to make them disappear. When interacting with items in the space, visitors perceive them floating and struggle to feel the boundaries of their own bodies until they get outside.	(i) Environmental information (physical characteristics of the room, objects, materials, textures) - (t) **Action selection** - (o) Projections to make it look like the visitor disappeared	Very much	Somewhat	Prominent role	Sensorial perception	Perceptive distortion (physically perceivable) + Empowering	None	Somewhat	None

Tab. 5.2 | Synthesis of the concepts generated in the Superpowered Museums workshop and related researcher's analysis.

The primary language used in the workshop was Italian, so the author translated the collected data. However, the tools provided were all in English to be aligned with research purposes.

5.3.1.4 Limitations, reliability, and validity

Of course, the study reveals some limitations. First of all, the sample is biased by a relatively small number of participants and a lack of diversity in their academic backgrounds. However, as the workshop was compulsory for the students enrolled in the Final Synthesis Studio, a skewed sample of people interested in the topic (like in the case of the *ML Pills for Designers* workshop) should have been avoided. Additionally, the data collection and analysis might suffer from a single researcher's perspective and the translation from Italian.
Referring to (Creswell, 2014), though, the adopted strategies to achieve qualitative validity of the research are the triangulation of different data sources, also by using mixed methods (again, with the limitation in the number of participants), and the attempt to provide a report of the methods and findings which is as transparent and objective as possible. To ensure the investigation's reliability, a careful and rigorous protocol has been followed for collecting and interpreting data. Starting from raw information and pursuing a coding activity to outline recurring themes and general descriptions, these were interrelated to infer the results.
Moreover, the study is structured to grant generalizability. Indeed, besides the unnecessary specific connection with the Final Synthesis Studio design goal, the workshop can be replicated in any interior design school following the same modalities and with no additional requirements.

5.3.2 Results

5.3.2.1 Concept generation overview

To facilitate the construction and communication of the concept scenario, the *movie trailer* was based on the SFPs , which are useful for envisioning and possibly shaping technological futures (Dourish & Bell, 2014). An adaptation of Johnson's science fiction prototyping method (Johnson, 2011) was proposed. It developed in five steps (Fig. 5.19): (i) *pick your science* – selection of the technology to be explored with the prototype and definition of the scenario set-up; (ii) *scientific inflection point* – radical change marking the transition to a new and advanced scenario; (iii) *ramifications of science on people* – exploration of the implications and ramifications of science in the

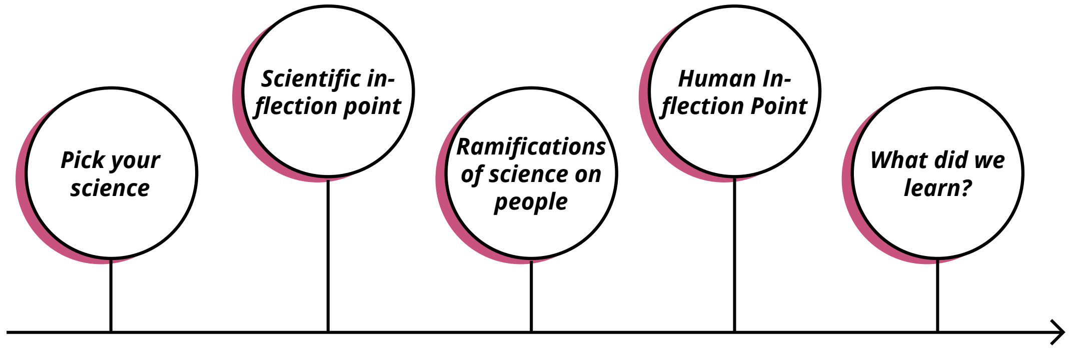


Fig. 5.19 | Steps to develop a Science Fiction Prototype (SFP).

imagined world: scenario evolution; (iv) *human inflection point* – repercussions on people's experience and sense of place at physical, personal, social, and cultural level (Tuan, 1977); and (v) *what did we learn?* – epilogue in which implications and solutions are evaluated.
Graphically transposed into an annotated storyboard, it included: *introduction* (presenting an unexpected opportunity modifying an initial situation), *situation setting* (depicting the scenario after the introduction of the superpower), *resolution* (how the characters and the scenario develop in the long-term), and *denouement* (bringing evidence of results – benefits and risks implied). Though, despite the indications, most of the works did not include the reflective steps of SFPs.

Fig. 5.20 | Movie poster and trailer presenting group #05's concept: Metamorphosis – Discovering a new sensory experience.

Fig. 5.20 portrays a complete *movie poster* and *trailer*, while all the concepts are summarized in Tab. 5.2. In most cases, the titles represent the envisioned superpower as suggested.
In general, the students seemed very engaged in the design activity, and some reported they had fun in the process. The outputs reflected this state of mind and, as they were asked to be daring in developing futuristic applications without worrying about the feasibility of the ideas, the solutions were adventurous and depicted scenarios even based on alternative realities or involving technologies that do not exist. They spanned from less than twenty to more than a thousand years from now and, combined with the not-yet-widespread practice of implementing ML to enhance the UX of spatial experiences, the ideas were inevitably more original than those developed in the *ML Pills for Designers* workshop.

5.3.2.2 Contents assimilation and consistent application of ML

As expected, the interior design students involved had no prior preparation on the topic: of the 35 respondents to the pre-workshop questionnaire, most declared to have little knowledge about ML and the remaining not at all (in a ratio of 60% to 40%). What is interesting is that their perception quite changed after the workshop (Fig. 5.21). Only 30 people answered the post-workshop questionnaire, but to the same question, the majority answered "somewhat" and one even "a lot," with the remaining five attesting to "little." This improvement in the understanding of ML was qualitatively confirmed by a pre- and post-workshop test aimed at assessing whether students could recognize ML systems in common products and services. The average of correct answers marked an already quite good 59% (before the workshop) and

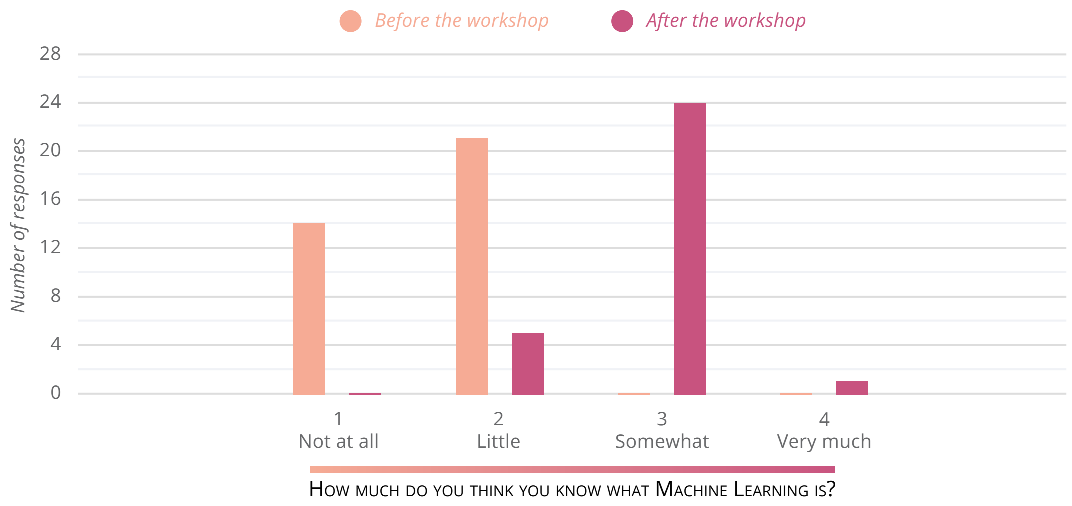


Fig. 5.21 | Self-assessment of ML knowledge before and after the workshop.

reached 75% afterward. The percentages for each question are depicted in Fig. 5.22. Interestingly, the recognition of photo editing features like brightness and contrast as non-ML systems passed from 94% of correct answers to 33%, probably biased by the explanation of ML generation systems related to this field of application. While the actual generation system to improve image resolution still was the least correctly identified.

Turning instead to the work that the participants developed during the workshop, the concepts were evaluated by the researcher on a 4-point scale (Not at all, Little, Somewhat, Very much) to understand if the solution would require the integration of ML and if the selected ML capability was consistent with its application (fourth and fifth columns of Tab. 5.2).

Reminding that for this activity the *ML Suitability Matrix* was not provided, not to add too much information, the coherent use of ML solely depended on the students' comprehension of the basic notions they were presented and the explanation of *ML Agents*. Positively, all the groups identified suitable possibilities for ML to contribute. Just three cases (#02 *The fear inductor*, #03 *What's this?*, and #08 *Drop 720*) could have been addressed effectively even without using ML. Still, they were correctly framed to exploit it.

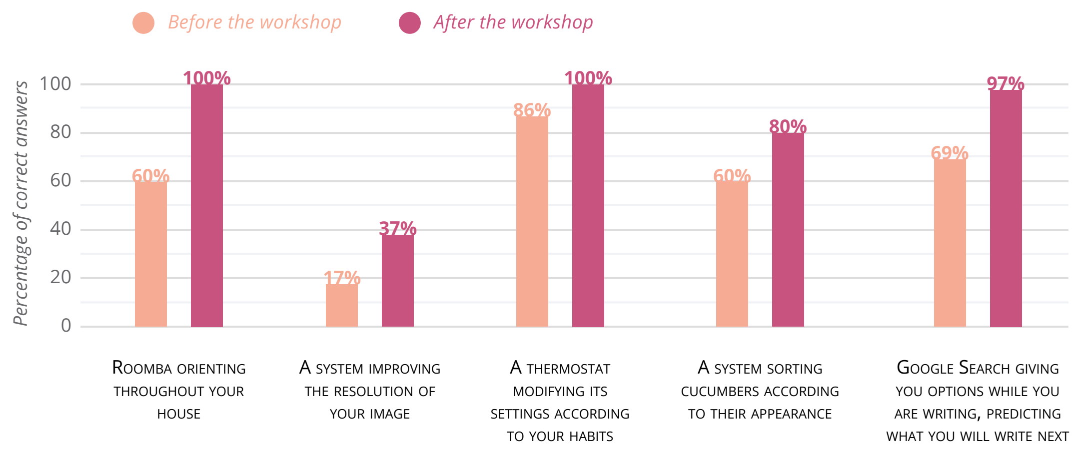


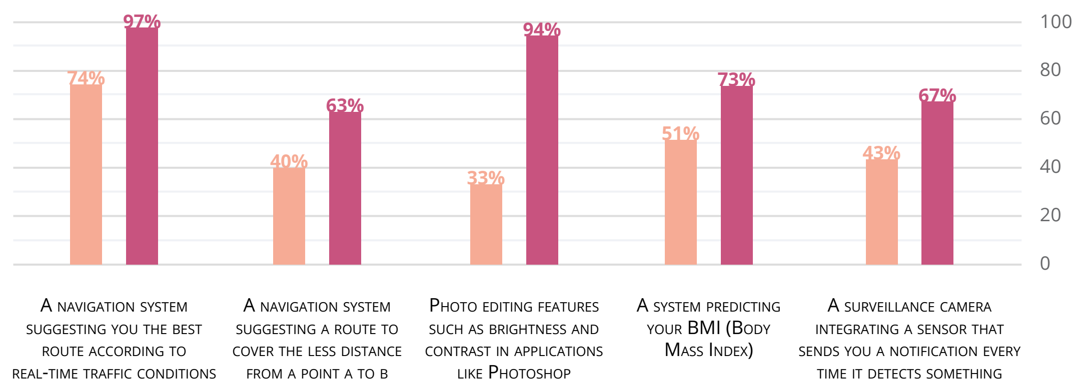
Fig. 5.22 | Percentage of correct answers to the ML identification test before and after the workshop.

The recognition of coherent ML capabilities to serve the students' ideas was equally satisfactory. Although all the choices are understandable, different *ML Agents* could have solved some cases. For instance, clustering would have been more convenient for #01 *Flashforward* and #07 *Insight*, as it is unclear how historical data could be retrieved. At the same time, a generative system could benefit #10 *Reverse invisibility*. However, how the concepts were developed and framed was not evaluated in the table, and some discrepancies between ML tasks and the outputs emerged. Indeed, even if ML capabilities were correctly applied, some groups did not properly frame the technical structure (input-task-output). As visible in Tab. 5.2 (in which the specifications of inputs and outputs cards are reported as the students wrote them on the boards), in some instances, instead of the outputs of the ML systems, those of the entire experiences were depicted. It is the case of #04 *Portable worlds*, where classification and generation agents transport visitors in different worlds, or #08 *Drop 720*, which output reports *"knowledge is acquired through active experience."* Though, these inconsistencies did not emerge during the reviews, when the groups presented their ideas to the researcher. This suggests that possibly the cause of the misinterpretation might be the recap box in the *concept specification space*, as it did not directly connect to the systems cards as intended.

5.3.2.3 ML integration in spatial exhibition experiences

Another factor to understand whether the contents were properly assimilated and if the approach was flexible enough to adapt to the educational context in which the workshop was hosted is to assess how ML was integrated with the spatial exhibition experiences. Indeed, ML applications should reflect how this technology can be a useful asset for designers without disrupting the flow, the objectives, or the principles of the design process.

During the second cycle of content analysis, four relevant themes were identified to define the effective application of ML to address the given brief. As depicted in columns 6 to 9 of Tab. 5.2, they point out how much ML impacted the overall concept, how the quality of the environmental UX and the typology of experience were affected, and whether the interface was present or dissolved, in light of the ubiquitous computing purposes that were introduced to inspire the experience development.



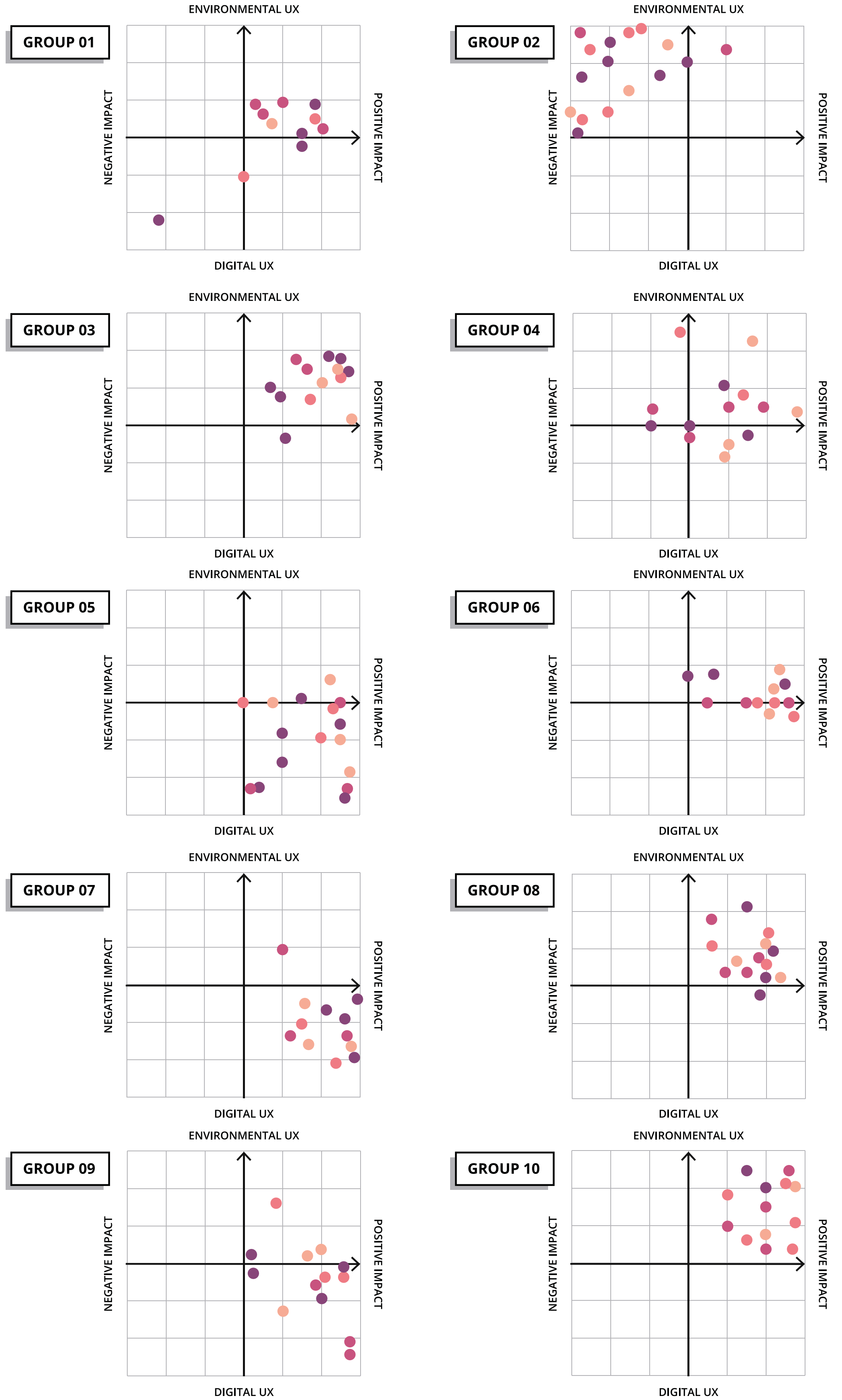


Fig. 5.23 | Peer evaluation matrices assessing the concepts according to their environmental or digital UX and positive or negative impact on people.

In relation to ML impact, the use of *ML Agents* could be distinguished into three main categories: (i) the experience is not in the physical realm – it is the case of #09 *Synderesis*; the experience has physical repercussions and (ii) the ML system is an indirect activator – that has been the most frequent option (6 cases); or (iii) ML plays a prominent role, by directly intervening in the environment, which happened in group #01, #03 and #10's concepts.

From a phenomenological perspective, all the concepts presented a highly immersive character. As visible in the peer-evaluation matrices (Fig. 5.23), the greatest part of the scenarios aimed at an environmental UX, with the disruptive exception of #09 *Synderesis* (in which the relevant sense of space is within the body itself). Instead, #05 *Metamorphosis* and #06 *Self vision* relapsed into mediated hybrid realities, similar to AR or VR. Interestingly, also #07 *Insight* was assessed closer to the digital realm, probably because of the dematerialized artifacts they presented in the physical space. Overall, the groups interpreted the UX in three main ways: by emphasizing or altering the sensorial perception (groups #05, #06, #08, #09, #10), maintaining natural interaction modalities in a changing environment (groups #01, #02, #03, #04, #07), or associating a cognitive load (groups #01 and #09).

As a result of the thematic analysis, the proposed experiences were categorized into two overarching typologies: spatial distortion and perceptive distortion. The former, the most recurrent, was implemented in the form of adaptive environments (groups #02, #03, #07, and secondarily #06), displacement (group #04 and #08), and environmental augmentation (group #01). The other was configured as a distortion of self-perception that could be physically perceivable (group #10), at a cognitive level (group #09), or a transfiguration (groups #05, #06, and #04 – slightly). Additionally, a sense of empowerment was fostered in the scenarios by groups #08 and #10.

Finally, it is interesting to understand how the interface was declined with respect to ubiquitous computing. Most groups (i.e., #01, #02, #03, #07, #09, #10) managed to make the interface disappear. While groups #04 and #08 fostered a spatial and even temporal displacement, they introduced invisible interfaces in the form of thresholds (a mirror in #04 *Portable world* and a display/portal in #08 *Drop 720*). Instead, the groups struggling with hybrid spaces inserted concrete filters, namely a suit in #05 *Metamorphosis* and visors in both *Metamorphosis* and #06 *Self vision*, even if they were transparent and unobtrusive from a first-person viewpoint.

So, once again, the young design students proved capable and at ease in facing an unusual challenge (compared to their educational background) and using a new material consistently. All the solutions responded to the brief, and ML did not cause the participants to be diverted from the objective of creating new experiences for entertainment in cultural institutions. The Sci-Fi perspective also helped students not get stuck on current ML solutions (primarily linked to the service dimension). However, it also made their imagination drift, risking losing focus on the spatial experience. Specifically, sometimes, from transposing digital functionalities to the physical world, they got to transfer physical perceptions to the virtual one and totally overlook design possibilities in the natural environment. For instance, in group #05's concept, the visitors not wearing the suit and visor were not considered to be part of the augmented experience or to interact with visitors impersonating fantastic creatures (e.g., they could be prey or hunters). This brought out an unexpected

disposition and interest in the sample students to design environments also in the virtual realm. As one noticed: *"It leaves a lot of freedom of expression, and the physical rules can be completely overturned* [so that] *spaces will give emotions that cannot be given with normal set-ups."*
Anyway, ML was always perceived as a tool in this process, not the end of the whole experience. Therefore, implementing the necessary expedients, CBB can be a flexible means for the translation, as it can support even not obvious design challenges. For didactic purposes the integration of ML was mandatory but, if combined with a tool to reflect on the appropriateness of implementing this technology, like the *ML Suitability Matrix*, this could become an actual possibility to consider for augmenting (or not) any kind of project.

5.3.2.4 Tools evaluation

The specific assessment of the tools further helped to shed light on how the ML translation was perceived. Participants were asked about it in the concluding questionnaire, which was answered by 30 out of 38 people, and in the discussion at the end of the workshop.
The *ML Agents* as knowledge transfer tools did not receive much feedback, but the majority of the students stated that they helped them very much in understanding ML capabilities. To complement the positive assessment, 9 would have preferred more exhaustive contents and 3 advocated for greater clarity (Fig. 5.24). Also the appropriate participation during the presentation of *ML Agents* and their consistent application in the projects reflected a good assimilation of the related contents. Moreover, comments from the participants revealed that the practical experimentation with the case studies also triggered reflections on the potential limits of ML systems, and they recommended spreading this kind of information to a larger public as it is not just a computer science domain.
The evaluation of the CBB was more articulated. The participants were asked if the tool helped them envision an experience of museum spaces that could integrate ML and be significant. To both questions, the responses were quite positive (Fig. 5.25). Only two people (for each question) found them of little use, and, in the first case, a third of the respondents asserted it was very helpful in generating a ML-infused solution. The 27 comments received about the CBB tool mostly highlighted its value because of the offered guidance and support in developing ideas. Also positively

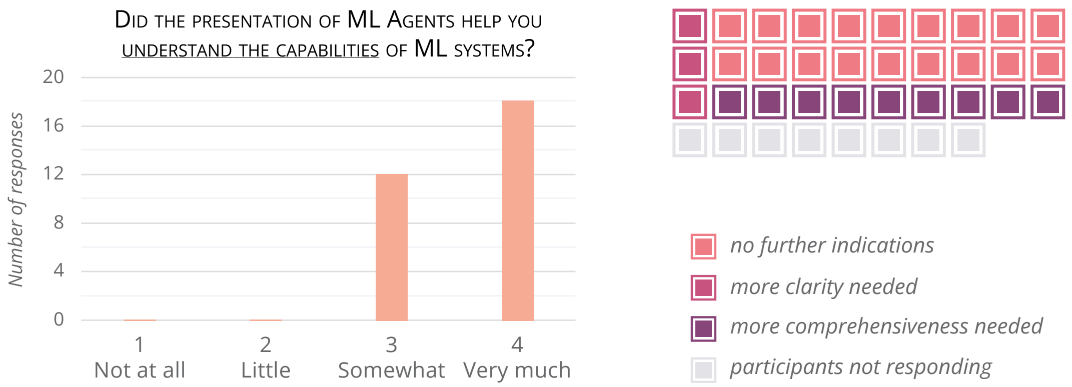


Fig. 5.24 | Students' evaluation of ML Agents.

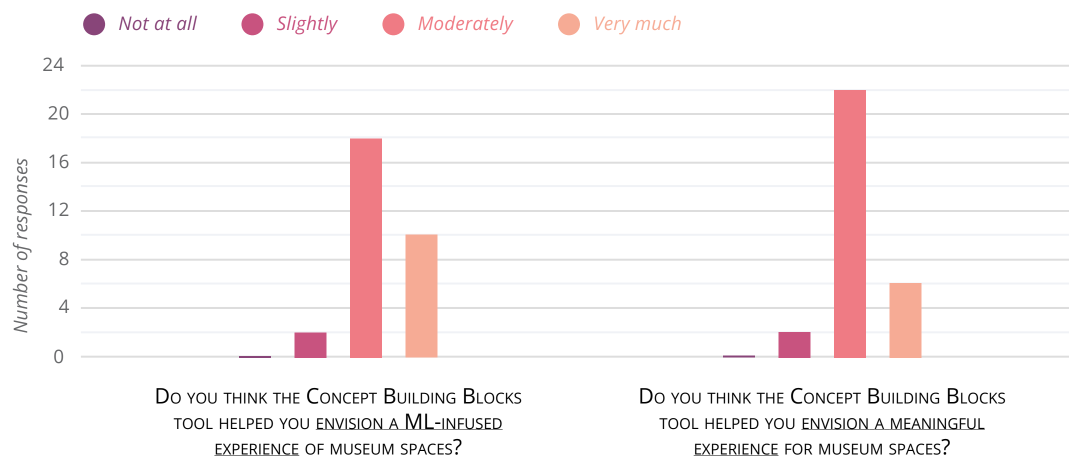


Fig. 5.25 | Students' evaluation of the CBB.

remarked was the fact that it enabled to create consistent connections and to reason "in blocks" as a way to simplify the process. Also appreciated were its capabilities to visualize, experiment, clarify, present, and create solidity in the process and stimulate effective and innovative ideas.
Greater customization and a different organization of the technical structure were identified as room for improvement. Indeed, both reflect two pain points. The first relates to the possibility of building systems integrating more than one capability, while the second underlines the high subjectivity of the mental process. In fact, others were more comfortable with a task–output–input structure as proposed.
Another relevant aspect in assessing the CBB tool was understanding how it was used. To avoid prescriptiveness, in fact, it was made clear that the path suggested was only a possibility and not compulsory. However, most respondents followed the design- and tech-driven approaches to build their idea (Fig. 5.26). From their feedback, the design-driven approach appeared linear and effective (each term recurring 5 times in the 16 responses), and it was also defined as exhaustive, intuitive, inspiring, captivating, helpful, and creative. Only one negative comment stated it left little freedom of thought.
Among those with the technology-driven process, no one arrived at a satisfactory solution right away. From the 10 dedicated comments, no prevalence of terminology emerged but the process was most commonly appreciated for its intuitiveness and

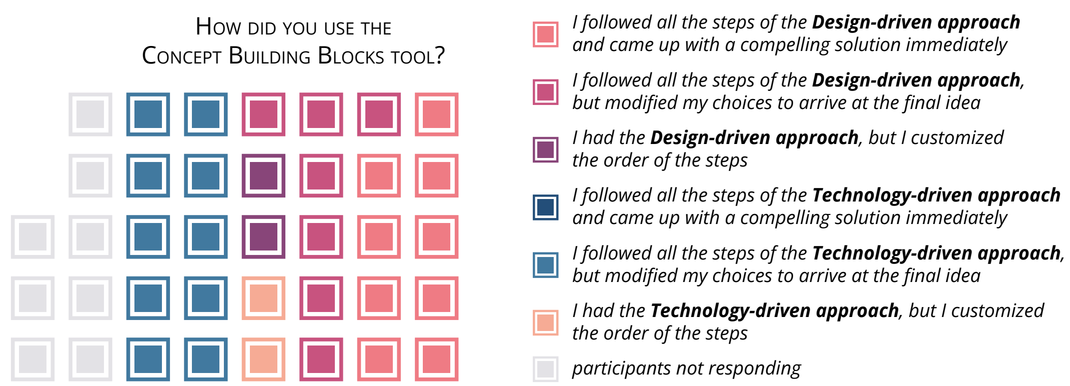


Fig. 5.26 | Report of use of the CBB.

immediacy, as well as for its effectiveness in guiding towards a solution. Interestingly, in this case, one feedback remarked how free it was.
Eventually, the design process was more commonly approached in an iterative way (Fig. 5.27), as the participants needed to modify their choices along the way. Only 7 stated it was completely linear, and 3 used it randomly to get inspiration.

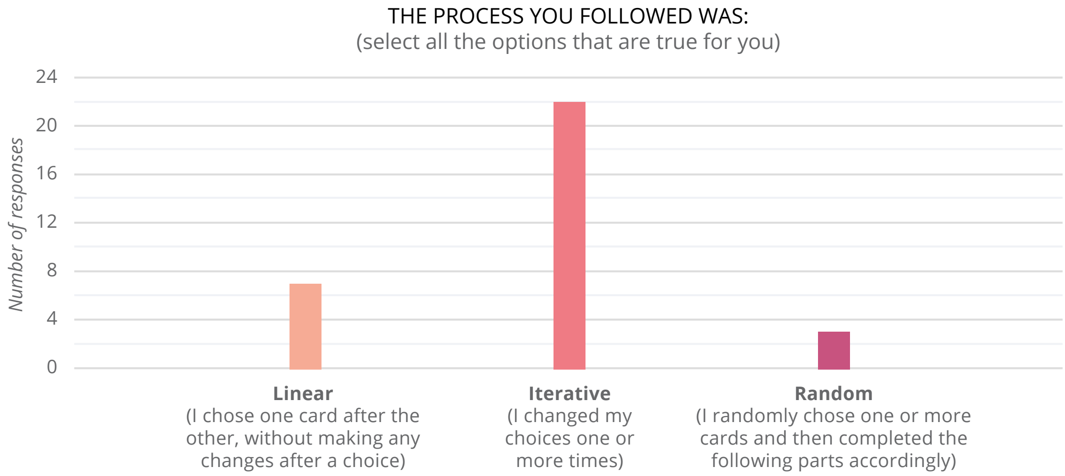


Fig. 5.27 | Assessed characteristics in the CBB use path.

5.3.2.5 Ethical concerns

Even though the generated concepts successfully outlined new kinds of ML-infused experiences in exhibition spaces, more or less focusing on the physical dimension, all the ideas had, in different measures, some value for people. The researcher tried to evaluate at what cost they were reached. In fact, albeit very briefly, the students were introduced to basic principles of responsible design and, at the end of their SFPs, they were asked to ponder the possible implications of their ideas. However, this part fell by the wayside and many groups did not develop it. Also the positive and negative connotation of the impact in the peer-evaluation matrix (Fig. 5.23) was not useful in stimulating critical thinking (suggesting that a low or high level of value added to people's experiences could have been a less misleading definition).
Significantly, according to the researcher, half of the groups presented applications whose value was not really commensurate to the cost that the implementation would require of users (Tab. 5.2, column 10). Indeed, a thought-provoking remark comes from the observed peculiar correlation between the disappearance of the interface and an apparent user's loss of agency. Except for #05 *Metamorphosis* and #10 *Reverse invisibility*, all the scenarios envisioned some level of personal and even intimate data giveaway. #01 *Flashforward* and #09 *Synderesis* even brought this sharing to a totalizing amount, including all past history, preferences and even conscious and unconscious attitudes as input for ML systems (Tab. 5.2, column 11).
Because the ethical responsibility of the solutions was not the central topic of the workshop, this matter was not discussed with the participants. Then, no precise idea about the rationale behind these choices could be inferred. Nonetheless, it is a clear indication that making reflections about values and implications explicit during the design process is necessary, as suggested by van den Hoven (2013). It is especially true with students in earlier stages of their design education, as doctoral students naturally manifested a deeper level of ethics-related reasoning.

5.3.2.6 Overall workshop experience and participants' perception evolution

Although the workshop's subject matter might be unconventional for the target audience and the result of the experimentation was unpredictable, attention and curiosity were perceived from the beginning during the introductory presentations. Then, all seemed very involved in the concept development activity. Only minor doubts emerged while getting started with the CBB tool, and, in the end, no negative remarks arose towards the workshop topic and experience. On the contrary, *interesting*, *stimulating*, and *fun* were the most recurrent words in the comments, and some students underlined the importance and usefulness of technology-related skills for interior designers.
Indeed, the participants' attitude towards the subject appeared very positive already from the pre-workshop questionnaire (Fig. 5.28), which received 35 answers. The majority believed ML to be a tool (74.3%) or a great opportunity (26.7%), which is already everywhere in our lives (68.6%). Moreover, despite a varied perception of the current role of technologies in the physical world, the majority envisioned a future where the digital dimension supports and improves life in the real one (77.1%), and only one claims that it should not exist (2.9%).
Curiously, after the workshop, the number of students sustaining that ML is everywhere decreased from 24 to 14 (on a total of 30 responses instead of the 35 of the first questionnaire) in favor of more people believing that it is concealed in advanced systems or not really diffused, answer that no one selected before. A possible explanation might be that, after understanding the real capabilities of

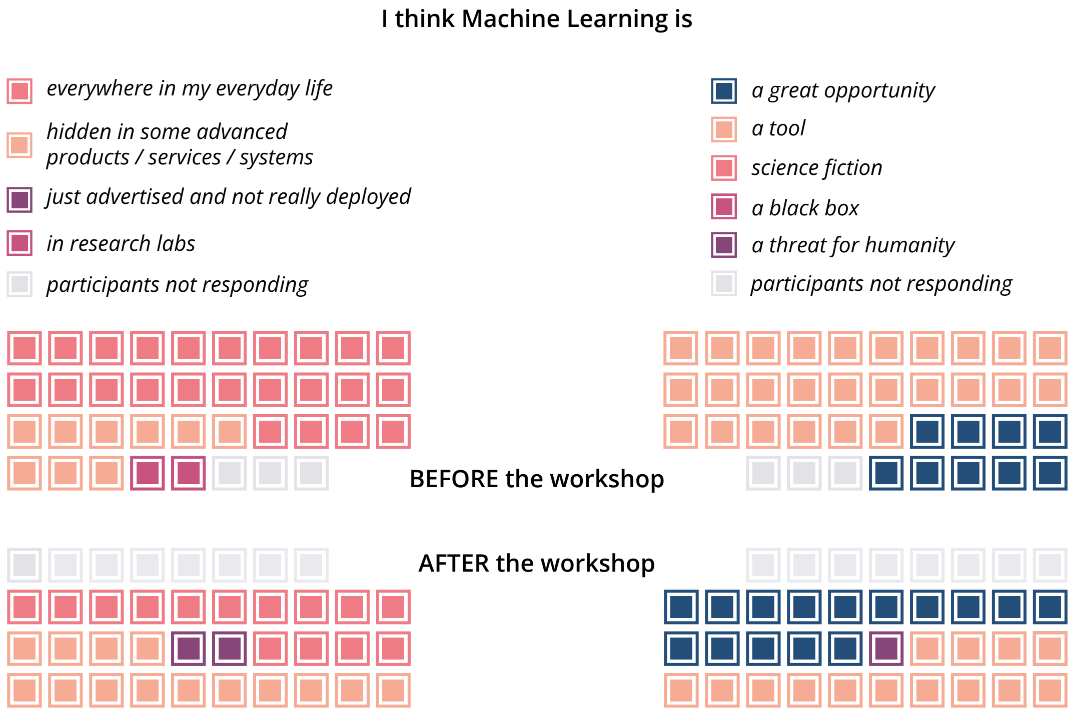


Fig. 5.28 | Participants' perception of ML before and after the workshop.

ML, some wanted to underline that the actual potential of this technology is still unexpressed, as supported by the increase (from 9 to 15) of the respondents who believe ML is a great opportunity. Another possibility is that some unintended messages were inferred from the experimental translation, as the answer marking ML as a threat to humanity appearing only in the second questionnaire may suggest.
Though, the predominant positive sentiment towards ML and especially its relevance for interior design students exceeded expectations. In response to whether they would like to learn to work with ML as a tool for their future design career, several answered that it would be advantageous to integrate it into any kind of project. Also, those not enthusiastic about technology recognized the opportunities ML might offer in the future. A comment finely synthesized the most common position: *"Having an awareness of the use of these systems, at the design level, is certainly a strength, and it would be most useful to be thoroughly acquainted with their full potential. Of course, however, they should remain accessories and aids to the project, not totalizing systems."*
Finally, about the relationship between technology and interior design education, the entire class was unexpectedly favorable to introducing the digital dimension in their projects and formative path. One stated that there is an inevitable *"progression towards an integration of physical and digital,"* and a couple of girls highlighted how this duality characterizes their generation. The climax was reached with the counter-question: *"Non-use* [of technology in interiors] *is limiting, so why NOT use these systems?"*

5.3.3 Discussion

In light of the students' closing remarks after the two-day workshop on ML to enhance spatial experiences, a strong interest in deepening technology-related understanding and skills was evident, even if their educational course usually does not involve that kind of topic. Reinforcing Meyer and Norman's (2020) assumption that design education, in general, should provide appropriate knowledge of technology, it was highlighted that the next generations of designers are born in the digital era. Therefore, it has always been part of their everyday life and keeping it separated from their educational path, thus, out of their control, would be anachronistic.
This is why, beyond the researchers' initial intents, the translation of **ML knowledge needs to be accessible to a broad audience of design students,** also if it does not represent a typical material of their specialization. In this regard, the **flexibility and modularity of the approach and the developed tools gain even greater importance, as they need to be able to complement several design contexts and challenges.**
The experimentation presented above only focused on a soft introduction of ML basic knowledge and capabilities, framed in the realm of ubiquitous computing and pervasive interaction to better suit the scope of interior design and using SFPs as a form of materialization to adapt to the capabilities of the students participating in the activity. The responsibility requirement to achieve meaningful ML applications remained in the background because of the limited time and not to overwhelm the participants with too much information – leaving more space to learn in practice. The result was that the envisioned solutions were generally aligned with the technology and the brief but raised some concerns from an ethical point of view. Because of this, the next steps of the research should focus on more value-sensitive perspectives.

5.4 Reflections on the educational experiences

Towards the construction of an educational method to test the research outputs presented so far, some deeper reflections on the educational experiences might be useful to pinpoint the essential traits and requirements it should have.
Undoubtedly, the cross-disciplinary translation of ML knowledge to enable a consistent and responsible design of ML-infused solutions can be described as a complex learning task, as it involves the *"integration of knowledge, skills, and attitudes"* (Kirschner & Van Merriënboer, 2009) that learners should acquire. Interdisciplinarity is indeed one of the premises of the research, which aspires to merge ML, ethics, and design knowledge in practical and experimental environments. A **holistic approach** was found to be a requirement for dealing with the complexity of this integration. Blending into a multifaceted project, the technical and ethical concerns can be encompassed in a systemic design perspective, in which not just a product or service but an ecosystem of artifacts and the people that revolve around it need to be taken into consideration to shape any design decision. With the proper support to explicitly trigger ethical reflections, thoughtful and responsible solutions can emerge.
Moreover, to handle the differences between design disciplinary contexts, students' skills and maturity, time at disposal, or educational purposes, two further requisites have been identified: **flexibility and modularity.** The contents addressed, the tools, the format, and the language used should support the educational method in being easily adaptable to the specificity of the situation. As Annik (2006) affirms, *"[...] design is fluid. And the teaching of a fluid thing must also be fluid."* Thus, *"it is nonsense to look for fixed formulas, write them down, frame them, and set off based on these formulas toward a particular goal"* [translation from Italian by the author].
In any of the three experimentations previously described, the tone and the level of the challenges were crafted to fit with the students' predispositions and settings. But the definition of the characteristics the educational activities should portray cannot always be defined beforehand. The educational designer might need some contextual information to adapt the didactic experience and must be able to **play with the educational models and tools to encounter their needs**.
Indeed, also the kind of experience that these convey is part of the delineation of an educational method. The sensitive tailoring of educational activities should be intended to create a **familiar environment for learners**, leveraging their domain knowledge and helping them build confidence in dealing with an unfamiliar topic (D'Ignazio, 2022). In this sense, inserting ML capabilities in a human-centered design process and a studio setting facilitates students' creative experimentation and learning, making ML part of something they are used to. In fact, paraphrasing Shirky's (2008) words, profound changes happen when technology becomes normal, which, in the case of designers, means that it becomes a tool they are comfortable working with.
The **informality and openness** of the tasks are further characterizing and effective traits in creating constructive conditions for the recipients, allowing design students to explore the paths they prefer.
Not to compromise the learning process, the **immediacy of the tools** is equally relevant. They need to be effortless and enjoyable to deal with. For this, **they should**

encounter designers' visual and practical requirements. Moreover, their **physical dimension** dramatically improves the smoothness of the experience, as it naturally **favors communication and collaboration** within a design team. As boundary tools to play with, move around, and support decision-making, they inherently encourage discussion between design students and activate a productive reflective environment. On the contrary, digitally mediated interaction can present technical inconveniences. Eventually, regardless of the challenging nature of translating and merging ML knowledge with ethics and design, the focus in shaping an educational method should **concentrate on the needs and predisposition of design students to foster their personal understanding and sense-making.**

TO SUM UP

The chapter addresses RQ3: ***How can theoretical constructs be operationalized into models and tools to be implemented and tested in educational contexts?***

- As a first step in the operationalization, an *Introductory Game to ML Responsible Design* was developed to assess how design students might respond to a multi-layered ML design knowledge that includes elements from all three disciplines from the earliest phase of a simulated design process. The experimentation provided relevant insights:
  - The holistic approach was effortlessly assimilated by the testers, who comfortably navigated across disciplines, confirming that design educational background is functional to deal with complex systemic issues.
  - Enough agency and freedom are essential for learners to explore the contents, form their own idea on the subject matter, and discuss about it to make decisions. Thus, tools should be precise and straightforward. They should engage, support reflection, facilitate decision-making, and expand design thinking without prescribing design actions.
- Accordingly, the *Concept Building Blocks* (CBB) tool (one of the main outputs of the research) was constructed to be flexible to adapt to different contexts and design purposes, and modular to break it down into parts conveying specific pieces of knowledge independently.
  - Its primary function is to guide design students to envision consistent ML applications, giving hints about how the elements of the theoretical construct combine with each other but leaving freedom for interpretation and practical implementation.
  - It can be approached in multiple ways (*Problem-based, Data-driven, Technology-driven, Artifact-driven, Value-based, or Human-driven*) and includes:
    - > a *concept framing space* to define the aim, context, and target audience of an idea (design A.C.T.);
    - > a *concept exploration space* (reflecting the *ML Designerly Taxonomy*) to help build the structure of a concept through a system of matching cards;
    - > and a *concept definition space* to specify and personalize it.
- A *VALUable by Design Expansion* (VDE) is added to integrate a responsible and value-driven approach to the design of ML systems and to synthesize the insights elaborated in section 4.4.
  - It adds the United Nations' Sustainable Development Goals (*SDGs*) and *principles* to the *concept framing space* to drive the concept generation.
  - It finally provides *concept specification* boards including *Reality Check* and *Value* cards to induce higher levels of reflection and encourage responsible additions to the concept.
- The *Superpowered Museums* workshop was designed to test the flexibility and modularity of the envisioned approach and tools, particularly the CBB. Hence, it was set in a challenging predetermined design context (third year of the bachelor design program in Interior Design) and aimed at augmenting the quality of the physical experience of exhibition spaces. It produced interesting findings:
  - The contents and tools to introduce ML were complemented by basic knowledge about ubiquitous computing to facilitate the connection between technology and interior spaces, and Science Fiction Prototypes (SFPs) as means for students to put theory into practice, without worrying about the feasibility of the solutions.

- Students with no previous knowledge of ML and not used to dealing with technological solutions had fun and were able to envision ideas generally consistent with the technology and the brief but raised some concerns from the ethical point of view.
- The application of ML reflected how this technology can be an asset for designers without disrupting the flow, the objectives, or the principles of the design process.
- *ML Agents* helped students understand ML capabilities, while CBBs were effectively used to visualize, experiment, clarify, present, and stimulate consistent and innovative ideas.
- The translation of ML knowledge can be accessible to a broad audience of design students, even if it does not represent a typical material of their sub-discipline. In this regard, the flexibility and modularity of the approach gain even greater importance to complement several design contexts and challenges.

These findings imply that educational method should concentrate on the needs and predisposition of design students to foster their personal understanding and sense-making. It should leverage their domain knowledge, helping them to build confidence in dealing with an unfamiliar topic, and leave agency to the educational designer to adapt the educational experiences (modules and tools) to different contexts.

6

6. TOWARDS AN EDUCATIONAL METHOD TO FRAME MODELS AND TOOLS

A sound education provides the Intelligence Amplification (IA) that balances the Artificial Intelligence (AI) that now surrounds us. In our time, design is the pathway of IA.

(Buchanan in Frascara, 2020)

C H A P T E R S I X

Previous experiments highlighted that active didactic experiences with procedural information in a setting where design students have enough agency to explore and make decisions are beneficial characteristics for translating ML and related ethics knowledge in the educational context. To **better frame and test the most appropriate modalities**, the selection of a constructive and project-based pedagogical method in conjunction with mixed-methods evaluation research is discussed (6.1).

Taking advantage of the convergence of formative and research objectives, some educational models were developed to validate the insights from previous experiences and the developed tools. Specifically, they were built on the theoretical assumption of the two main requirements for a meaningful ML-infused solution. Thus, a *consistency* and a *responsibility model* were envisioned and tested independently and subsequently combined to create an integrated model (6.2).

Always maintaining two separate branches for research purposes, all the models were implemented in workshops held in different European universities (namely the École de Design Nantes Atlantique, FH Joanneum University in Graz, Universidade da Madeira, and Politecnico di Milano) targeting students enrolled at different levels of multiple design specializations. The methods, unfolding, and results of the workshops, considered case studies, are discussed in 6.3. In particular, the assessment focuses on compliance with the ILOs, the tools, and the students' overall experience and knowledge acquisition.

Finally, the iterative research process and findings are synthesized in the framing of an educational method to convey ML and ethics knowledge to design students (6.4). Being the ultimate research output, it sets the premises for further investigations.

6

6.1 Identifying the didactic method and the research protocol

6.1.1 A constructive and project-based method

Having gathered valuable insights about the most promising strategies and character for educational activities to translate and transfer basic ML knowledge to design students, a more formal definition of the educational method used to support and validate the research outputs is required.
Following consolidated approaches to the design of educational experiences (Sancassani et al., 2019) and tackling it as any other design problem, the focus is on the recipients. Therefore, according to Biggs's constructive alignment (Biggs, 2003), there are three main phases to designing effective educational activities. First, one should set **(i) the objectives**, the intended learning outcomes (ILOs) that students should reach. To those, **(ii) the evaluation processes** and **(iii) teaching-learning activities** should be aligned.
Hence, in support of a more structured method, the ILOs from the previous experiments have been reframed and enriched on the basis of Dublin descriptors and the revised Bloom's Taxonomy (Krathwohl, 2002). Depicting the cognitive dimensions and the expected student actions, they facilitate the precise identification of the modalities in which the achievement of the ILOs can be assessed and the necessarily related teaching-learning activities.
Based on the foundational research assumptions, the ILOs are differentiated according to the requirement to reach meaningful ML-infused solutions they respond to, *consistency* (Fig. 6.1), and *responsibility* (Fig. 6.2). Four synthetic ILOs are depicted for each strand, and they are more finely articulated into knowledge, skills, and values-related learning outcomes. Each is associated with a Dublin descriptor (in pink) and a Bloom's learning typology (in blue). These specifications allow the identification of suitable assessment processes.
Although formal evaluation is not the purpose, this process enables the researcher to measure how the educational method is performing and the students to have feedback reinforcing their understanding. In particular, quick, informal, structured, and semi-structured formative tests (e.g., through multiple choice or brief open questions) are helpful to appraise capabilities in the domain of *remembering, understanding*, and (partially) *applying*. They can assist students to check their comprehension and the teacher to reinforce some key concepts. In the case of

SHORT-TERM RESEARCH OUTCOMES		Enable design students to: imagine **CONSISTENT** design-driven ML-infused solutions			
ILOs		Students should be able to:			
		Understand the core characteristics of ML systems (*knowledge & understanding* | ***understand, remember***)	**Understand ML capabilities** (***knowledge & understanding, apply*** | ***understand, remember, apply, analyze***)	**Identify what problems can be solved with ML systems** (***applying knowledge and understanding, making judgments*** | ***apply, evaluate***)	**Generate relevant, consistent, and effective design concepts including ML systems** (***applying knowledge and understanding*** | ***apply, create***)
	Knowledge	*Define the main characteristics of ML systems (knowledge & understanding | remember)*	*Remember the capabilities of ML systems (knowledge & understanding | remember)*		
	Knowledge	*Understand the differences between ML systems and traditional programs (knowledge & understanding | understand)*	*Identify the main capabilities of ML systems in the current scenario (knowledge & understanding | understand)*		
	Skills		*Recognize possible applications of ML (agents), given the description of the functionality of existing products and services (knowledge & understanding | understand)*	*Identify potentialities and relevant problems for ML systems based on their current capabilities (apply knowledge & understanding | apply)*	*Envision useful and consistent applications of ML systems to solve relevant problems and outline their input, task and output (applying knowledge and understanding | apply, create)*
	Skills		*Choose consistent ML agents to solve a given problem (applying knowledge and understanding | apply)*	*Infer whether ML is a proper tool to solve a given problem, considering its capabilities (making judgments | evaluate)*	*Design ML systems applications based on their capabilities (applying knowledge and understanding | apply, create)*
	Values	*Support an entry-level discussion with peers and ML experts (communication skills | apply)*			
	Values	*Recognize ML as a manageable asset for design (knowledge & understanding | understand)*			
	Values	*Maintain a design-driven approach to solve problems with ML (applying knowledge and understanding | apply)*			
EVALUATION PROCESSES (for didactic & research)	**During TLAs**	**Formative tests**			**Collective presentation and peer evaluation**
	During TLAs	**Observation & mentoring**			
	After TLAs	**Questionnaire + oral & written feedback**			
TOOLS (res. outputs)		(presentation)	ML Agents	ML Suitability Matrix	Concept Building Blocks
EXPECTED OUTPUT					**ML Hero concept presentation**
CONTENTS		Intro to ML: *definitions* (what it is) and *demystification* (what it is not)	*ML capabilities* (tasks, case studies and functioning of ML systems)	Procedural info	
		Supportive info			
		Familiarity level of mastery			
LEARNING MODULE STRUCTURE		**2-day workshop** (DAY 1: 1-6; DAY 2: 7-8)			
		1|Getting in touch with ML	2|Training - step 1 (recognizing) 3|Training - step 2 (identifying)	4|Training - step 3 (inferring ML suitability) 5|Groups' problem framing	6|Groups' concept generation 7|Groups' representation of the Hero 8|Groups' presentation and peer evaluation activity
PEDAGOGICAL FRAMEWORK		**Constructivist approach** - Problem/Project-based learning within a structure inspired to Gagné Events of Instruction			

Fig. 6.1 | Theoretical framework underlying the consistency model.

SHORT-TERM RESEARCH OUTCOM		Enable design students to: imagine **RESPONSIBLE** design-driven ML-infused solutions			
ILOs		Students should be able to:			
		Understand ML systems as socio-technical systems and their capabilities (***knowledge & understanding, apply*** | ***understand, remember, apply, analyze***)	**Identify and use values to drive the design of ML systems** (***knowledge & understanding, apply*** | ***understand, apply***)	**Identify and anticipate ML possible impacts systems in practical, personal, social, cultural, and eco-systemic dimensions** (***applying knowledge and understanding*** | ***apply***)	**Generate ethically acceptable, sustainable, and desirable design concepts including ML systems** (***applying knowledge and understanding*** | ***apply, create***)
	Knowledge	*Identify the main capabilities of ML systems in the current scenario (knowledge & understanding | understand)*	*Identify the values that should drive the design of ML systems to achieve meaningful and responsible solutions, starting from technical, designerly and ethical perspectives (knowledge & understanding | understand)*	*Understand the implications of ML systems in their practical, personal, social, cultural and eco-systemic dimensions (knowledge & understanding | understand)*	
	Knowledge	*Understand and describe the role of humans in the development process of ML systems (knowledge & understanding | understand)*			
	Skills	*Understand the - technical, ethical and UX - limitations of ML systems in relation to the tasks they are meant to solve (knowledge & understanding | understand)*		*Identify and anticipate undesirable outcomes of ML systems given their characteristics and context of use (applying knowledge and understanding | apply)*	*Identify possible solutions to avoid or counteract undesirable outcomes of ML agents in pursuit of pre-defined values (applying knowledge and understanding | apply)*
	Skills			*Identify strengths and flaws of existing and imagined ML systems, based on their properties (input, task, output, context, limitations, potentialities...) by anticipating their impact (making judgments | evaluate)*	*Activate a value-driven design process to the development of solutions that include ML systems (applying knowledge and understanding | apply)*
	Skills				*Envision responsible and meaningful applications of ML systems to solve relevant problems and outline their input, task and output (applying knowledge and understanding | apply, create)*
	Values	*Support an entry-level discussion with peers and ML experts (communication skills | apply)*			
	Values	*Design with a responsible and value-driven approach to anticipate undesirable outcomes in the concept generation phase (applying knowledge and understanding | apply)*			
EVALUATION PROCESSES (for didactic & research)	**During TLAs**	**Formative test**	**/**		**Self-assessment, collective presentation and peer evaluation**
	During TLAs	**Observation & mentoring**			
	After TLAs	**Questionnaire + oral & written feedback**			
TOOLS (res. output		(presentation) + ML Agents	Concept Building Blocks + VALUable by Design exapansion		
EXPECTED OUTPUT			**ML Hero concept presentation**		
CONTENTS		Intro to *Responsible Research Innovation* (RRI), *value sensitive design*, ML as a *socio-technical system* and its *capabilities*	Procedural info		
		Supportive info			
		Familiarity level of mastery			
LEARNING MODULE STRUCTURE		4hour workshop			
		1|Getting in touch with responsible design 2|Getting in touch with ML	3|Groups' responsible and value-driven concept generation 4|Groups' presentation and peer evaluation activity		
PEDAGOGICAL FRAMEWORK		**Constructivist approach** - Problem/Project-based learning within a structure inspired to Gagné Events of Instruction			

Fig. 6.2 | Theoretical framework underlying the responsibility model.

outcomes that require more critical reflection-in-action (Schön, 1983), instead, project development is the most natural learning environment for design students. In this regard, collective presentations, self, and peer evaluation on predefined, significative parameters contribute to a more profound and lasting acquisition of knowledge, skills, and competencies, including effective communication. Observation and mentoring complement these activities.

Subsequently, teaching-learning activities are shaped according to pedagogical theories and models, or the educational designer's philosophical perspective on knowledge and learning, as Sancassani (2019) puts it. The envisioned educational method embraces a constructivist approach, as the previous experiences and the research itself testify. It is funded on the assumption that knowledge results from a construction of meaning that each person makes based on the internal elaboration of feelings, prior knowledge, value systems, and beliefs, which is fundamental for designers to foster innovation. In this sense, education recursively involves students in processes based on experiences, abstractions, inferences, problem-solving, information recombination, and peer collaboration (Sancassani et al., 2019).

The specific references to build the educational models for the conclusive validation of the research outputs (better introduced in the next section) are Gagné's *events of instruction* and *problem-based learning*. The former involves nine steps. From an initial (1) engagement phase aimed to gain the students' attention, it (2) informs them of the objectives (ILOs) before (3) recalling prior knowledge. Then, it includes (4) presenting useful contents to reach the ILOs, (5) providing guidance (e.g., through the facilitation of the activity or tools), (6) eliciting performance, (7) providing feedback, (8) assessing performance, and (9) enhancing retention and transfer. Furthermore, a problem-based or, better, *project-centered approach* informs the events 4 to 6 from the previous framework, depicting a common studio format. Indeed, in Kirschener and Norman's (2021) perspective, a project-centered approach entails a wider scope than just solving a problem, as it includes *"social, societal, economic, ethical, ecological aspects and so further of that solution"* (Kirschener & Norman, 2021). Moreover, this is what the authors identify as a critical point to emphasize the role of design as part of a whole system, no matter how little it is.

In this setting, a more exploratory and hands-on character of the educational activities would have been pursued, but it was not possible to develop and test them within the frame of the research.

6.1.2 Evaluation research to assess the educational method

Intrinsically, the proposed educational method is both an output of the study and the basis for its assessment, framed as evaluation research (Robson & McCartan, 2015).

As a comparative study (the most common format for this research method) was not feasible due to inherent difficulties in creating the conditions for a consistent control group, the investigation has been envisioned as a **multiple case studies analysis**. A series of four workshops, based on the educational models and with little differences between one and the other, were held in different European universities and were used as case studies. Because of the richness of retrievable information, a mixed-method research strategy was adopted.

The evaluation research aimed at **(i) testing the effectiveness of the educational method** in terms of the appropriateness of contents (typology, structure, format), tools, and teaching-learning activities; **(ii) understanding how the didactic models unfolded in practice** (which results were produced, whether some unexpected techniques to reach the goal emerged, or something was missing); and **(iii) identifying spaces for improvement** of the elements listed before with the addition of time.

As the research and educational objectives are strongly intertwined, as portrayed in Fig. 6.1 and 6.2, the parameters and evaluation activities conceived for the didactic models overlap with the research requirements and offer relevant material for assessment and validation. For the same reason, the two main components of the educational method (*consistency* and *responsibility*) have been tested separately, providing more flexibility to the evaluation process.

To collect data, different techniques have been implemented. Specifically, (i) the researcher's observation, supported by a dedicated sheet, aimed at gathering information about students' responses to teaching-learning activities, contents, and tools (focusing on doubts and difficulties, consistent and unexpected uses, and possible shortcomings). (ii) Formative tests, peer, and self-evaluations were additional sources of possibly quantitative data to be compared with the researcher's or their own perceptions as expressed in (iii) pre- and post-workshop questionnaires. Similarly to the previous workshops, the former was intended to portray an initial assessment of students' ML knowledge and disposition. The latter explicitly asked for evaluations of the workshop and its tools (in general and according to the declared ILOs). Finally, (iv) students' delivery and presentation of ML-infused concepts were material for qualitative inferences. In particular, all the oral and written feedback and productions were subject to a content analysis organized with *attribute coding* to record the fieldwork settings, *structural coding* to highlight the specific topic of inquiry (e.g., ILOs, tools), *descriptive coding* to infer relevant issues, and *in vivo coding* to pop out interesting quotes (Saldaña, 2009). Affinity maps were ultimately generated to cluster and make sense of the codes.

For the analysis, a strategy halfway between *convergent* and *exploratory* (Creswell, 2014) has been assumed, meaning that qualitative and quantitative data are collected concurrently for subsequent triangulation. Still, the former is emphasized in accordance with the prevalence of qualitative sources of information.

In terms of limitations, reliability, and validity, the study shares most of the considerations stated in section 5.3.1.4 for the *Superpowered Museums* workshop, with some differences outlined in the following. It tries to improve the quality of the sample and to overcome the lack of diversity by replicating the format in different universities, with students enrolled in different programs and at different levels in their educational paths. Though, in most cases, the number of participants is still limited.

To increase the reliability and validity of the qualitative investigation, the research protocol has been further detailed and carefully structured beforehand to maintain homogeneity throughout all the workshops and get consistent data.

ML Hero Agency

Contextual information

WHAT	Two-day workshop (16 hours).
WHEN	10-11 March 2022.
WHERE	École de Design Nantes Atlantique.
WHO	The author.
STUDENTS INVOLVED	18 international first-year students enrolled in the Digital Design MDes (working in groups of 3-4 people). Participation was mandatory for them.

Structure of the educational activity

ILOs	***Knowledge***	• Define the main characteristics of ML systems. • Understand the differences between ML systems and traditional programs. • Remember the capabilities of ML systems. • Identify the main capabilities of ML systems in the current scenario
	Skills	• Recognize possible applications of ML (agents), given the description of the functionality of existing products and services. • Choose consistent ML agents to solve a given problem. • Identify potentialities and relevant problems for ML systems based on their current capabilities. • Infer whether ML is a proper tool to solve a given problem, considering its capabilities. • Envision useful and consistent applications of ML systems to solve relevant problems and outline their input, task and output. • Design ML systems applications based on their capabilities.
	Values	• Support an entry-level discussion with peers and ML experts. • Recognize ML as a manageable asset for design. • Maintain a design-driven approach when using ML systems to solve problems.
EXPECTED IMPACT(S)		• Build basic awareness on designerly modalities to approach the conceptualization of ML systems toward consistent solutions. • Instill curiosity towards a still unusual tool for design.
CONTENTS		• ML definition • ML demystification • ML capabilities • Procedural information
TOOLS	***Knowledge transfer***	• Theoretical introduction • ML Agents • ML Suitability Matrix
	Design activities	• CBB
OUTPUT		**ML Hero concept**: *core structure*, (CBB) + *visual representation* synthesizing its superpower and meaning) + *storyboard* (portraying the UX)

VALUable ML Heroes

Contextual information

WHAT	Half-day workshop (3 hours).
WHEN	9 March 2022.
WHERE	École de Design Nantes Atlantique.
WHO	The author.
STUDENTS INVOLVED	10 first-year MDes UX design students (working in groups of 2-3). Participation was mandatory for them.

Structure of the educational activity

ILOs	***Knowledge***	• Identify the main capabilities of ML systems in the current scenario. • Understand and describe the role of humans in the development process of ML systems. • Identify the values that should drive the design of ML systems to achieve meaningful and responsible solutions, starting from technical, designerly and ethical perspectives. • Understand the implications of ML systems in their practical, personal, social, cultural and eco-systemic dimensions.
	Skills	• Understand the - technical, ethical and UX - limitations of ML systems in relation to the tasks they are meant to solve. • Identify and anticipate undesirable outcomes of ML systems given their characteristics and context of use. • Identify strengths and flaws of existing and imagined ML systems, based on their properties (input, task, output, context, limitations, potentialities...) by anticipating their impact. • Identify possible solutions to avoid or counteract undesirable outcomes of ML agents in pursuit of pre-defined values. • Activate a value-driven design process to the development of solutions that include ML systems. • Envision responsible and meaningful applications of ML systems to solve relevant problems and outline their input, task and output.
	Values	• Support an entry-level discussion with peers and ML experts. • Design with a responsible and value-driven approach to anticipate undesirable outcomes in the concept generation phase.
EXPECTED IMPACT(S)		• Build basic awareness on designerly and value-driven modalities to approach the conceptualization of ML systems toward consistent solutions. • Highlight the importance of early reasoning about values and impacts. • Recognize ML agents as design assets to reach meaningful goals.
CONTENTS		• Intro to Responsible Research Innovation (RRI) • Value-sensitive design • ML as a socio-technical system • ML capabilities • Procedural information
TOOLS	***Knowledge transfer***	• Theoretical introduction • ML Agents
	Design activities	• CBB • VDE
OUTPUT		**ML Hero concept** (CBB and VDE completion)

VALUable ML Hero Agency

Contextual information

WHAT	3 days (18 hours) | 3 days (16 hours)
WHEN	10-12 May 2022 | 21, 24, 25 October 2022
WHERE	FH Joanneum University of Applied Science, in Graz (Austria) | Universidade da Madeira (Portugal)
WHO	The author.
STUDENTS INVOLVED	7 students in their first year of master's degree (6 enrolled in the interaction design program, 1 in media design). They worked in two groups of 2-3 people. They voluntary decided to participate in the workshop. | 15 third-year bachelor design students (working in groups of 5). Mandatory attendance.

VALUable ML Hero Agency – condensed

Contextual information

WHAT	Half-day workshop (3 hours).
WHEN	22 December 2022.
WHERE	Politecnico di Milano, School of design, Design & Engineering MSc program, *"Design Theory and Practice"* course held by Prof. Lucia Rampino.
WHO	The author.
STUDENTS INVOLVED	104 students MDes UX design students (working in 18 groups of 5-6). Participation was mandatory as they were attending the *Design Theory and Practice"* course that hosted the workshop.

Structure of the educational activity

ILOs	• Understand the core characteristics of ML systems. • Understand ML capabilities. • Identify what problems can be solved with ML systems. • Generate relevant, consistent, and effective design concepts including ML systems. • Understand ML systems as socio-technical systems and their capabilities. • Identify and use values to drive the design of ML systems. • Identify and anticipate possible impacts of ML systems in practical, personal, social, cultural, and eco-systemic dimensions. • Generate ethically acceptable, sustainable, and desirable design concepts including ML systems.

Structure of the educational activity

EXPECTED IMPACT(S)		• Build basic awareness on designerly and value-driven modalities to approach the conceptualization of ML systems toward consistent solutions. • Instill curiosity towards a still unusual tool for design. • Highlight the importance of early reasoning about values and impacts. • Recognize ML agents as design assets to reach meaningful goals.
CONTENTS		• ML definition • ML demystification • ML capabilities • Intro to Responsible Research Innovation (RRI) • Value-sensitive design • ML as a socio-technical system • Procedural information
TOOLS	***Knowledge transfer***	• Theoretical introduction • ML Agents • ML Suitability Matrix
	Design activities	• CBB • VDE
OUTPUT		**ML Hero concept:** *core structure*, (CBB + VDE completion) + *visual representation* (synthesizing its superpower and meaning) + *storyboard* (portraying the UX)

Overall

Research Rationale

RQ UNDER INVESTIGATION	***RQ4: Which design education method can support the conceptualization of ML-infused solutions?***
RESEARCH OBJECTIVE(S)	(i) Test the effectiveness of the educational method in terms of appropriateness of contents (typology, structure, format), tools, and teaching-learning activities. (ii) Understand how the didactic models unfolded in practice (which results were produced, whether some unexpected techniques to reach the goal emerged, or something was missing). (iii) Identify space for improvement.
TARGET AUDIENCE RELEVANCE	Create a diverse sample of design students (for educational context, cultural background, level of education, design specialization, etc.).

Methodological framing

EXPLORATION STRATEGY	The workshop activities and contents were presented and supported by the researcher. After ice-breaking questions to assess the class's level of knowledge and predisposition toward ML, an interactive presentation introduced the topic. Formative tests were included to reinforce understanding, and most of the time was dedicated to a hands-on design activity, developed in groups, to elaborate a technologically consistent and/or responsible *ML Hero* concept and representation. The results were presented to the class for a peer review based on predefined parameters (*relevance* of the addressed problem, *consistency*, overall *effectiveness* in bringing a positive impact in people's lives; as well as *ethical acceptability, sustainability,* and *social desirability*). To conclude, a general discussion, the completion of a final questionnaire, and optional written feedback wrapped up and evaluated the entire educational experience.
DATA COLLECTION	• Observation (supported by a structured observation sheet aimed at gathering information about students' responses to teaching-learning activities, contents, and tools) • Formative tests, peer, and self-evaluations (quantitative data to be compared with the researcher's assessment) • Pre- and post-workshop questionnaires. • Students' delivery and presentation of ML-infused concepts. • Oral and written feedback (subject to a subsequent content analysis).
RESEARCHER'S ROLE	Participant observer (facilitating the educational activities while gathering feedback and insights).

Findings

KEY INSIGHTS	• All the ILOs were quite positively met, with more successful results in the workshops adhering to the educational models. The condensed one, presenting several differences, was a useful counterevidence for some of the basic assumptions. • The tools were helpful for their purposes. Feedback on *ML Agents* confirmed the primary role of examples in understanding the subject. Central to the design activities, the CBB was particularly appreciated for the guidance and process provided, which were inspiring but leaving the freedom to make personal decisions. The VDE successfully elicited reflection and oriented the concept development toward responsible solutions. • Most participants felt an increase in their knowledge about ML and underlined the effectiveness of the formative tests in supporting their learning process. • The tools' physical nature and playful interaction are among the most valuable features of the method, as they favor discussion and collaboration. • The proposed approach proved versatile to be assimilated by design students from different specializations, with only a difference in aptitude between undergraduate and graduate students, more reflexive in the second case. • The studio format was consistent with the educational purposes not only for the practical application of knowledge but for the preferable dimension of the class and the essential support of a facilitator. • Modularity worked fine, but the holistic approach, merging the disciplinary perspectives from the very beginning, expressed the full potentialities of the educational models.
ISSUES FOR FUTURE INVESTIGATION	• Enhance learning-by-doing activities. • Depict more complex and plausible systems with interconnected capabilities. • Introduce the producer's perspective (to include financial and technical feasibility). • Find a way to synthesize, communicate, and hand over ideas to ML experts. • Make designers and ML experts collaborate. • Include the Operative Knowledge level of the ML Designerly Taxonomy. • Increase the innovative nature of the outputs. • Concretely address relevant challenges.

6.2 Defining the educational models for the assessment

Always set in action research modalities, building on previous experiences and reflections, more precise educational models have been developed to synthesize and concretize the objectives and assumptions of the research based on the presented theoretical constructs and tools. Even though more iteration cycles could be developed to investigate all the nuances of the research problem, these educational models represent the borderline and main output of the doctoral work (as they encompass all the others). As anticipated, they also include the prerequisites for their validation. Workshops' synthetic sheets are presented in the previous pages (212 - 217) while the detailed description follows.

6.2.1 General layout overview

Following the theoretical assumptions underlying the research, a dual strategy for implementation was envisioned. The approaches oriented towards consistent and responsible ML solutions could and have resulted in both separate and complementary educational modules. Specifically, the first two pilot workshops individually embedded a consistent and responsible approach, while the next ones offered a unique educational experience, differentiating the two parts to facilitate evaluation.

However, all share some common features. First, they are all set in a narrative according to which a fictional *ML Hero Agency* calls students to be intern designers. It is presented as a transnational organization developing ML-infused products and services for good (*ML Heroes*) to improve the quality of life on Earth at all scales and guarantee a better future. In accordance with this purpose, after a short training, the interns are asked to work in groups to imagine their first *ML Hero*, consolidate their learning and test their understanding. The metaphorical definition of ML-infused artifacts intends to intuitively evoke the overall objective of reaching positive achievements by leveraging super-human capabilities. Moreover, it is sufficiently general to avoid any explicit categorization (e.g., suggesting the development of a particular artifact), which guarantees students' freedom. For the responsible approach, the *WACSI division* (*World Agency for Challenging and Strategic Issues*) steers the *ML Hero* concept development towards a high-level aim (SDG) and a value-driven approach.

To implement the educational and research method requirements, all the workshops open with few multiple-choice questions to assess the class's level of knowledge and predisposition toward the topic. Then, an interactive presentation provides introductory notions essential to the following activities. It is interspersed with formative tests (questions and simple exercises) supported by the Wooclap platform. After a brief explanation of the expected tasks and a real-time practical example of how to use the tools, the hands-on design activity is developed. The results are presented to the class, which then evaluates the concepts based on predetermined parameters.

To conclude, a general discussion, the completion of a final questionnaire, and optional written feedback wrapped up and evaluated the entire educational experience, and the *ML Hero Agency* interns are awarded the qualification of *ML Hero designers*. As learned from previous experimentations, multiple strategies to collect comments and perspectives were necessary because of the difficulty of encouraging students to participate in conversations or answer the conclusive form when nothing compels them. To this goal, tiny paper envelopes with two questions inside were prepared and distributed to the participants who could leave fast, anonymous feedback in a dedicated box at the end of the workshop. The physicality and enjoyability of this in-person activity possibly motivated more students to write their comments.

Based on this structure, the educational models could adapt to different requirements, objectives, timings, and contexts by slightly adjusting their components. The specific details and strategies adopted for each workshop are described in the following.

6.2.2 Consistency model

To specify what is illustrated in Fig. 6.1, the educational experience unfolding is summarized here.

Three questions, working as icebreakers, ask the design students to self-assess their knowledge of ML and aim at understanding their perspective on the topic (like it was done in the *ML Pills for Designers* and *Superpowered Museums* workshops, except for the longest question about identifying ML applications).

ILO C1 is addressed by a presentation problematizing why ML should be a material for designers, introducing what ML is (in relation to AI, traditional programming, and providing definitions), and uncovering myths (like multitasking, complexity, autonomy, and highlighting the role of people). This is integrated with a couple of questions (formative tests) focusing on the definition of "agent" and some characterizing qualities of ML.

Analogously, the explanation of *ML Agents* and simple exercises contribute to ILO C2. Specifically, *training activity 1* requires recognizing ML systems capabilities in existing products and services, while *training activity 2* asks students to identify desirable applications for ML systems in their daily life. *Training activity 3,* challenging them to infer the suitability of ML for some suggested cases, with the support of the *ML Suitability Matrix*, responds to ILO C3. Addressing the first three ILOs takes approximately half a day.

Instead, the more articulated ILO C4 involves the elaboration of a technologically consistent *ML Hero*, for which the basic version of the CBB tool is explained and provided. The groups need to keep track of the way they use the tool for research purposes. Introducing the brief and letting students build their *ML Hero* structure can require an hour and a half to 3 or 4 hours. Additionally, a visual representation (drawing, model, etc.) and an annotated storyboard of the *ML Hero* are requested, respectively, to highlight the hero's importance and meaning and to depict the human-hero interaction (Fig. 6.3). These activities, the final presentations, peer evaluation, and conclusive session take up about a day.

The parameters according to which the ideas should be developed and assessed are the *relevance* of the addressed problem, the *consistency* of their concept technological core, and its overall *effectiveness* in bringing a positive impact in people's lives.

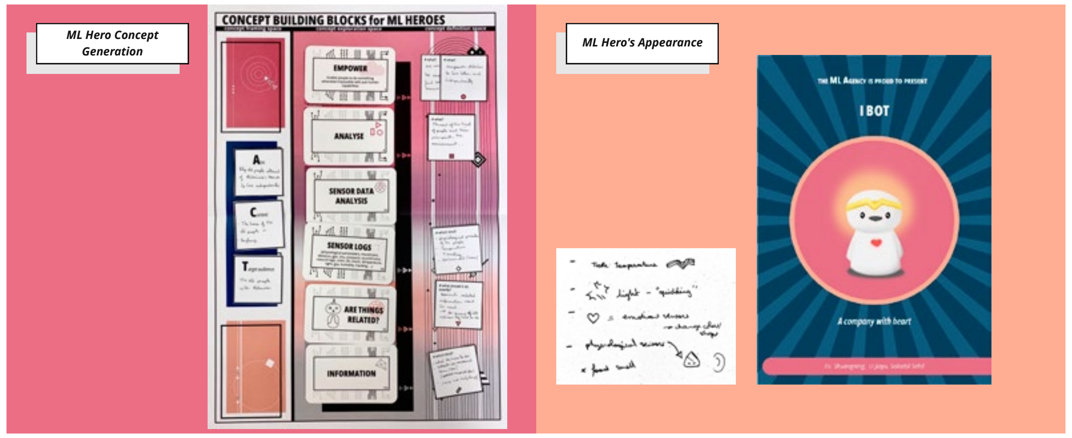
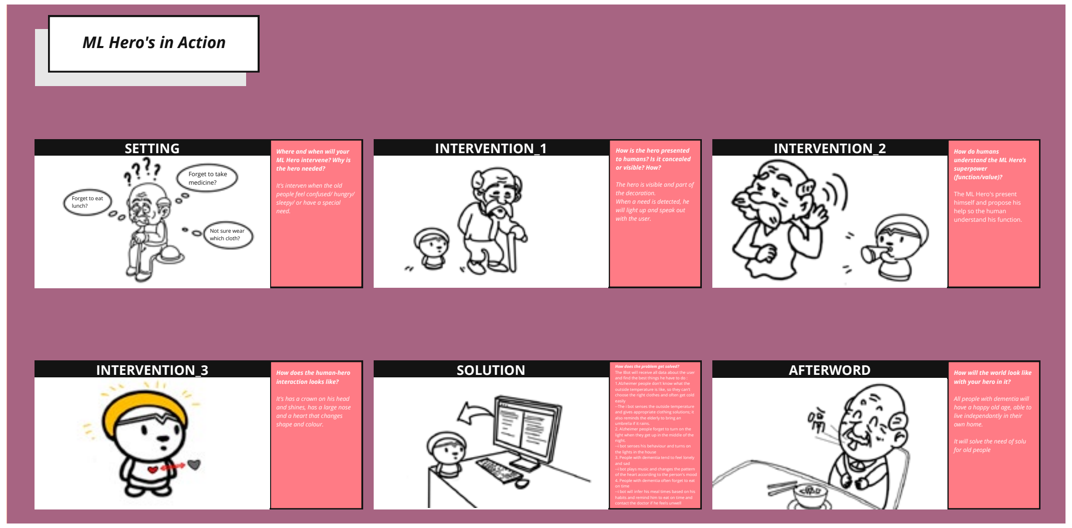


Fig. 6.3 | Example of the materials delivered for the ML Hero Agency workshop in Nantes (from left and continuing in the next page: completed Concept Building Blocks board, visual representation of the conceived ML Hero in the form of a poster, and annotated storyboard).

This model has been tested as a self-standing, two-day workshop with 18 international first-year students (9 females, 9 males) enrolled in the Digital Design MDes at the École de Design Nantes Atlantique.

6.2.3 Responsibility model

Detailing what is represented in Fig. 6.2 and following the general outline from section 6.2.1, the students were first asked about their knowledge of ML and their opinion of a value-driven approach to designing ML systems.
Then, ILO R1 was addressed through a very brief theoretical introduction of the challenges posed by ML, RRI and value-sensitive design as ways to counter it, and a definition of ML as part of a sociotechnical system and its capabilities (using *ML Agents*). As for the previous model, a couple of formative questions were submitted via Wooclap to reinforce the understanding of the definition of "agents" and "sociotechnical systems". This could be synthesized in less than one hour.
Within half a day, most of the time was dedicated to the practical design activity as the predominant educational component, supporting ILOs R2, R3, and R4. After explaining the task at hand and showing how to use them, the CBB tool and the VDE (as presented in 5.2) guided the whole design and reflective activity toward envisioning a *VALUable ML Hero*. Completing the tool boards is enough to get a responsible concept aligned with the expectations. To increase critical thinking, the groups are required to self-assess their idea according to all the *Principles cards* (promotion of flourishing, prevention of harm, attention to fairness, respect for human autonomy, and increase of intelligibility) before presenting them to the class. Then, they are involved in a peer evaluation activity based on the RRI requirements of *ethical acceptability, sustainability,* and *social desirability*. The workshop concludes with a feedback session analogously to the *consistency model*, with oral and written comments and through an online questionnaire.
Also in this case, the experimentation took place in the École de Design Nantes Atlantique and involved ten first-year MDes UX design students (5 females, 5 males). For the circumstances, this self-contained workshop had to last three hours, however, four would be ideal. For the limited time at disposal, all the groups (made of two or three students) had to address the Peace, Justice and Strong Institutions SDG, aiming



to prevent conflicts based on cultural differences and misunderstandings in the context of contemporary or future international or intercultural relationships and targeting people with prominent roles in the society.

6.2.4 Integrated consistency and responsibility model

The integration of the two models consists of a unique didactic experience in which the introduction of basic ML knowledge and its application precedes the implementation of a responsible and value-driven approach. In practice, the *consistency* and *responsibility models* are proposed subsequently, with the latter complementing the work developed in the first.
While the *consistency model* can be provided as it is, the *responsibility* one needs few adaptations. Of course, repetitions of the definition of ML systems and their framing as agents are avoided. Thus, the presentation only focuses on RRI and the value-driven approach.
Additionally, as the second module builds on the *ML Hero* concept already generated in the *consistency* one, some slightly different mechanics are introduced. The SDG is selected to fit the idea and the *Principle card* according to what the group thinks is the most suitable to fulfill their goal. At this point, some changes in the *ML Hero* structure or specifications might already be needed to make it coherent with the selected SDG and principle. They can be operated before passing to the *Reality Check* and *Value cards*. Afterward, the *responsibility model*, as described in 6.2.3, is regularly followed.
This integrated version, embodying the interdisciplinary spirit of the research, has been tested in three different contexts. In FH Joanneum University of Applied Science, in Graz (Austria), it involved seven students (5 females and 2 males) in their first year of master's degree. Six were enrolled in the interaction design program, and one was in media design. They were divided into two groups of two people and one of three. Here the workshop developed over three days, for a total of 18 hours (instead of the planned 24). In the *responsibility module*, to encourage the first iteration of the *ML Hero* (after the selection of the *SDGs* and *Principles cards*), the explanation of the VDE was divided into two parts.
The second iteration of the workshop presented no substantial changes. Only the VDE was introduced all at once. It was held in the Universidade da Madeira, with 15 third-

year bachelor design students (13 females and 4 males), organized in 3 groups of 5 people. A peculiarity is that there is no design specialization here. Although more time might have been necessary for the young students, out of the four scheduled days, three were actually employed, for a total of 16 hours.
Finally, a condensed format has been conceived to encounter the necessities of the *Design Theory and Practice* course held by Prof. Lucia Rampino in the first year of the Design & Engineering MSc program at Politecnico di Milano. It targeted 104 students (65 males, 39 females) who worked in 18 groups of 5 or 6 people, with the addition of one student who individually developed and delivered the required output. There were only 3 hours available for the workshop, so it underwent substantial adjustments but still retained the same structure.
The students were asked the preliminary questions and to explore the case studies from the *ML Agents* in advance. During the workshop, the introductory presentation and explanation of the CBB tool and VDE were completed in one hour. The former included the reason why ML should be a design problem, its basic definition, its capabilities (through *ML Agents*), its framing as a sociotechnical system, and essential notions from RRI and value-sensitive design. For the latter, the support of a projected Miro board was necessary to give the live example. Of course, no training activities or formative tests could be performed, and for working with the boards, a hybrid methodology was prepared, providing the students with the physical CBB cards and giving them all the other materials in a digital format on the Miro board (with a random selection of five *Reality Check cards*).
To condense the needed time, only one way to use the tools was provided: the value-based one, which started from the *SDGs* and *Principles cards*, to complete all the steps, including addressing two *Reality Check cards*. Also the SDGs were already provided as a starting point, and #04 Quality Education, #11 Sustainable Cities and Communities, and #15 Life on Land were assigned to 6 groups each. To complete the activity related to the *consistency model*, 40 minutes were suggested, including a check with the *ML Suitability Matrix*, tracking how the tool was used, and a self-assessment on the usual parameters (relevance, consistency, and effectiveness). Indeed, no poster or annotated storyboard could be prepared in this short time.
For the other, 30 minutes were esteemed sufficient, again including a self-assessment of the *ML Hero* according to the RRI requirements. Even though there was no time for oral presentations, a written one was delivered (identifying strengths and weaknesses of the proposed solutions and possible improvements). Only the conclusive questionnaire was submitted to the participants for feedback. Despite the differences, this organization guaranteed the collection of consistent data with respect to the other two integrated workshops.

6.3 Assessing models and tools. Workshops as case studies

The educational models have been applied as presented, neatly separating the assessment of the technological *consistency* and *responsibility* strands. As no significant criticalities emerged during the first workshops, the iterations mostly consisted in the adaptation to the context and availabilities of the hosting universities and courses, trying to maintain sufficiently homogenous conditions for an integrated analysis.

This was possible for the experiences in Nantes, Graz, and Madeira, which followed the premises of the educational models more accurately. The extreme conditions and choices for the condensed workshop in Polimi, instead, reflected in the results, which are somewhat inconsistent with the previous cases. Stressing the character of the models proved particularly effective in highlighting actual or potential criticalities and made the Polimi workshop a perfect counterevidence for some of the basic assumptions of the educational method.
Accordingly, in the following, when this divergent nature of the results emerges, a **double interpretation is presented** to show what happens if the ideal minimum conditions are met or not.

6.3.1 ML Heroes overview

Analyzing the concepts produced as required outputs of the educational activities (presented in Tab. 6.1), no substantial differences can be spotted among the different contexts. From a qualitative point of view, even though some were conditioned by predefined briefs or impacts, a diverse set of ideas was created. The vast majority demonstrated that the participants assimilated enough information about what ML can do to succeed in the development of technologically consistent solutions.
Keeping into account the limited time the students had at their disposal, and that it was not the goal of the activity, no disruptive ideas came out. On the contrary, some systems can already be found in our lives (like ML-based simultaneous language translators, NC-01). What is relevant, though, is that all the solutions result from a problem or necessity that the students considered relevant to tackle, not from the attempt to emulate ML-infused systems they already knew, which denotes a success for the fostered approach.
The more daring or particular the problem framing, the more original the result. It is the case of *F.R.E.P.* (MA-01) that, targeting arachnophobes, resulted in a frog-shaped robot using computer vision and classification to detect and capture spiders in the house. Of course, it does not aim to be the most meaningful project. Nonetheless, as a first experimentation, it led to an unconventional application of ML.
In general, the adopted choices denoted some interesting tendencies of design students. Most *ML Heroes* (13 out of 34) were fed sensor logs, indicating how the design students tried to connect ML to world-based applications. The ML tasks that have been used the most are classification (11), sequence prediction (11), and clustering (6), producing several recommendation systems (11). From the outputs, instead, no relevant information can be inferred. They are quite varied and reflect the intentions expressed by the selected *ML Agents* with a prevalence of categories – *Typology* (5) and *Quality* (4) cards were among the most used – and recommendations – with five *Suggestion* and one *Recommendation* cards employed.
From a thematical point of view, leaving the concepts with a given brief aside (NR-00s), the most recurrent topics were recycling (4 cases plus one about finding the nearest bin in public spaces) and sustainable cultivation (3 cases), with several ideas referring to the organization of people's lives, their time, personal spaces or commuting habits.
Tab. 6.1 collects all the peer and self-evaluations that the participants expressed during the workshop activities (in the Polimi case, codes MI-00, peer reviews were not possible due to time limitations) and that the researcher defined to estimate

Code	Title + Concept	Input-Task-Output	ML solution	Relevance (res. ev.)	Relevance (peer/self-ev.)	Consistency (res. ev.)	Consistency (peer/self-ev.)	Effectiveness (res. ev.)	Effectiveness (peer/self-ev.)
NC-01	**LAN' BUDDY. The buddy for languages the better communication you achieve |** To overcome communication barriers caused by speaking different languages, Lan' Buddy is embedded into a pair of AR glassess. It receives the speech and translates it in real time on an earphone.	(i) Audio contents (foreign languages) - (t) Sequence prediction - (o) Typology (user's mother language)	4	**4**	3,47	**4**	3,21	**3**	3,00
NC-02	**FEELY. Music the feel |** As a virtual tool, it connects to different devices (smart watch and phone) to collect the user's physiological parameters, launch and adapt the music to improve people's mood.	(i) Sensor logs (pulse, respiration) - (t) Sequence prediction - (o) Audio content (music, ambient sound)	4	**4**	3,29	**3**	3,23	**4**	3,17
NC-03	**DEEP FREEZE. No more waste! |** To preserve food, a fridge integrate the ML Hero, which determines whether the food is good or not and adapts the internal conditions.	(i) Sensor logs (humidity, pressure, aspect, temperature, smell, date, usage) - (t) Classification - (o) Quality (of food)	4	**4**	3,20	**4**	3,62	**3**	3,38
NC-04	**ENERGIZE. Energize your energy |** To help women in their postpartum care, a custom-built device monitors the condition of the pelvic floor and, based on data collected from tests, suggests exercises to restore a healthy condition.	(i) Sensor logs (pelvic floor contraction strenght) - (t) Sequence prediction - (o) Suggestion (exercise advice)	4	**4**	3,50	**4**	3,00	**4**	3,21
NC-05	**I BOT. A company with heart |** It helps people with dementia or Alzheimer's disease by monitoring user and external conditions, reminding them what they need to do, and supporting them in daily life (eating, dressing properly, taking medication, etc.).	(i) Sensor logs (physiological parameters, temperature, environmental) - (t) Clustering - (o) Information (easy and daily things one has to do)	4	**4**	3,58	**4**	3,08	**4**	3,25
NR-01	The system suggests judges the possible fairest decision, based on similar cases in a dataset curated by several jusdges. The system does not take any decision.	(i) Historic data (past fair judjements, decided by judges) - (t) Sequence prediction - (o) Suggestion (to orient the decision in a fair way)	4	**3**	/	**3**	/	**2**	/
NR-02	Using several different law systems as a reference, the system comes up with new laws that encounter also the necessities of minorities.	(i) Written content (penal codes) - (t) Action selection - (o) Suggestion (new law that best match the situation)	3	**3**	/	**1**	/	**3**	/
NR-03	To bring together people with different cultures, the system suggests activities that might interest different people and to which they can decide wether they want to participate.	(i) Environmental information (center of interest) - (t) Action selection - (o) Suggestion (activities to do together)	3	**4**	/	**1**	/	**3**	/
NR-04	The system transforms depositions into visual reconstructions that should help decision-makers in trials to spot inconsistencies.	(i) Audio contents (audio recordings of people's stories) - (t) Generation - (o) Performance (of the scene)	3	**2**	/	**3**	/	**3**	/
GR-01	**LINCOM. Listen to me! |** Lincom can be activated to analyze its user's skills in speaking a foreign language. Better working in specific contexts, it can point out the most common mistakes the person has made and suggest ways to improve the environment-related vocabulary.	(i) Audio contents (speech) - (t) Classification - (o) Typology (mistakes)	4	**3**	3,86	**4**	3,71	**4**	3,86
GR-02	**CAPTAIN ALIKE. Don't spoil the soil, baby! |** With sensors to collect multifaceted soil information, the system elaborates maps of correlations to help human experts understand how to grow healthy food in a sustainable way.	(i) Sensor logs (to be defined with experts in soil management) - (t) Clustering - (o) Relation (map of factor correlations - e.g. fertilizer & agricultural yield)	4	**4**	4,00	**4**	3,14	**4**	3,71
GR-03	**THE QUALITIZER. Qualify for quality by dunking useless quantity |** System that learns what pictures have value for the person and what can be deleted to free them of unnecessary data. It co-evolves with the user.	(i) Visual contents (images, pictures on the phone) - (t) Classification - (o) Quality (value or no value)	4	**4**	4,00	**4**	3,71	**4**	3,86

Tab. 6.1 | [On the left] synthesis of the concepts generated in all the workshops applying the educational models. [On the right] ML Heroes assessments by the researcher (res. ev.) or the students (peer/self-evaluation in the Polimi case) based on their relevance, consistency of the adopted ML Agent, and effectiveness. Additionally, the author established whether the proposed ideas were solutions for ML to achieve or not. Codes stand for: NC > Nantes - Consistency model; NR > Nantes - Responsibility model; GR > Graz; MA > Madeira; MI > Polimi workshops.

Code	Title + Concept	Input-Task-Output	ML solution	Relevance (res. ev.)	Relevance (peer/self-ev.)	Consistency (res. ev.)	Consistency (peer/self-ev.)	Effectiveness (res. ev.)	Effectiveness (peer/self-ev.)
MA-01	**F.R.E.P. (Frog Removes Every Pest). Gotta catch 'em all! |** To help arachnophobes get rid of spiders, F.R.E.P. can be activated to detect and catch spiders that it collects in the belly to release them later.	(i) Sensor logs (movement, color, touch, tracking, form) - (t) Classification - (o) Subject (spiders)	4	**3**	2,75	**4**	3,06	**4**	2,80
MA-02	**ORGANIZER. Disorganized home? I can help! |** The system suggests how the photographed space could be organized and gives instructions.	(i) Visual contents (images of spaces, objects) - (t) Action selection - (o) Decision (what, where to organize / separate by sections)	4	**3**	3,86	**1**	3,57	**3**	3,87
MA-03	**REECO. Reduce, reuse, ycle with Reeco. Your trash is in good hands now. |** It automates the recycling process in public spaces, automatically sorting the trash that citizens throw in.	(i) Sensor logs (image, weight, color, texture, temperature) - (t) Classification - (o) Typology (paper, glass, plastic, organic, electronic, batteries, oil)	4	**4**	3,86	**4**	3,93	**3**	3,92
MI-01	It creates a personalised study plan based on interests, current courses, and time availability. It suggests courses and how to organize one's time.	(i) Activity logs (weekly schedule, classes, home assignments,...) - (t) Sequence prediction - (o) Suggestion (most suitable activities and timing organization)	3	**3**	3,50	**2**	3,33	**2**	3,00
MI-02	**YouLearn |** It transforms online information in free visual educational content, to make education accessible to everyone that has access to internet, despite language and age barriers.	(i) Written content (online content) - (t) Sequence prediction - (o) Visual content (accessible educational material)	3	**3**	2,60	**1**	3,00	**2**	4,00
MI-03	**Mario Attention Officer |** The system provides teachers with data regarding the attention of the classroom, so that they can improve their lectures.	(i) Visual contents (Real time recording of the class, and pictures of bored/unattentive people) - (t) Classification - (o) Quality (general or individual attention level during class)	3	**3**	3,20	**4**	3,40	**3**	3,75
MI-04	**PLANTASKIC |** The system allows students with extra curricular activities to optimize time balancing personal life activities and good performances at school.	(i) Historic data (personal habits, hobbies, activities assignements, exams divided by priority, time spent in each task) - (t) Sequence prediction - (o) Visual content (A weekly/monthly color-coded schedule with all the tasks that needs to be done)	3	**3**	3,17	**3**	2,83	**3**	3,33
MI-05	**BT - Bat Translator |** Provide support for blind people to study by themselves by vocally describing visual contents.	(i) Visual contents (Image, video, drawings and handwriting) - (t) Classification - (o) Subject (speech or voice message)	3	**4**	3,40	**3**	3,40	**3**	3,60
MI-06	**ML UNI_CO |** The system suggests relevant educational resources and materials based on students' interests and goals to help them choose the right educational path.	(i) Written content (Resume, school marks, interests, job descriptions databases) - (t) Clustering - (o) Affinity (set of suggested career or study paths)	4	**3**	2,67	**4**	2,83	**3**	3,00
MI-07	**Traffic Manager |** It regulates traffic at an intersection by adapting traffic lights to the best way to facilitate circulation at a given time.	(i) Numerical properties (amount of queued vehicles, provenience, time of day) - (t) Sequence prediction - (o) Visual content (traffic lights color)	2	**4**	3,43	**4**	3,43	**4**	3,71
MI-08	**FBB (From the Bin to the Bus) |** System that properly separate garbage to help the environment.	(i) Environmental information (garbage categories) - (t) Action selection - (o) Decision (garbage separation)	4	**4**	3,17	**2**	3,17	**4**	2,67
MI-09	**TrashTalk |** To improve correct recycling in the domestic environment, an app suggests the correct container in which to throw waste based on uploaded pictures.	(i) Visual contents (discarded materials) - (t) Classification - (o) Typology (correct bin)	4	**3**	3,33	**4**	3,00	**3**	3,33
MI-10	**FLOWME |** The system aims at providing personalized solutions and plans to the transportations companies and to the consumers in order to make the traffic flow better and increase the comfort of public transportation.	(i) Sensor logs (amount of people using public transport, weather, time...) - (t) Classification - (o) Quality (obtained by reducing mental stress)	4	**4**	2,58	**1**	3,08	**2**	3,25
MI-11	**MR. SORTY |** To sort any kind of trash properly in the domestic environment.	(i) Sensor logs (factors to recognize materials, e.g. magnetic response for metals, light for glass, etc.) - (t) Classification - (o) Typology (of trash)	4	**4**	3,40	**4**	3,80	**3**	2,20

Code	Title + Concept	Input-Task-Output	ML solution	Relevance (res. ev.)	Relevance (peer/self-ev.)	Consistency (res. ev.)	Consistency (peer/self-ev.)	Effectiveness (res. ev.)	Effectiveness (peer/self-ev.)
MI-12	**ECO - Pilot |** Inspire car drivers to choose the most sustainable transportation or to choose different times on which use the private transpotations.	(i) Sensor logs (GPS, movement, pollution data collector, distance) - (t) Clustering - (o) Relation (between the distance and the emitted pollution)	4	**4**	3,75	**3**	3,00	**3**	2,83
MI-13	**Spotty |** To help zoologists and researchers monitor the behaviour of animals, the system analyzes people's photos to spot wildlife in particular zones and in relation to time.	(i) Audio contents (Websites, social media, smart phones) - (t) Sequence prediction - (o) Recommendation (when and where to find species)	4	**4**	4,00	**4**	3,50	**4**	3,42
MI-14	**AGRO |** The system suggests suitable best crop rotation for the next season to preserve soil fertility, based on the soil composition.	(i) Sensor logs (pressure, temperature, humidity, light, gas, nitrogen in soil) - (t) Sequence prediction - (o) Written content (Prediction about the nutrients available in the soil/farm, and suggestion of suitable crops for next season)	4	**4**	3,67	**4**	3,33	**4**	3,67
MI-15	**Desertification Visual Forecaster |** The system aims to track and find patterns to prevent future desertfication by generating forecasts.	(i) Visual contents (satellite images) - (t) Sequence prediction - (o) Visual content (future Earth conditions)	4	**4**	3,67	**4**	3,33	**4**	2,83
MI-16	The system intends to improve the relation between human and nature by augmenting users' awareness on the conditions of their surrounding environment through the detection of anomalies.	(i) Sensor logs (location, chemicals, humidity, gas, pressure) - (t) Classification - (o) Anomaly (toxicity, possible actions/interactions)	3	**4**	3,50	**4**	2,50	**3**	2,67
MI-17	**Water/Mineral Configuration for homegrown Plants |** To assist people in home-growning plants and vegetables, the system automates water and mineral distribution & configuration based on the plants condition thereby maximising yield output.	(i) Sensor logs (Soil moisture, humidity, PH, temperature, greenness of leaves, mineral mixtures, historical data) - (t) Clustering - (o) Response (to improve poor plant conditions, measured optically, creation of a revised mineral mixture, change water distribution or warmth)	4	**4**	3,50	**3**	3,50	**4**	3,00
MI-18	**RAIN METER |** To empower people make decisions about water management and prevent bad consequences of possible futures droughts or heavy rains, the system identifies the areas that need water the most.	(i) Numerical properties (need of water in an area, from industry, agriculture, and environment) - (t) Regression - (o) Location (of the areas most in need of water)	4	**4**	2,60	**2**	2,50	**4**	3,20
MI-19	**BINFORM |** The system helps to find the nearest bin in order to not leave the thrash on the streets.	(i) Numerical properties (bins location, capacity/fullness of the bins,when they were last cleaned out and when the next cleaning up is programmed) - (t) Clustering - (o) Information (nearest bins, specifying fullness typology)	2	**3**	3,00	**2**	3,00	**3**	2,00

the groups' performances on the acquired capability to envision a concept that can be framed as a ML solution (exploiting the strengths of this technology), addresses a relevant problem (having value for people and improving the quality of life at any scale), consistently applies *ML agents*' capabilities to address their design A.C.T. (aim, context, target audience), and can be effective in achieving the goal of positively impact people's lives. Comparing the last three parameters between the researcher and students' evaluations (for which the means are considered), no significant distance can be measured, except for the cases in which the researcher reported no consistency of the *ML Heroes* with the selected *ML Agent*. The students did not detect these discrepancies at the moment of their peer and self-evaluations.
Even if some information about the ILOs can be elicited from this overview, a more specific analysis is presented in the following.

6.3.2 *Using ILOs to assess the effectiveness of the educational method*

As anticipated, the educational and research purposes mostly coincide, especially in relation to the fulfillment of the ILOs. As multiple assessment methods have been implemented, a qualitative triangulation of data is reported, starting from the direct question submitted through the survey after the conclusion of the activities. In this matter, the discrepancy between the workshops applying the educational models as outlined and the condensed Polimi case is quite evident, so a dual layer of interpretation is necessary.

6.3.2.1 ILO C1: understanding the key features of ML systems

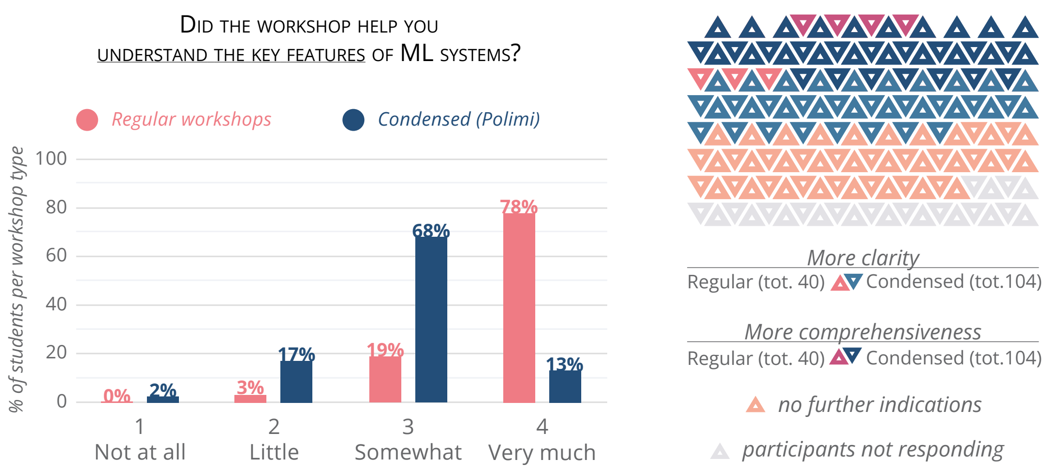


Fig. 6.4 | Students' evaluation of the workshop's effectiveness to help them understand the key features of ML systems.

Based on their responses, represented in Fig. 6.4, the participants mostly felt to have grasped the main characteristics of ML systems, especially those who had longer presentations. This is also confirmed by the formative tests that students in Nantes, Graz, and Madeira had after the theoretical introductions. Aimed at reinforcing some key points, the questions and exercise *training activity 1* have been structured to be tricky and foster reflection. Nonetheless, the majority of the participants identified the correct answers (Fig. 6.5).

The first training activity, requiring the identification of ML systems in existing products and services (including the correct ML capabilities for the last 6 out of 10 examples), was more insidious. Here, the most uncertainty emerged in the distinction between artifacts integrating ML or based on traditional programming, like in the case of the presented navigation systems. In fact, most respondents marked the one only calculating the path with the shorter distance between two points as ML-infused.

Being partially useful to evaluate students' comprehension of ML key features, the formative tests have been highly recognized by the participants as essential components to achieve ILO C1, as visible in Fig. 6.6. Indeed, underlying that they were

not included in Polimi didactic experience (blue triangles), the training activities are surprisingly second only to the CBB tool for the attendants to the other workshops. Of course, the practical understanding conveyed by the interaction with the tool is unparalleled, but also the theoretical introduction of the contents and the *ML Agents*, especially for the provided examples, played a significant role in both kinds of applications of the educational models. In the condensed case, having a relatively small percentage of people declaring that this ILO was little or not achieved and several asking for more clarity and comprehensiveness (Fig. 6.5), three comments explain the motivations for this. Two highlighted that the explanation was complex, possibly due to the short time at disposal, and the other expressed that vagueness had to be avoided, including more details. Indeed, the key features presented were reduced to the definition of agents and their structure, with a quick mention of the capability of ML systems to improve their performances with experience. Hence, not enough elements or time were dedicated to this ILO in the condensed version, while the interactive format of the other workshops proved successful.

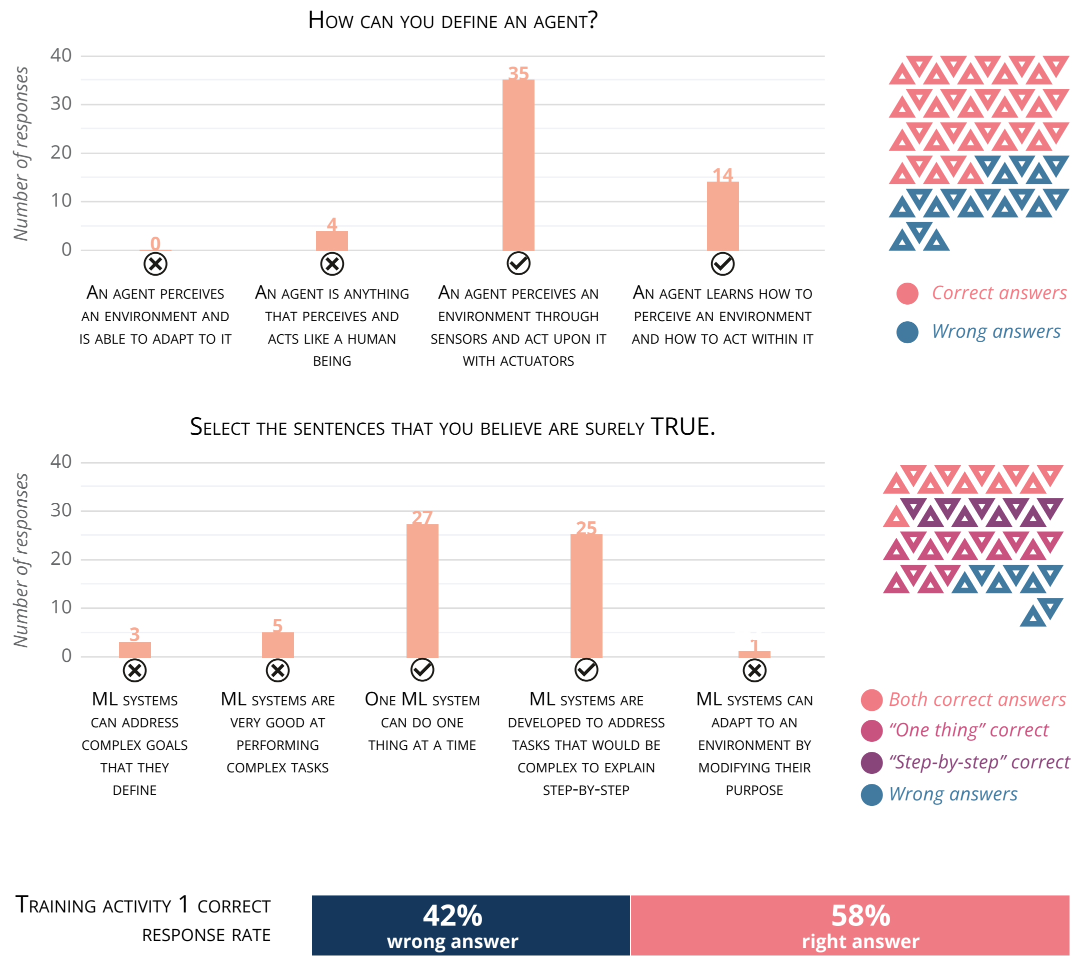


Fig. 6.5 | Participants' correctness rate to the formative questions of the consistency model.

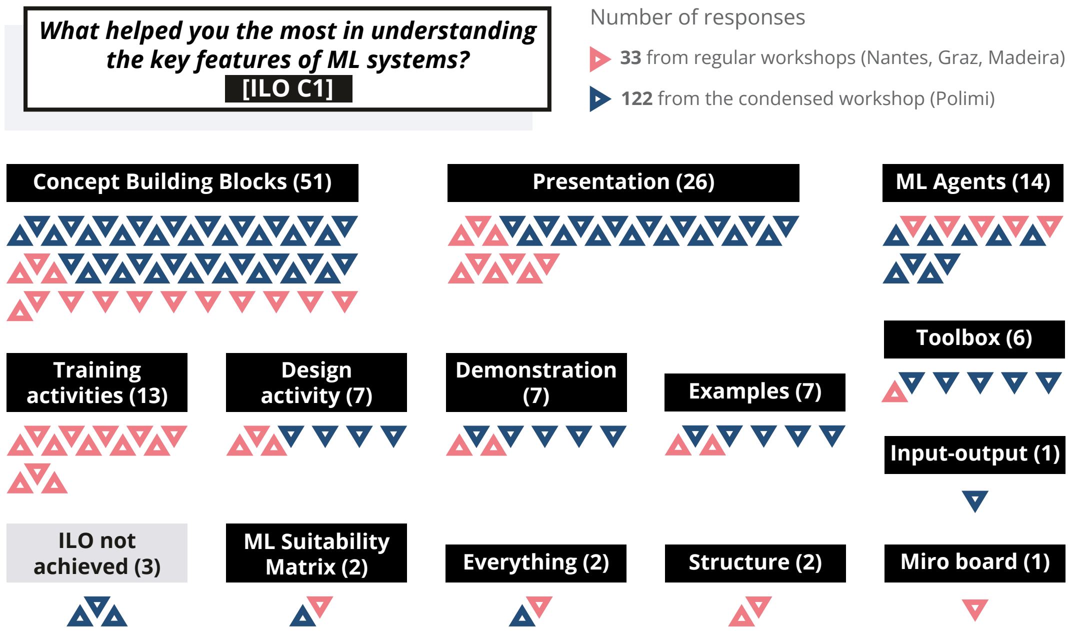


Fig. 6.6 | Identification of the elements conveying the fulfillment of ILO C1.

6.3.2.2 ILO C2: understanding the capabilities of ML systems

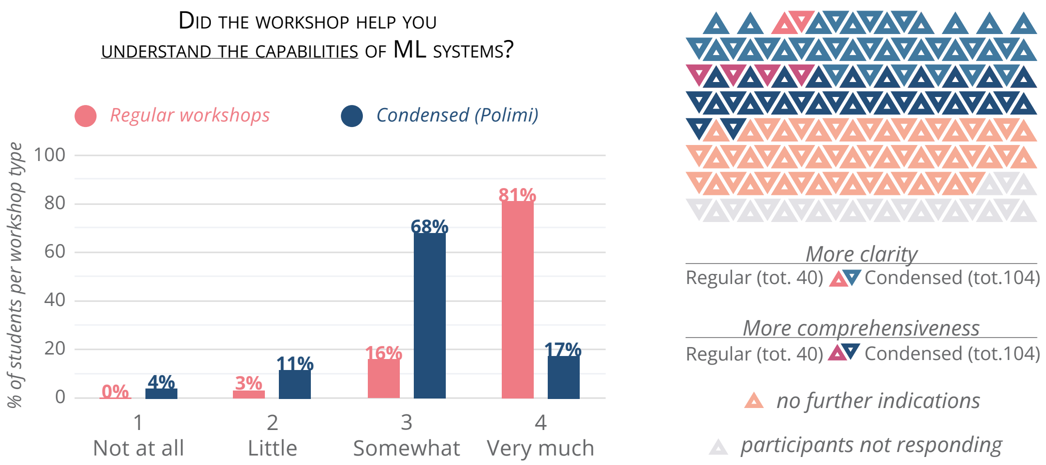


Fig. 6.7 | Students' evaluation of the workshop's effectiveness to help them understand the capabilities of ML systems.

Some indications of whether this expected learning outcome has been met can already be inferred from observing the produced concepts. As shown in Tab. 6.1 before, most of the ideas (18 out of 34) consistently integrated ML capabilities to address the problem identified in the design A.C.T., and in seven cases, they were quite aligned, but maybe a different *ML Agent* could have been better to match the described project. From an applicative point of view, then, this ILO seems fulfilled, and the same can be affirmed from comparing the self and peer assessment of the technological consistency with those by the researcher, as anticipated in 6.3.1.

However, both students' performances and self or peer evaluations highlighted some difficulties in understanding how to apply some ML capabilities. Specifically, *action selection* (out of four proposed cases, three are not consistent, and one is loosely consistent with this capability, mainly mistaken for a recommendation system) and *regression* systems (the only idea integrating it does not output numerical properties). This is also confirmed by the few requests for clarification during the reviews, which were mainly related to these *ML Agents*, and often discarded afterward for not being in line with the idea.

In general, though, the researcher as facilitator did not have to intervene much in the ML capabilities selection as they were mostly coherent with the presented concepts. For the workshops following the envisioned educational models, the final inconsistencies were mainly due to a change of narrative from the elaboration of the *ML Hero* structure (day #01) and its representation in the storyboard (day #02), suggesting that it is important to ensure that students recall the reasons behind their choices.

Also from the participants' perspective, this ILO was quite positively achieved (Fig. 6.7), with the submitted explicit question reporting 13 negative ratings out of the 88 respondents of the condensed workshop and none from the others. Again, most of the requested clarification and comprehensiveness came from the former, and examples are commonly recognized as valuable resources for this purpose.

Fig. 6.8 illustrates that even if CBB remained the most quoted item to enhance the comprehension of ML capabilities, the presentation and provided examples closely follow. To be noted is the fact that case studies and examples were specifically pointed out, even more frequently than *ML Agents* themselves. As a tool precisely envisioned for transferring knowledge about ML capabilities, the *ML Agents* have been curiously mentioned as useful for this purpose as much as the *ML Suitability Matrix*.

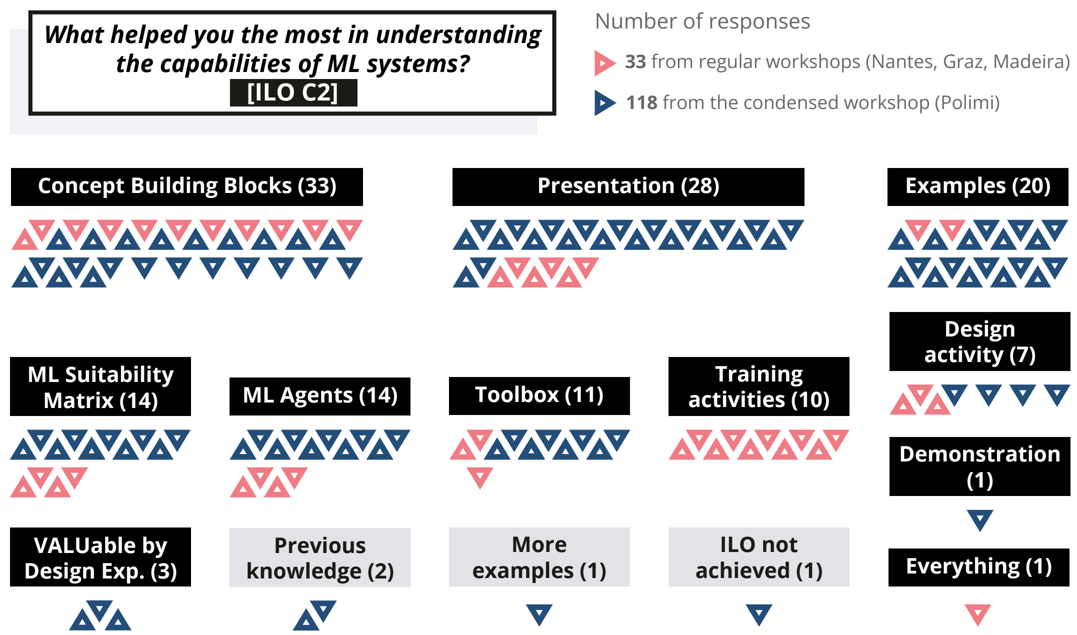


Fig. 6.8 | Identification of the elements conveying the fulfillment of ILO C2.

6.3.2.3 ILO C3: identifying relevant ML problems

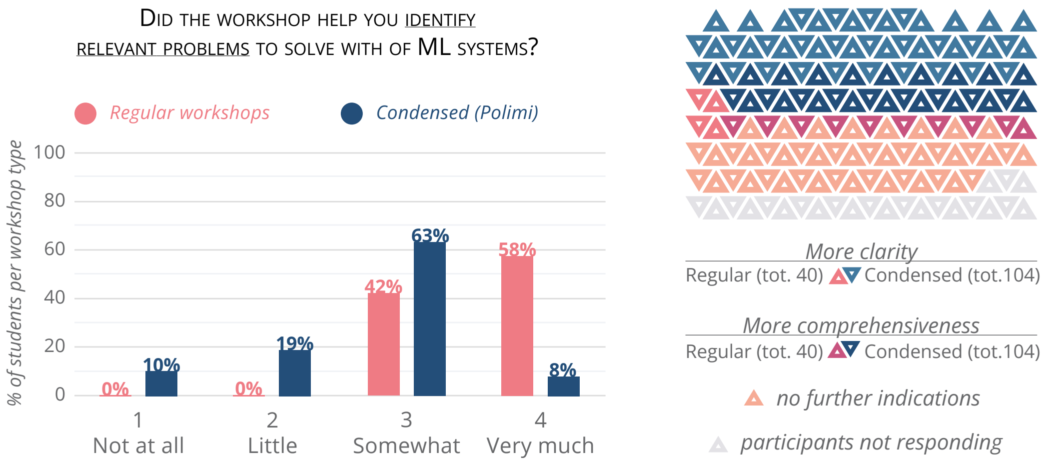


Fig. 6.9 | Students' evaluation of the workshop's effectiveness to help them identify relevant ML problems.

Among the ILOs of the *consistency model*, this received the lowest results (Fig. 6.9), coherently reflecting the least time spent on its explanation. Indeed, it was openly presented only during *training activity 3* (the last activity in the morning of the first day, which in Madeira was not even completed by all the groups nor discussed together as the time was over). It coincided with the quick introduction of the *ML Suitability Matrix* to support the assessment of six examples of ML applications, according to the value they bring to people and the consistency with ML capabilities. In the Polimi workshop, the subject was not even touched upon during the presentation, and the only support provided was the *ML Suitability Matrix* to be interpreted based on few instructions written on the Miro board. In both cases, though, the related activity was highly subjective, with no clear right or wrong answers and no fixed rules to be identified. Despite the importance of the topic, its ambiguity and weight in the educational model brought the result that more comprehensiveness has been broadly recognized as necessary, and more clarity is fundamental in the condensed version, as the tool probably needs some kind of explication.
Also the variety of elements indicated to be helpful for identifying relevant ML problems sustains the fuzziness of the issue (Fig. 6.10). Overall, CBB and the presentation have been cited the most in the comments, but considering the integral application of the educational model, the training activities are protagonists. Indeed, while *training activity 3* specifically addressed this ILO, few students considered the *ML Suitability Matrix* to assess their ideas in the rest of the design activities. On the contrary, it was the sole complementary tool provided in the condensed version, so it gained more importance among its participants.
Interestingly, along with the always essential examples, also the designerly framing of the concept was recognized as noteworthy of a specific mention.
Finally, out of 119 comments, eight stated that nothing was helpful in reaching this ILO, and five did not answer the question, all from participants of the Polimi workshop.

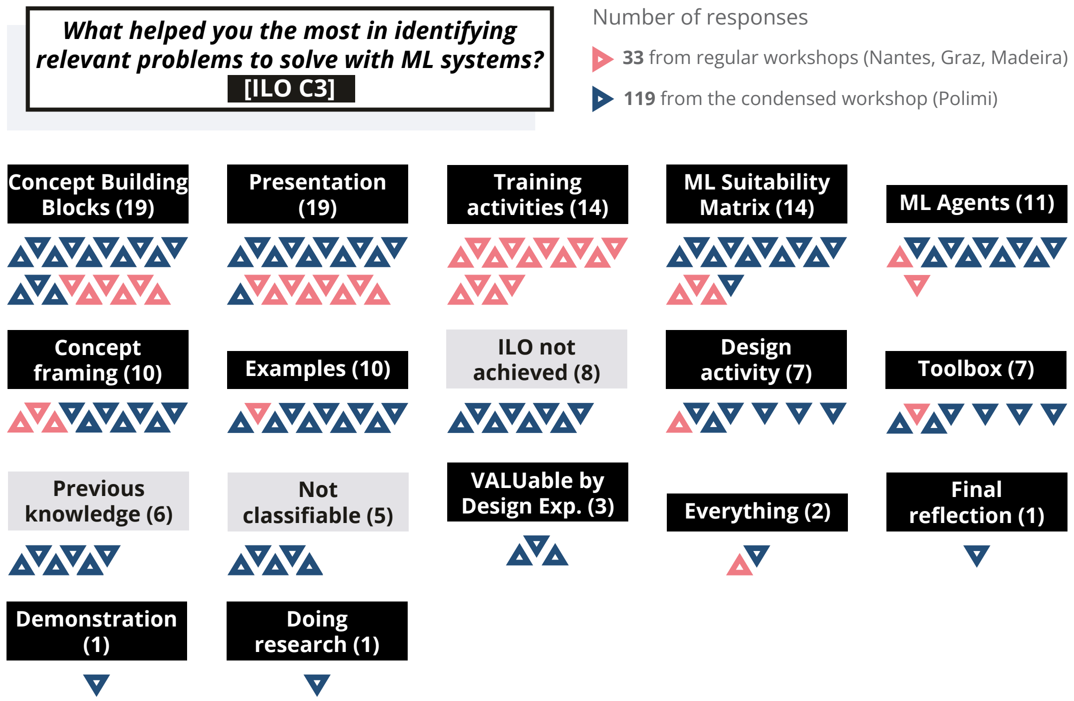


Fig. 6.10 | Identification of the elements conveying the fulfillment of ILO C3.

However, the analysis of the conceived ML Heroes contradicts the understandable remarks. As reported in Tab. 6.1, only two of the 34 concepts would not really benefit from the application of ML systems, and most of them (21) were judged as proper ML solutions by the researcher.

6.3.2.4 ILO C4: generating relevant, consistent, and effective design concepts including ML systems

Of course**, few hours spent dealing with ML systems from a highly conceptual level cannot be enough to master the generation of relevant, consistent, and effective ML-infused solutions**. However, with proportionate expectations, almost all the conceived ML Heroes succeeded in all the parameters, most of them marking the highest score for relevance and consistency in the researcher's personal opinion (Tab. 6.1).
Also the self-recognition of this skill is generally positive (Fig. 6.11), even if not with the conviction that characterized the responses for ILOs C1 and C2, which is definitely justifiable for the extent and complexity of the task.
Not surprisingly, CBB supported students the most in the development of their concepts, as this is their intended purpose (Fig. 6.12). The process and guidance provided were particularly prominent in respondents' feedback and, strongly connected, were essential in a first approach to designing with a new material.
Interestingly, the comments from some students in Madeira and Nantes place the training activities as the second most useful element for the fully applied *consistency modules*, probably reinforcing the importance of practice for acquiring applied knowledge.
Finally, once again, for some Polimi students (less than 20%), it has been difficult or not possible to achieve this ILO, and the motivations emerging from some comments

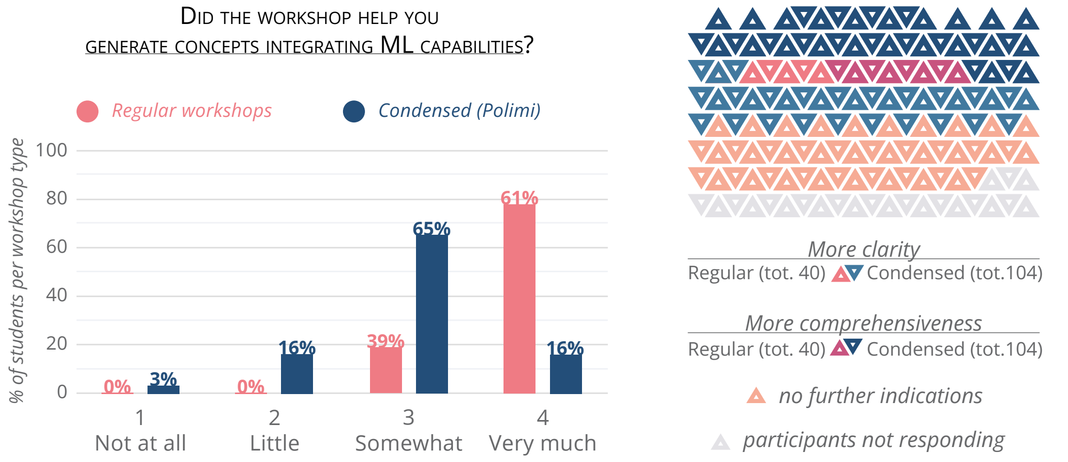


Fig. 6.11 | Students' evaluation of the workshop's effectiveness to help them generate relevant, consistent, and effective design concepts including ML systems.

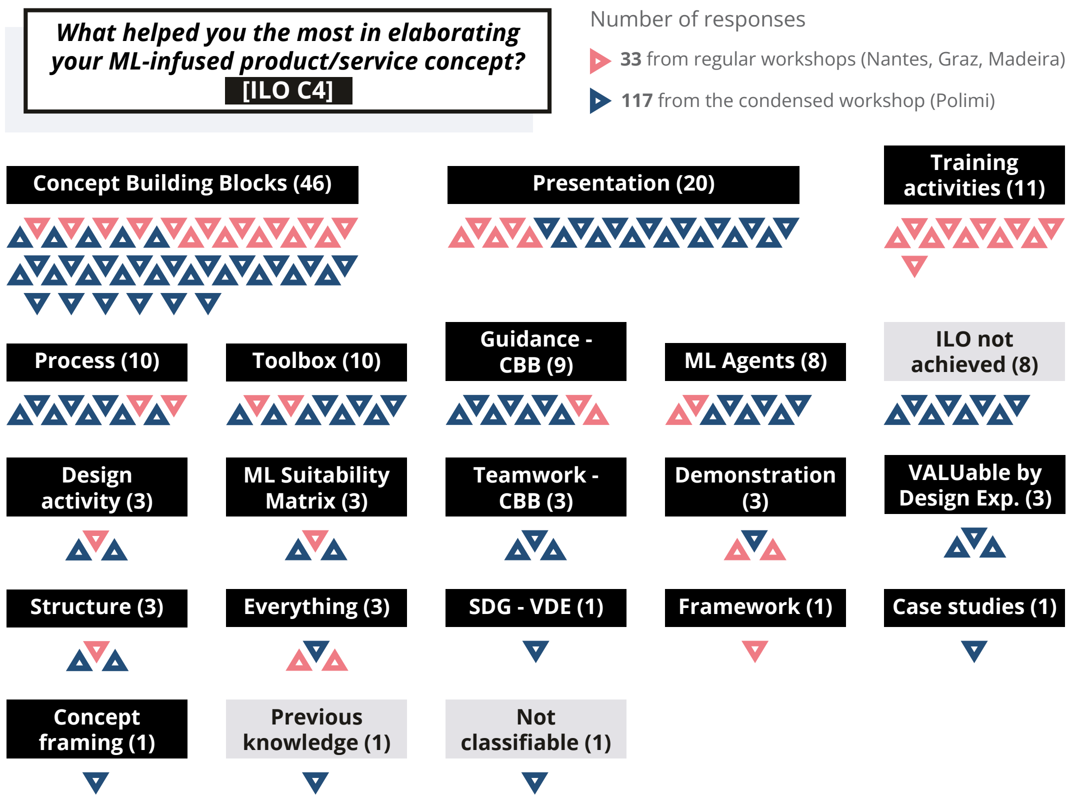


Fig. 6.12 | Identification of the elements conveying the fulfillment of ILO C4.

echoed the previous ones: limited time and more explanations are needed. Indeed, the demonstration of how to use the tool, highly appreciated in the other workshops, was rather short and necessarily less interactive than the other cases, as the elevated number of students did not allow otherwise. Additionally, the researcher could not adequately support all 18 groups in the class during the 40 minutes provided for the activity, which is also relevant for assessing the educational model. It was observed that some difficulties were encountered with the digital interface of the tool (not detectable with the physical version) at the expense of the available time.

6.3.2.5 ILO R1: understanding ML systems as socio-technical systems

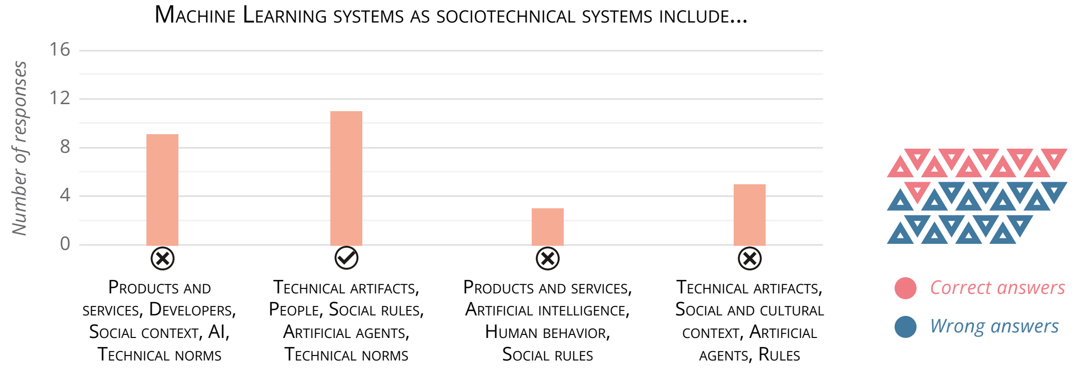


Fig. 6.13 | Students' responses to the formative question of the Responsibility model.

In all the tested cases, the responsible model could develop in a shorter time with respect to the consistency counterpart because essential knowledge about ML capabilities was necessary to eventually design *VALUable ML Hero* concepts. This is the reason why few theoretical contents were provided to introduce the module, leaving little space for assessing the acquired knowledge, or none like in the Polimi workshop. Understanding the definition of ML systems as agents (considered in the evaluation of ILO C1) was the first step into getting both a technical and ethical understanding of ML as part of larger sociotechnical systems defined in the relations with people and their environments: a milestone for building the entire responsible discourse. This is why the second formative test revolves around this concept, once again subtly focusing on the details of its composition to strengthen the concept in students' educational path. Fig. 6.13 demonstrates the complicated nature of the issue by showing how, overall, the correct answer has been the most recognized but not by the majority of the respondents, especially in Madeira, where most people assumed that just the developers were enough to represent human agents in the sociotechnical systems. However, it is possible to observe how this concept was applied in practice in the presented *ML Heroes*, even if not all had an introduction to responsible design (namely, not the NC-00s). Indeed, already framing the problem in a designerly way was sufficient to broaden the perspective from the technology itself to the ecosystem in which it had to be inserted. Additionally, most of them (22) went beyond the consideration of an individual user to make the envisioned solution part of a broader scope and aimed to have an impact also at a social or environmental level (Tab. 6.2).

6.3.2.6 ILO R2: identifying and using values to drive the design of ML systems

All the other ILOs of the *responsibility model* were supposed to be conveyed by the practical activity supported by the VDE of the CBB and thus assessable by the participants in the conclusive questionnaire.
The majority of the participants were still positive about the tool's effectiveness in helping identify and use values to drive the design of ML systems (Fig. 6.14). Many, from both modalities of application of the educational model, underlined the clarity and usefulness of the *Principles* and *Values cards* to instill a reflection process about

Code	SDG	Principle	Scale of impact	Risks are addressed	Final reasoning	Acceptability (res. ev.)	Acceptability (peer/self-ev.)	Sustainability (res. ev.)	Sustainability (peer/self-ev.)	Desirability (res. ev.)	Desirability (peer/self-ev.)
NC-01	/	/	Personal + Social	/	/	/	/	/	/	/	/
NC-02	/	/	Personal	/	/	/	/	/	/	/	/
NC-03	/	/	Personal + Social	/	/	/	/	/	/	/	/
NC-04	/	/	Personal	/	/	/	/	/	/	/	/
NC-05	/	/	Personal	/	/	/	/	/	/	/	/
NR-01	#16 Peace, justice and strong institutions	Attention to fairness	Personal + Social	Yes	/	3	3,10	2	2,90	2	2,80
NR-02	#16 Peace, justice and strong institutions	Attention to fairness	Social	Yes	/	3	2,90	2	3,20	3	2,90
NR-03	#16 Peace, justice and strong institutions	Attention to fairness	Social	Yes	/	4	3,30	3	3,10	3	3,00
NR-04	#16 Peace, justice and strong institutions	Attention to fairness	Personal + Social	Yes	/	2	3,10	2	3,10	3	3,20
GR-01	#4 Quality education	Increase of intelligibility	Personal	Yes	/	3	3,14	3	3,86	4	3,86
GR-02	#15 Life on land	Promotion of flourishing	Social + Environmental	Yes	/	4	4,00	4	4,00	4	4,00
GR-03	#12 Responsible consumption and production	Increase of intelligibility	Personal + Environmental	Yes	/	4	4,00	3	3,29	3	3,29
MA-01	#15 Life on land	Prevention of harm	Personal	Yes	/	3	2,82	3	3,45	3	3,36
MA-02	#3 Good health and well-being	Respect for human autonomy	Personal	Yes	/	3	3,18	3	3,18	3	3,36
MA-03	#11 Sustainable cities and communities	Promotion of flourishing	Social + Environmental	Yes	/	4	3,91	3	3,73	4	3,91
MI-01	#4 Quality education	Respect for human autonomy	Personal	No	Yes	2	3,67	2	2,33	3	3,17

Tab. 6.2 | ML Heroes responsible factors (SDGs, Principles, scale of impact) and related evaluations by the researcher (res. ev.) or the students (peer/self-evaluation in the Polimi case) based on their ethical acceptability, sustainability, and desirability. The author also assessed if the drawn risks were actually addressed and if Polimi students critically elaborated the final reasoning. Codes stand for: NR > Nantes - Responsibility model; GR > Graz; MA > Madeira; MI > Polimi workshops.

Code	SDG	Principle	Scale of impact	Risks are addressed	Final reasoning	Acceptability (res. ev.)	Acceptability (peer/self-ev.)	Sustainability (res. ev.)	Sustainability (peer/self-ev.)	Desirability (res. ev.)	Desirability (peer/self-ev.)
MI-02	#4 Quality education	Increase of intelligibility	Personal	Yes	Yes	**2**	2,60	**3**	3,40	**2**	4,00
MI-03	#4 Quality education	Promotion of flourishing	Personal + Social	No	Yes	**2**	1,25	**2**	2,75	**2**	2,50
MI-04	#4 Quality education	Respect for human autonomy	Personal	No	Yes	**3**	2,92	**2**	2,75	**3**	3,67
MI-05	#4 Quality education	Respect for human autonomy	Personal	Yes	Not really	**3**	3,40	**3**	2,60	**3**	3,50
MI-06	#4 Quality education	Respect for human autonomy	Personal	Yes	Not really	**3**	3,67	**2**	2,50	**3**	2,67
MI-07	#11 Sustainable cities and communities	Promotion of flourishing	Personal + Environmental	No	Yes	**4**	2,83	**2**	3,50	**4**	3,00
MI-08	#11 Sustainable cities and communities	Attention to fairness	Social + Environmental	No	Yes	**3**	3,67	**2**	3,33	**2**	3,00
MI-09	#11 Sustainable cities and communities	Promotion of flourishing	Personal + Environmental	No	Not really	**3**	4,00	**2**	3,17	**3**	3,17
MI-10	#11 Sustainable cities and communities	Respect for human autonomy	Personal + Social	Yes	Not really	**3**	3,08	**2**	3,08	**2**	3,17
MI-11	#11 Sustainable cities and communities	Prevention of harm	Personal + Environmental	Yes	Yes	**3**	4,00	**2**	3,80	**3**	2,80
MI-12	#11 Sustainable cities and communities	Promotion of flourishing	Personal + Environmental	Yes	Yes	**3**	3,00	**2**	3,83	**3**	3,00
MI-13	#15 Life on land	Prevention of harm	Environmental	Yes	Yes	**3**	3,50	**2**	4,00	**3**	3,00
MI-14	#15 Life on land	Promotion of flourishing	Environmental	No	Yes	**4**	3,67	**3**	3,83	**4**	3,50
MI-15	#15 Life on land	Prevention of harm	Environmental	No	Yes	**4**	3,08	**4**	3,58	**4**	3,00
MI-16	#15 Life on land	Increase of intelligibility	Environmental	Yes	Not really	**4**	3,33	**4**	2,50	**4**	3,50
MI-17	#15 Life on land	Promotion of flourishing	Personal	Yes	Yes	**4**	3,83	**4**	3,83	**4**	3,17
MI-18	#15 Life on land	Prevention of harm	Social + Environmental	Yes	Yes	**4**	3,83	**3**	3,83	**4**	3,17
MI-19	#15 Life on land	Promotion of flourishing	Environmental	Yes	Yes	**3**	3,00	**2**	4,00	**3**	2,00
					Average	***3,21***	***3,30***	***2,62***	***3,33***	***3,14***	***3,19***

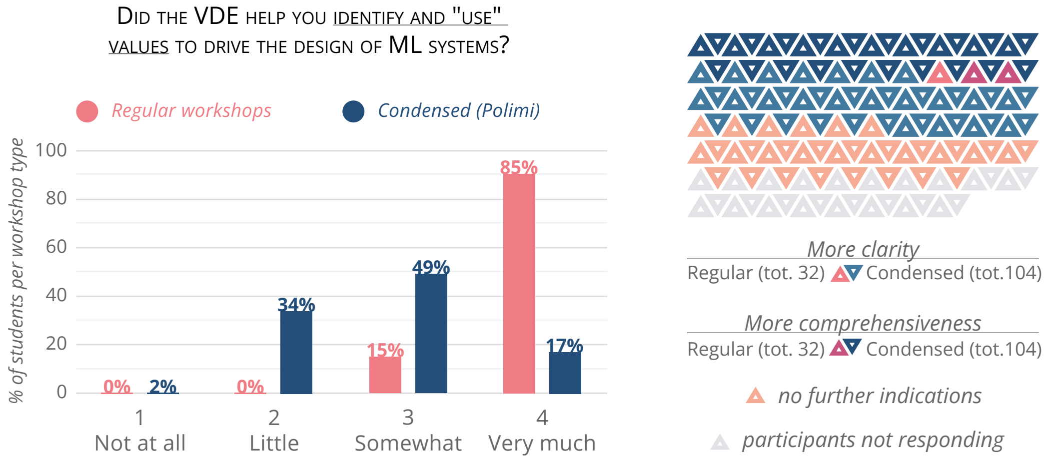


Fig. 6.14 | Students' evaluation of the responsibility module to identify and use values to drive the design of ML systems.

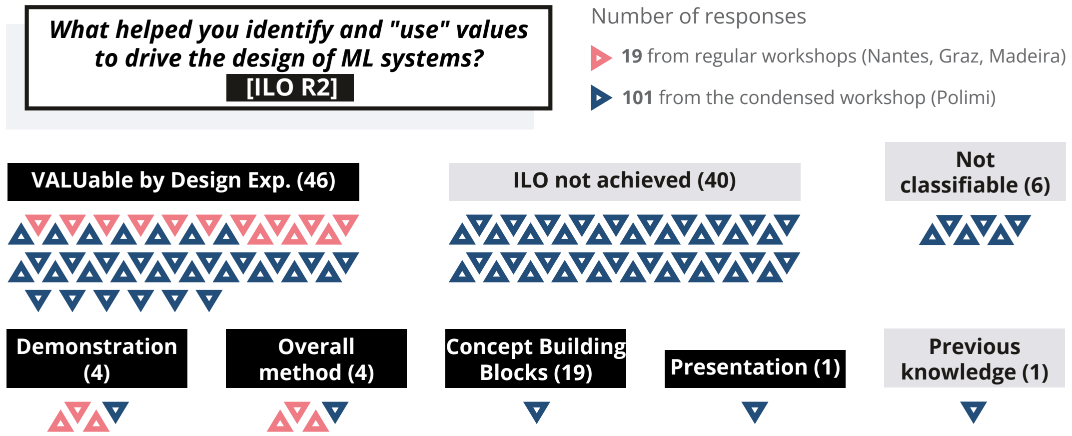


Fig. 6.15 | Identification of the elements conveying the fulfillment of ILO R2.

the desired impact (Fig. 6.15). It made design students pause to assume different perspectives and explore paths otherwise not investigated. As argued from the previous experiences, the value-driven process is not unfamiliar to designers' way of thinking. Still, it needs to be made explicit, as one comment demonstrates: *"The cards put in words what was already but kind of blurry in our mind."*

On an opposite note, a large portion of feedback from the Polimi workshop complained that what they had to do was not very clear and that they had too little time to handle it. The development of this part was indeed highly compromised in the condensed workshop as very limited time was left for its demonstration and for the students to properly dedicate to its completion. Then, while it is important to keep this in consideration in the analysis of the results, some interesting insights were prompted by this unfavorable situation. For instance, a comment reported that a group of students could not grasp how to connect the VDE to the concept structure of the *ML Hero*, suggesting that the materials per se did not express it well.

6.3.2.7 ILO R3: identifying and anticipating possible impacts of ML systems in practical, personal, social, cultural, and eco-systemic dimensions

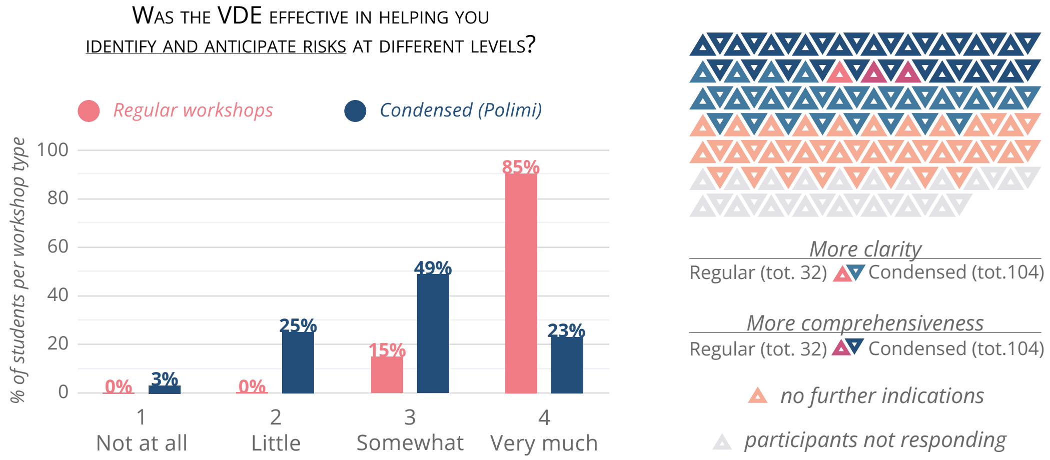


Fig. 6.16 | Students' evaluation of the responsibility module to identify and anticipate impacts of ML systems.

Also the anticipation of ML-related risks was a critical part. To shorten the weighing of options, the *Reality Check* cards had to be randomly drawn to decide later whether they could apply to the *ML Hero* concept or not. In the Polimi workshop, this process was further limited as only five *Reality Check* cards were available to each group (instead of the original 25), and they had to address a maximum of two. With enough time to ponder if and how the presented risks could manifest in a scenario where the *ML Hero* was deployed, the activity proved once again useful and successful (Fig. 6.16). Some remarks confirmed the necessity to make these critical reflections explicit during the design process because, otherwise, the attention focuses on other aspects of the project (Fig. 6.17). Also the mechanic of randomly drawing cards to check one's idea was appreciated for its dynamism, even if one of the respondents would have liked to go through all the cards to test all the possibilities. Indeed, this modality would enable students to gain more awareness, as they could have an overview of all the

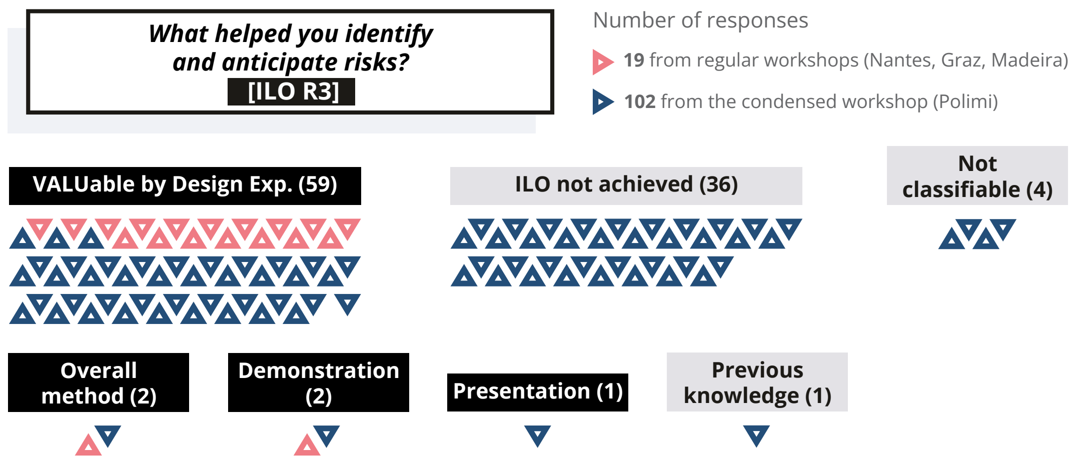


Fig. 6.17 | Identification of the elements conveying the fulfillment of ILO R3.

options before choosing those that are more relevant to their concept. However, more time should be dedicated to this activity.
Even if there are examples of in-depth and original responsible reflections also among Polimi students, and maybe the most thorough specifications come from a group of this condensed workshop, a fair part could not properly complete the exercise (and a group even left the *Specification cards* blank). The conditions did not favor the understanding of this activity, and eight groups out of 19 did not think about possible solutions to avoid or limit the risks from happening but were limited to echo their *ML Hero* qualities or the relevance of the threat or value selected. The comments even revealed that, to cope with the limited time available, a group split to deal with the *consistency* and *responsibility* parts in parallel, which is incompatible with the activity itself.

6.3.2.8 ILO R4: generating ethically acceptable, sustainable, and desirable design concepts including ML systems

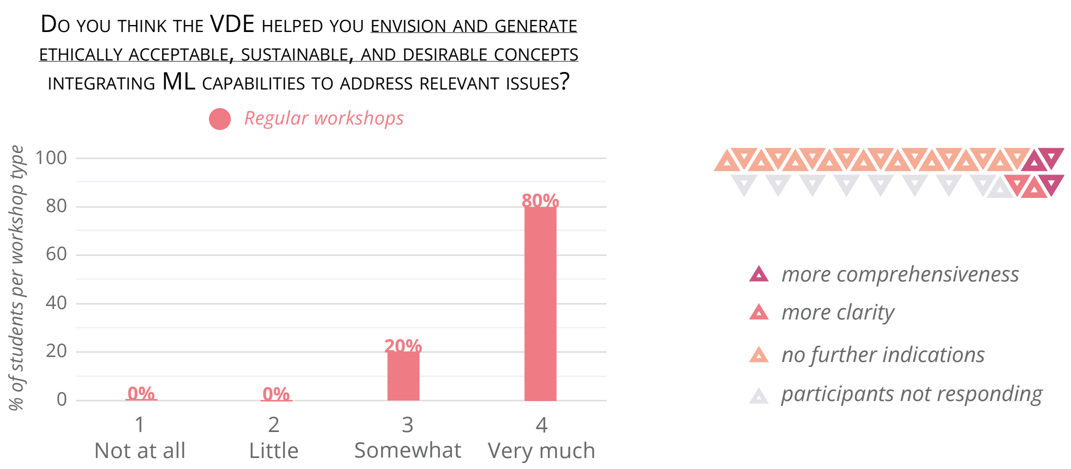


Fig. 6.18 | Students' evaluation of the workshop's effectiveness in helping them generate ethically acceptable, sustainable, and desirable design concepts including ML systems.

As tackling just two *Reality Check* cards in half an hour was not enough to justify the differentiation of the design process in the *consistency* and *responsible* parts, an explicit question was not directed to the participants of the condensed workshop. Thus, figures 6.18 and 6.19 only show the results from Nantes, Graz, and Madeira educational experiences. Their perception about the fulfillment of this ILO is again quite positive, with most of the respondents affirming that they felt that the VDE helped them very much to envision ethically acceptable, sustainable, and desirable concepts integrating ML. A comment even highlighted how the fostered process could lead to innovation. However, as one pointed out, more sessions would be needed to properly master the tools in relation to different projects.
In addition to the questionnaire results, other aspects can be considered to understand how the participants' performance relates to this ILO.
For instance, several students pointed out how the SDGs helped them to frame their idea in a responsible way, connecting it to real problems and suggesting the best principles to pursue. Indeed, in most cases, the influence of these objectives

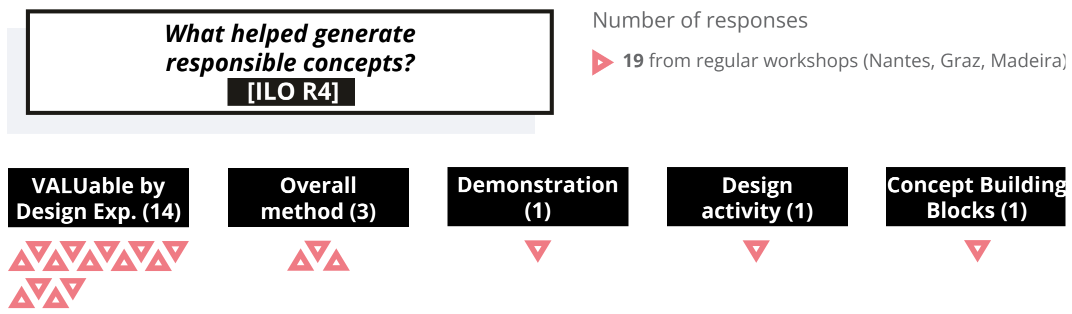


Fig. 6.19 | Identification of the elements conveying the fulfillment of ILO R4.

was clearly perceivable in the way the *ML Heroes* were developed. Selecting them encouraged one group to better frame their design A.C.T. while, among the groups who started with an SDG and a principle (Nantes and Polimi workshops), some decisions to limit potential risks were envisioned even before the *Reality Check* cards. Though, nobody felt the necessity to change the structure of their *ML Hero*.
The *responsibility* students felt invested with can also be observed in the principles they chose to guide their design process. It is peculiar that to deal with SDG #16 (*Peace, justice and strong institutions*), all opted for focusing on the *fairness* of their solutions. At the same time, the majority of groups decided to set their *ML Hero* on a positive and proactive path with the *promotion of flourishing* principle. Emblematic in this sense is the comment of a girl dealing with the *Life on Land* SDG. She noted that *"harm is already there"* and trying to prevent bad behaviors could lead to imposing solutions for the target audience. Hence, they changed their principle to *promotion of flourishing* so that their concept could develop in a less intimidating way.
Moreover, even though the activity with the VDE was partially effective for Polimi students, they demonstrated good critical skills in describing the strengths and weaknesses of their concepts. Due to time limitations, they were not requested to come up with the best possible ideas but to reflect on the positive and negative factors of their *ML Heroes* (Tab. 6.2). In this task, most of the groups (14 out of 19) demonstrated awareness on the impact of their choices and some envisioned ways to overcome undesirable issues, which indicates the achievement of this ILO.

6.3.3 Observing tools performances

6.3.3.1 ML Agents

ML Agents represent the main means envisioned to transfer ML knowledge. They were first explained during the introductory presentation and then provided in the form of a booklet to support the design activity and complement the CBB (in Polimi's case, they were digitally available on the Miro board). Several indications could be collected from the questionnaires, participant observation, and discussions with students. As Fig. 6.20 shows, their role in enabling the comprehension of ML capabilities has been mostly positively recognized, in combination with some issues.
The comments (Fig. 6.21) are aligned, and the most evident outcome is the relevance of examples. Confirming what emerged in the preliminary research conducted through the *ML Pills for Designers* workshop and reinforced in the *Introductory Game to ML Responsible Design*, the primary role of examples in understanding a new subject

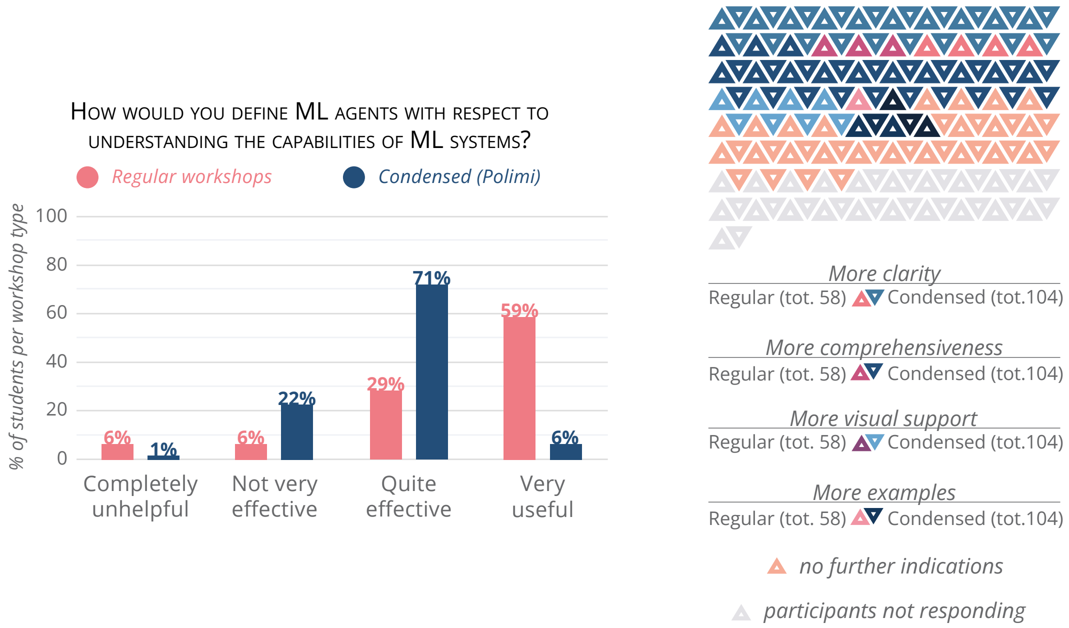


Fig. 6.20 | Students' assessment of the ML Agents.

is something that everybody agreed on. In fact, the request for more examples came from both positive comments underlying how the provided case studies were effective in making sense of ML capabilities, and from Polimi students who declared that they were not able to understand much of the topic. Some episodes that occurred during the workshops further emphasize this. For example, during the presentation of *ML Agents* in Graz, students intervened a lot to add and have confirmation about known applications of ML pertaining to different categories. In Nantes and Madeira, instead, some groups of students who finished their *ML Hero*'s structure earlier discussed further examples of ML capabilities with the facilitator.

Continuing to comment on the format used for communicating ML capabilities, the synthesizing questions, the simplified names, but also the reported skills and structure of the definitions (highlighting the input–task–output process) were cited as useful to foster comprehension and memory of the contents, as well as for inspiring possible applications. All these factors can justify why *ML Agents* were mainly perceived as easy and simple, helpful, interesting, effective, but also versatile to address different problems.

On the other side, few comments, exclusively from the Design & Engineering class who attended the condensed workshop, reported how *ML Agents* were not very clear. Specifically, some commented how they could be more comprehensive and with more technical and scientific language and contents, avoiding unnecessary simplification. As from the same group of students diametrically opposed feedback also arrived, for instance, underlining how the *"shocking simple terms"* but still accurate and comprehensive explanation made the contents accessible for *"non-nerds"*, it is relevant to notice how the personal background or predispositions might alter the perception. Interestingly, after nine months from their selection, one also commented that the examples could be more up-to-date, underlying how fast ML applications are spreading and multiplying. Of course, some parameters for the selection of the case studies (like the fact that they should be explorable and provide explanations on

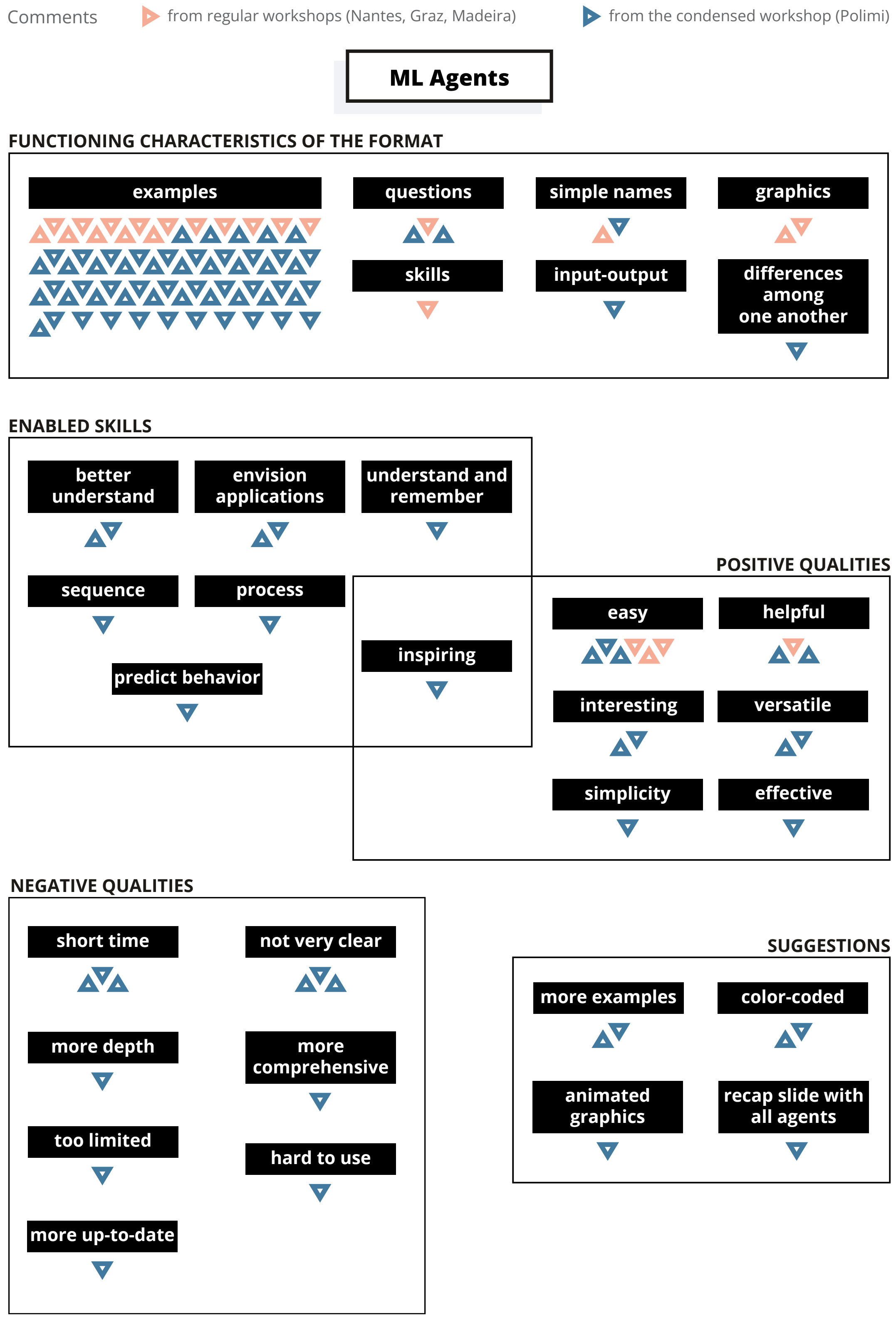
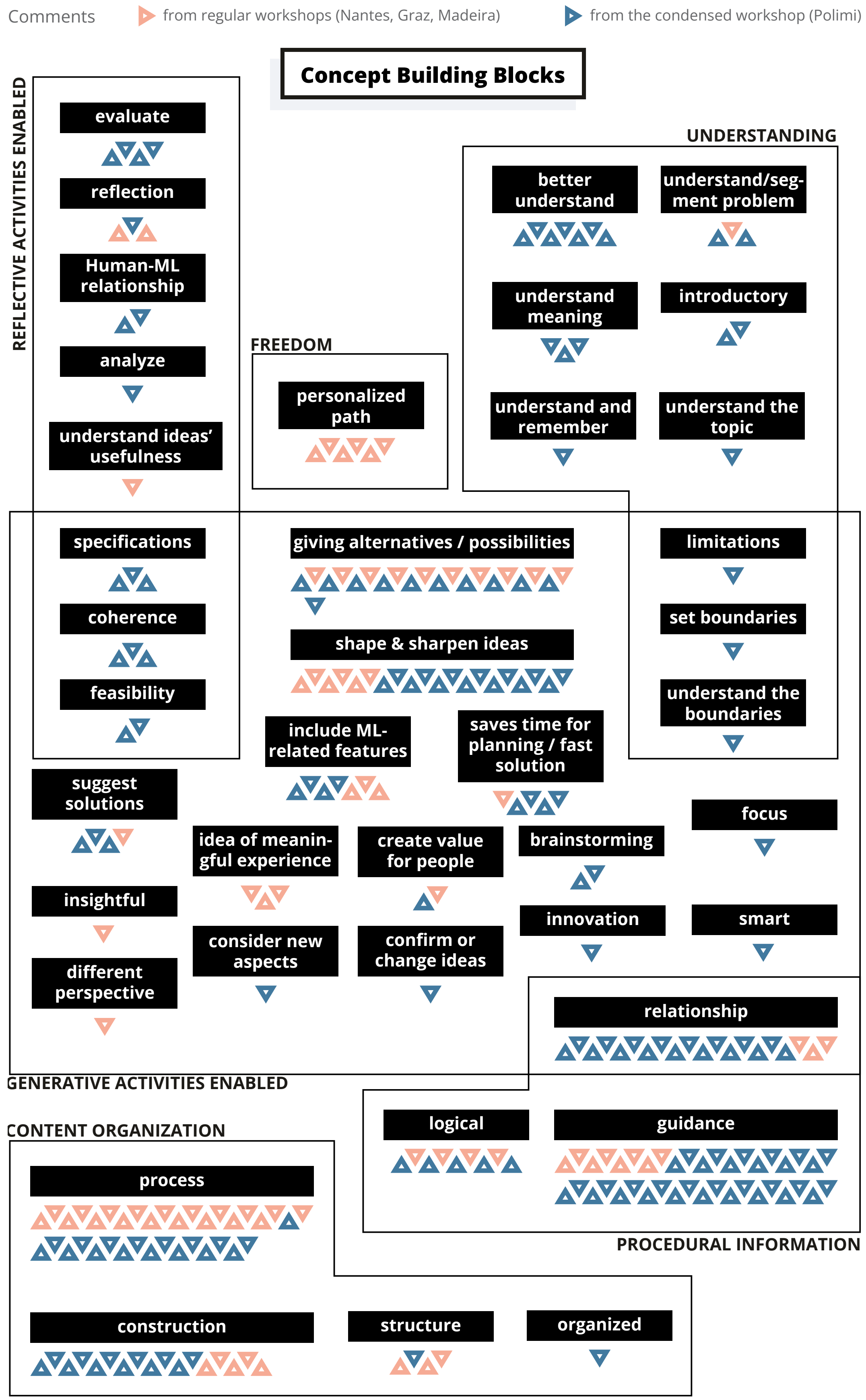


Fig. 6.21 | Summary of the content analysis of students' comments on the ML Agents.



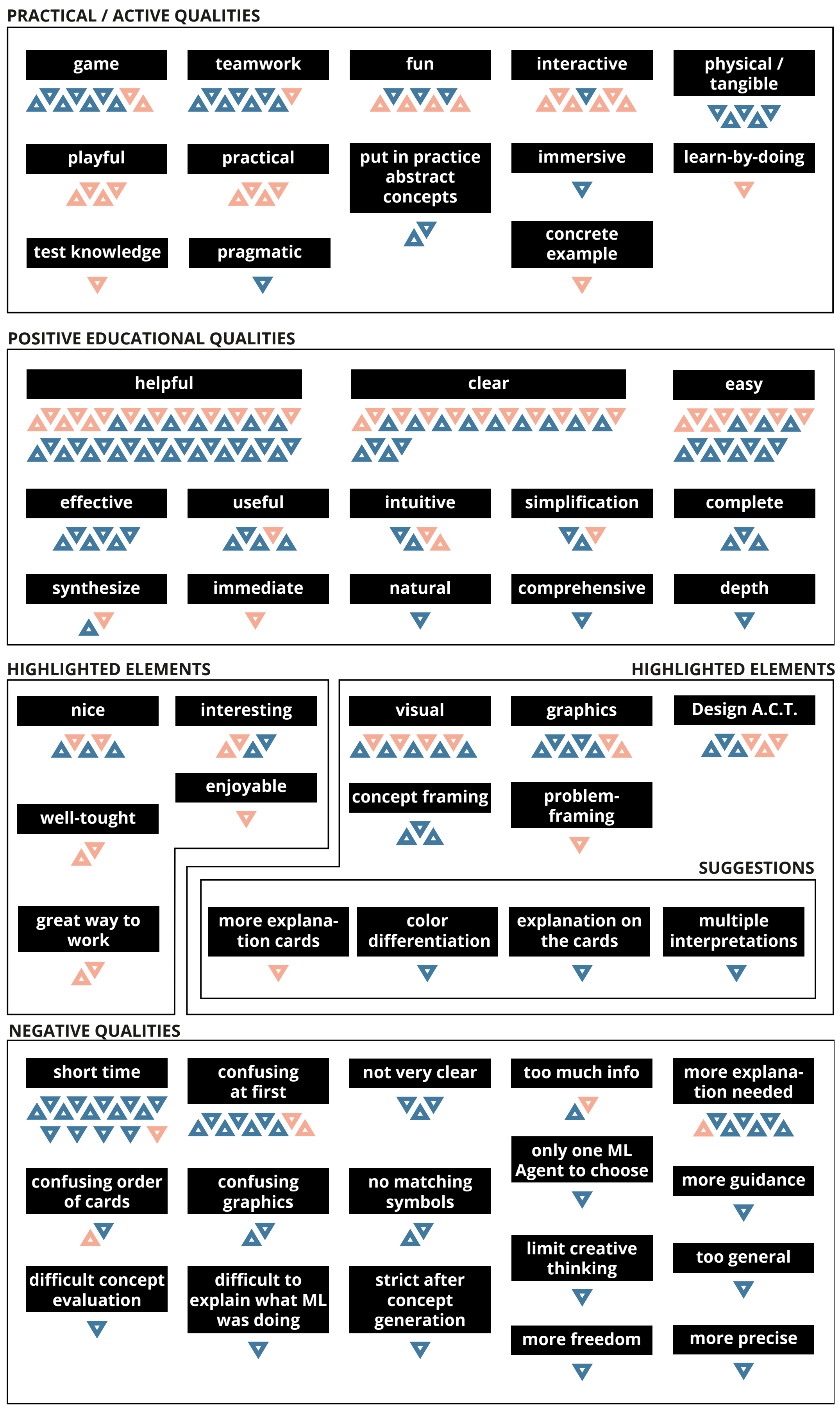


Fig. 6.22 | Summary of the content analysis of students' comments on the Concept Building Blocks.

how the ML capabilities were applied) are not often encountered, which is why, for educational purposes, "older" examples can be better than the latest ones. However, updating the resources and offering several ones, also to portray different purposes (especially in terms of responsibility), are crucial points to take into consideration.
Additionally, useful suggestions were also elicited. Although the graphics were mainly appreciated, some improvements could include color-coding the *ML Agents* (even if it would conflict with the CBB tool) or adding animations to simulate the functioning. Indeed, the visual support did not seem to have impacted comprehension.
Finally, from the point of view of students' fruition of *ML Agents*, the frontal presentation, the (possibly) collective discussion, and the test of the acquired notions before the design activity proved functional. As supported by different comments, this preparatory phase was essential to gather the necessary knowledge to apply afterward. Indeed, while developing their ideas, very few students consulted the booklet collecting all the *ML Agents*, and not many doubts emerged.

6.3.3.2 Concept Building Blocks

Materializing the *ML Designerly Taxonomy*, the CBB plays a central part in the educational method, and a lot of related comments were collected, as synthesized in Fig. 6.22. Despite the difficult experience of the condensed workshop, the majority testified that the tool was successful in its intent. Most commonly, it was recognized as helpful, clear, and easy. Its effectiveness is also remarked by the fact that it was described as intuitive, immediate, enjoyable, well-thought, complete, deep, and comprehensive. One even stated, *"It seemed natural to develop the ML Hero idea."*
According to several students, its character and the enabled design activities worked best. In relation to the former, the CBB has often been defined as a game (despite this word was never used by the author in any workshop), and using it was perceived as fun and playful, generating interactive and immersive design experiences.
In particular, the physical and tangible nature of the materials favored discussions, reflections, and teamwork in a dramatically better way with respect to the digital version. It was especially evident during the Polimi workshop, as the hybrid nature of the tool proved the flaws of a virtual-based experience (like having difficulties with the platform or being unable to control all the aspects of the concept at a glance). While the physical cards were still helpful in promoting collaboration and providing suggestions, the lack of the other components threatened the smoothness of the in-presence activity.
In practical terms, the CBB was mostly appreciated because of its support in the generative phase of the design process. Specifically, it was recognized as useful for portraying alternatives and possibilities, thus facilitating initial brainstorming to shape and sharpen ideas, allowing to get to a possible solution faster. Of course, it helped to include ML-related features in a project and encouraged considering different perspectives to generate meaningful artifacts and value for people. This was also confirmed by the positive responses that all groups (except those that exclusively tried the *responsibility model*) gave immediately after having created their first *ML Hero* (Fig. 6.23). In the extended educational experiences, it even translated into a unanimous maximum rating.

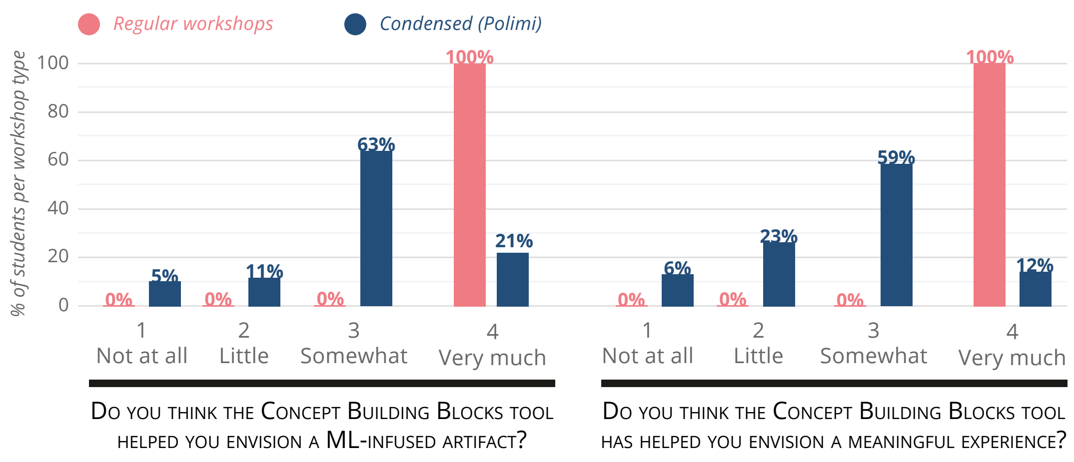


Fig. 6.23 | Students' assessment of the Concept Building Blocks performance.

Moreover, several participants affirmed that the CBB helped them reflect on their own ideas. They used it to analyze and evaluate their concepts, to find possible human-ML relationships, and understand the overall usefulness, feasibility, and coherence of the *ML Hero* they could specify as they pleased.
As a facilitator of comprehension, the tool also enabled the students to make sense of what they were doing and to understand *"the true meaning,"* the boundaries, and the limitations of ML systems. The procedural information that it was supposed to infer played a major role in this. Through the symbols on the cards, the participants had the possibility to uncover the relationships among the parts of the system and to build the structure of their *ML Hero*. They particularly valued the logic, organization, guidance, and process provided by the CBB as it was mostly perceived as inspiring but, at the same time, leaving the freedom to make personal decisions.
Except for Polimi students, who were only suggested a value-based approach or could undertake a personal one, most of the groups opted for the most familiar paths: the human-driven or problem-based one, the latter proving more effective in leading to the final concept (Fig. 6.24). A highly common feature of the participants' design processes is that they were mainly iterative, as students adapted and adjusted their ideas as they went on. Even if few concretized a completely personalized path in the end, several commented that they felt free to change their approach during the

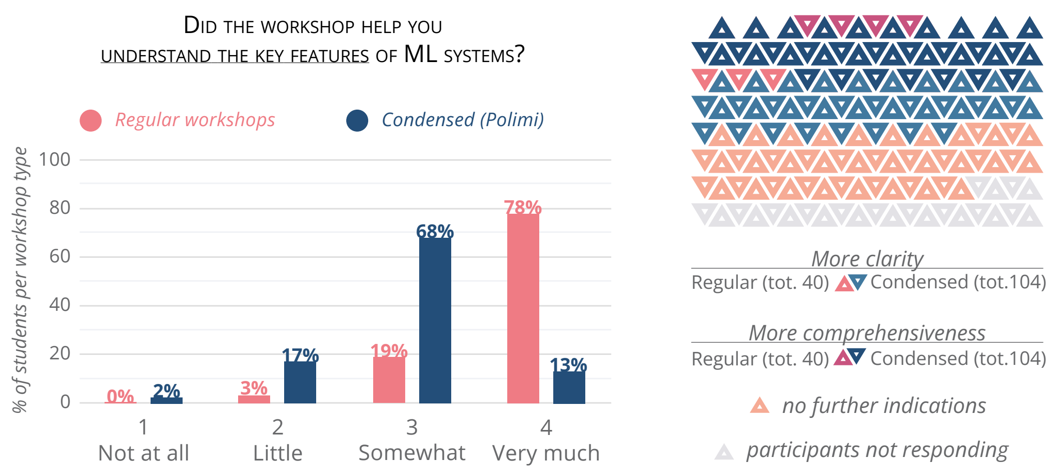


Fig. 6.24 | Students' description of the design process followed with the Concept Building Blocks.

design process, to better match their way of thinking, and only a couple warned about the possible creativity limitation of this procedure. In particular, one identified the actual scope of the tool, saying that it is *"great for an initial idea generation phase, but it could be somehow too strict after."*

Other significant elements of the CBB, possibly those that the design students felt most familiar with, were the *concept framing space* with the identification of the design A.C.T., the graphics, and the visual representation of the ideas.

Only few comments declared the opposite, describing it as confusing at first (from a content and graphical perspective), not very clear, or to be better explained. Again, the limited time accentuated these difficulties, but three aspects are worthy of attention. The first, elicited by a couple of comments, is that students were not always able to select the best cards with matching symbols. Unfortunately, this happened in the Polimi workshop, and the facilitator was not asked to clarify the situation. Being very interesting for the research, the tool is built to add further cards or to highlight previously unforeseen links. Still, in all the workshops where a closer observation was possible, it never occurred (despite the encouragement of the researcher). As a couple of cards have been overwritten on the Miro boards, these might be the cases to which the comments refer. However, the edited cards ultimately coincide with existing ones.

Another relevant issue is the demand for further descriptions for the cards, which emerged in different ways in all the workshops. Some felt the necessity of more specifications of the *capabilities level* cards, while others wanted to better understand the differences between the *intent* concepts or to have more information about *inputs* and *outputs*. Although the freedom to interpret and specify the cards at will was conceived to be an integral part of the tool, the recurring requests raise the possibility of adding further explanation, maybe in a separate booklet.

The last issue, explicitly pointed out just by one student, relates to allowing the construction of more articulated *ML Hero* structures by selecting multiple cards per level. Even if enabling the framing of more realistic systems would represent an interesting development for the tool, the results show that, for a first approach, this level of simplification may actually work.

6.3.3.3 VALUable by Design Expansion

For the problematic conditions illustrated in the analysis of the ILOs related to the *responsibility model* (sections 6.3.2.6 – 8), the VDE has been the tool most easily exposed to criticism.

However, Fig. 6.25 shows that the unfavorable issues directly attributable to the tool are not as much as the positive ones.

To a large extent, the comments mirror those of the CBB. After all, it is an expansion of it and shares the same nature. Therefore, helpfulness and clarity are the most common qualities, but the tool has also been described as well-explained, immediate, intuitive, cool, enjoyable, and fun (among others). Peculiar to this expansion is its identification as an essential instrument for its ethical-related nature. It was considered a reflection starter, able to provide a general overview and *"perfect to start from scratch,"* but also a synthetic tool *"that really put everything together."* Like the CBB, it offers a process to support the design activity, especially giving it significant direction and orientation. It also facilitates understanding *"because it forces you to be critical on your work,"* translating into practice abstract concepts. Of course, the most prominent activities that this tool prompts are reflective ones. Namely, it has been appreciated because it helped envision problems that might be unique to ML systems or that would otherwise be overlooked, anticipate impacts, evaluate, test, iterate, rethink, or even evolve one's ideas. It was also said that it usefully introduced new perspectives, being *"eye-opening"* and promoting innovation.

Indeed, it was recognized as having a responsibly nuanced generative capability, one that enables the inclusion of key principles in the decision-making process, brings values to people's lives, can *"push the ML Hero further,"* and even *"create something that could be sustainable and acceptable for society"* out of a *"funny concept."*

On the negative side, possibly time-related issues hinder the clarity and effectiveness of the tool. Interestingly, it is both identified as too complex, difficult, and confusing and as too general, simplified, and superficial. Indeed, the proposed modalities, at the same time, leave space for freedom in the interpretation of *risks* and *values* (which might be disorienting) and do not go into granular details of technical issues, with a level of simplification that can be perceived as superficial.

Very few comments requested further explanations and examples also in this case, while only one student highlighted that *"Yes, I understood the risks, but I don't really get the solutions to solve them."* All other cases confirmed the assumption that, given the problem and understood the systemic dimension of the project, design students are inspired to autonomously look for ways to tackle the challenges. Indeed, it is significant to notice how the envisioned solutions did not limit to the technical or artifact dimensions but involved different people (not just developers or designers) in taking active roles within the ML-enabled system.

Finally, from a graphical perspective, it has been suggested that the VDE might benefit from a second board to be more self-explanatory like the CBB. This might address the difficulties of those that found the graphics confusing.

6.3.4 Overall workshop experience, knowledge assimilation, and participants' perception evolution

From a didactic point of view, almost all the workshops were fulfilling. Expectations have always been exceeded; the students' perceived attention level and participation in Graz and the responsible module in Nantes were high and respectively moderate and low in Madeira. In Polimi, it was harder to determine as the class was numerous. Despite for few clarification questions, there was no room for participation. In the front rows, some seemed highly attentive, and others were very tired, maybe from the previous lecture. In the other workshop in Nantes, instead, some students presented language problems; hence, their focus seemed limited during the presentation. However, as the design activities went by, the participants of all workshops (at least those with whom the researcher interacted, which does not include everybody in the Polimi experience) seemed rather interested and receptive to the contents. Exceptions were represented by the lower responsiveness of a part of the Nantes' class involved in the *consistency module* (with high language barriers) and by Graz students, who, on the contrary, always demonstrated great enthusiasm and worked hard during all three days.

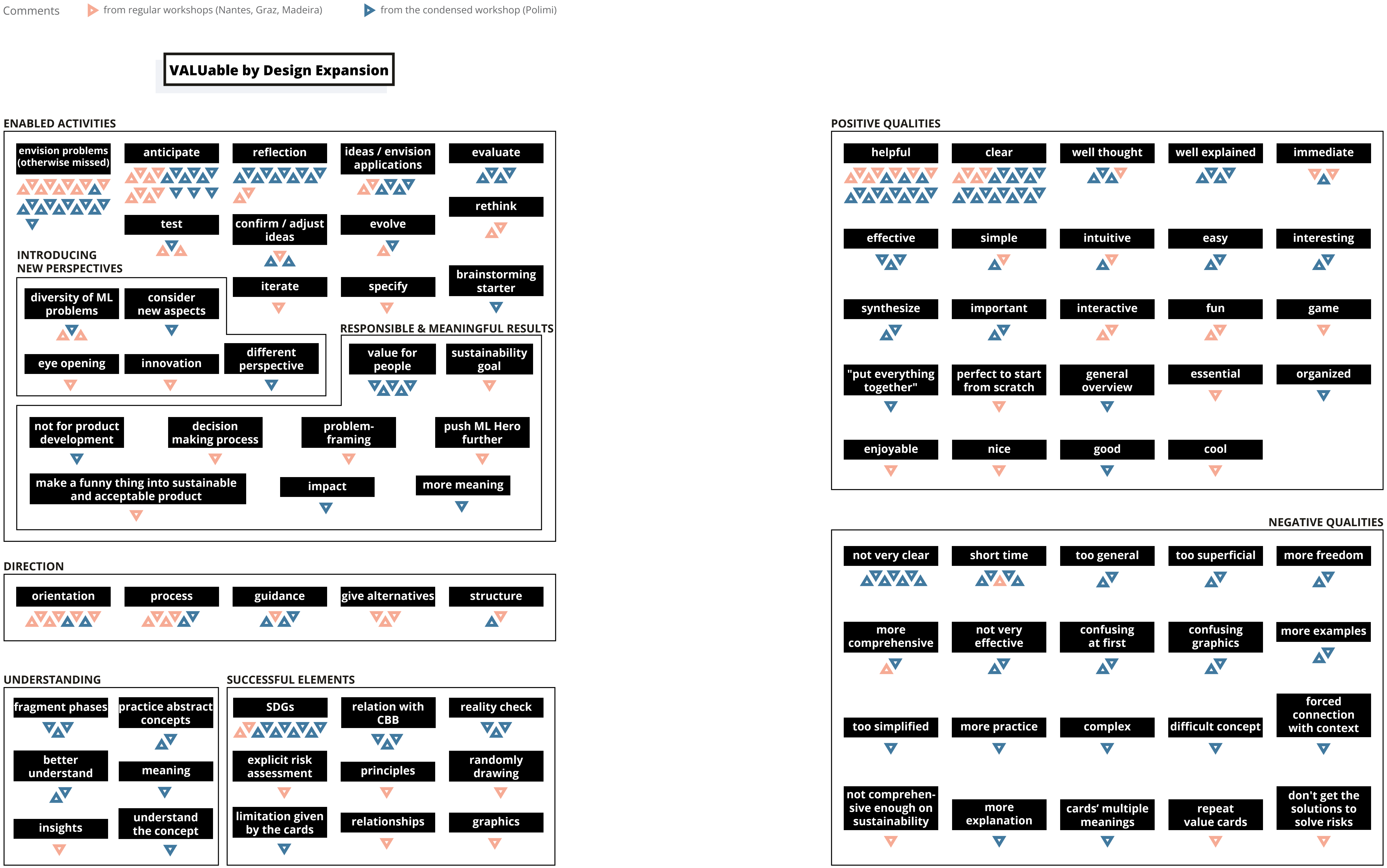


Fig. 6.25 | Summary of the content analysis of students' comments on the VALUable by Design Expansion.

Within an informal environment and with no pressure for a necessarily successful performance, almost all the participants got involved in the training activities, and the observed groups seemed to be engaged and having fun while developing their tasks. Even though no significant problems were found during the reviews, the reassuring or clarifying role of the facilitator was essential (as the Polimi case proves). These conditions generally led students to perceive an increase in their knowledge about ML (Fig. 6.26). Before the workshop, the greatest part of the participants knew nothing or little about ML, while after the two days of the *consistency model* in France, a student exclaimed, *"It's crazy how much we learned in such a short time!"* and in Graz, another said: *"My brother studies computer science, and now I feel like I can communicate with him."*

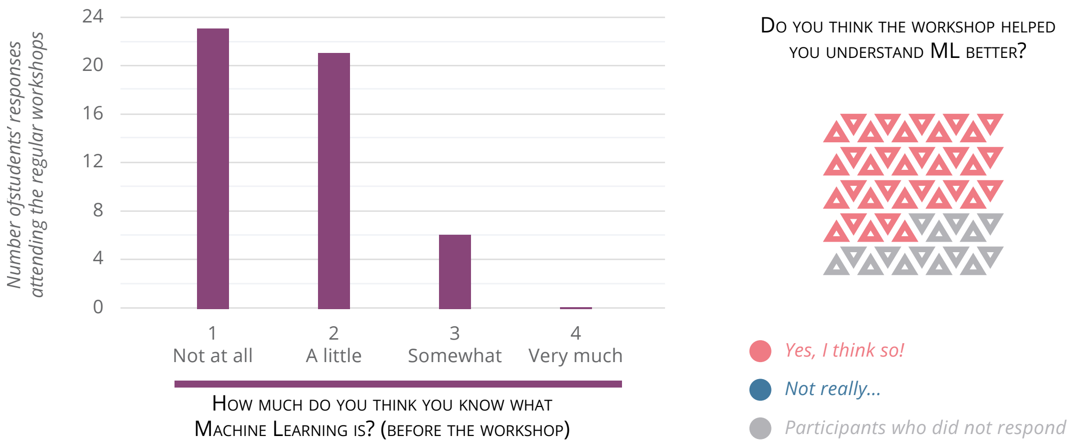


Fig. 6.26 | ML-related knowledge before the workshop and students' perception afterwards. Unfortunately, data from Polimi condensed workshop cannot be used because of a technical issue.

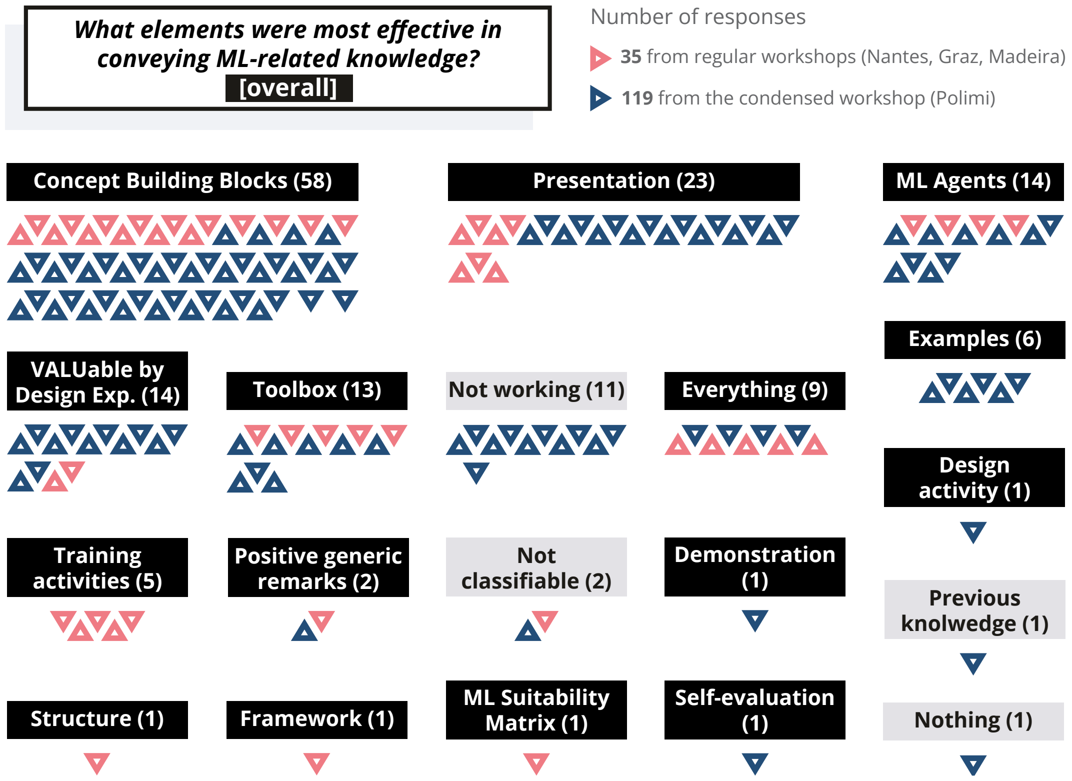


Fig. 6.27 | Summary of the content analysis of students' assessment of the elements that best conveyed ML knowledge.

Similarly to what was observed for the ILOs, it was esteemed that the CBB worked the best in conveying ML knowledge, followed by the introductory presentation, *ML Agents,* and the VDE (Fig. 6.27). Some comments even stated that everything was effective, while another (from the condensed workshop) said that the entire process needed a rethink.

Not much can be added with respect to the previous analyses. Besides some curiosities and an interest in the differences between ML, AI, and traditional programming, no issues were detected in the selected contents or transferring them. In the end, compelled or not to attend the workshop, most students approached the educational experience with a good predisposition (Fig. 6.28), and some finally commented that *"It has been the best course we had so far"* (from the *consistency module* in Nantes), or *"the best workshop I have attended"* (both from Nantes and Graz), or again *"I saw a lot of generative toolkits during my bachelor in design, and I sincerely think that yours is the best I happened to come into, both for the attention to detail and ease of use"* (from a participant of Polimi workshop). Far from suggesting that the educational models and tools are perfect as they are, these indications proved at least that the proposed activities, contents, and modalities are appropriate for the target audience, with no big differences connected to their background or program they are enrolled in.

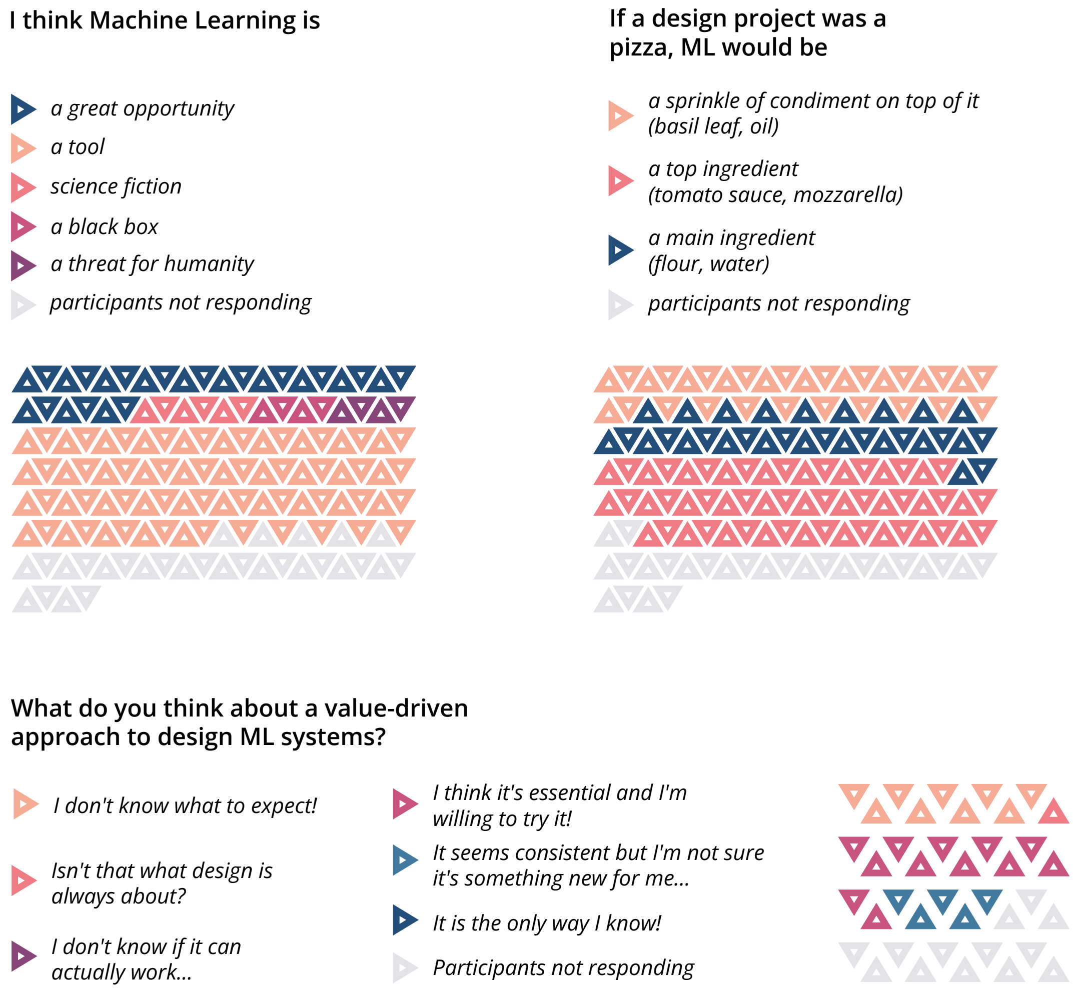


Fig. 6.28 | Students' preconceptions about ML and value-driven approach.

6.3.5 Discussing the results

6.3.5.1 Reflections from the different contexts and audiences

If, in theory, one may hypothesize that ML knowledge could benefit any kind of design subdiscipline as this technology has the potential to permeate every aspect of human life, developing an educational method that can address any design student (as long as they have assimilated the basis of the design approach) might seem a too broad assumption. For this, practical experimentations in different contexts and with varied audiences were essential. These confirmed what was already observed in the *Superpowered Museums* workshop (section 5.3) with an addition. Even if the educational activities can easily adapt to the didactic necessities, the proposed approach is sufficiently versatile to be assimilated by design students from different specializations, having a generic brief to unite them: add quality to life on Earth. Indeed, for the vast majority of the participants, **no problems in the assimilation of ML knowledge** were detected.
One main difference was observed between bachelor and master students, though. While the younger proceeded faster, as to "complete the game," master students were more reflective and produced more thorough and attentive considerations. This happened especially in Graz, where the participants spent some time discussing whether their *ML Heroes* were actually addressing ML problems or if other solutions might be more feasible. They also took advantage of the intimate context to exchange views and advice with other groups during the design phase, and a group started a comprehensive brainstorming to identify the best impact to set the design A.C.T. (at the very beginning, when they were unaware of the *responsibility module*). Albeit in a shorter time frame, the same critical thinking was manifest in the French class dealing with the *responsibility module* and in Polimi, thanks to the final reflections students had to deliver. This does not mean that bachelor students were not able to produce quality work or to reflect on their own concepts. What changes the most is the aptitude, always considering that no credits or evaluation depended on the results they produced.
Another confirmation concerns the **studio format.** This was consistent with the educational purposes not only for the practical application of knowledge – however appreciated – but for two additional factors. The first is the preferable dimension of the class. A comment from the Polimi workshop noted that the proposed activities were probably more suitable with fewer students. At the same time, another from the *responsibility* workshop in Nantes would have appreciated having the whole class participate (instead, they were ten working in groups of two or three). Certainly, a good balance of participants is needed to ensure exchange within and between groups. In fact, more discussion and peer review activities were requested by a couple of participants of the condensed workshop, in which these were not included. With more time available, also the other experiences would have benefited from more thorough peer review activities as, after the presentations, they only took the time to express their opinion on the Wooclap platform, and it was difficult to foster a proper discussion. The second essential aspect of the studio format is the support of a facilitator. Despite trying to set up educational models and tools that could be easily managed by the students without much teacher/researcher intervention, like in all the previous experimentations, the presence of a figure to turn to for confirmation, feedback, or clarification has always been sought.
Ultimately, the slightly different iterations of the educational models provided indications about the implementation of a holistic approach. The two separate models, as tested in Nantes, worked fine, validating the modularity of the method. However, it expressed its best in the integrated modules and, given the significance attributed to SDGs and principles to drive the concept framing (especially emerged from the Polimi workshop), merging the disciplinary perspectives from the very beginning could be the best solution.

6.3.5.2 Beneficial aspects and limitations

No matter how trivial or if already implemented somehow, achieving concepts that are consistent with the technology they integrate and developed with a designerly and responsible approach as a first attempt at dealing with ML is not a foregone achievement. This not only confirms that designers have the potential to be involved in teams developing ML applications, but it also suggests that the educational models positively contributed to their results. From the collected comments and observations, different strengths and room for improvement can be identified.
Starting from the didactic components (apart from the tools already discussed exhaustively), the Wooclap-based activities worked surprisingly well. Inserted to fulfill research objectives and because suggested by innovative teaching practices (Sancassani et al., 2019), the initial expectation was uncertain, having concerns that they would be perceived as forced or too trivial for the audience. On the contrary, in all the workshops in which they were proposed, students underlined the effectiveness of these formative tests in supporting their learning process. Some would have liked to have more, and it was suggested to make an internal competition out of it to elicit faster answers in a limited span of time. For the research goals, well-thought responses, possibly from the entire class, were needed to have a better overview. For this reason, annoying downtime occurred. However, for educational purposes, as making mistakes is part of learning, enhancing these activities to be more engaging, and finding a good balance between accuracy and time could ensure a better experience and encourage more people to promptly participate.
The elaborated posters and the storyboards were not mentioned in any feedback. Suitable for being developed in few hours, they probably failed to foster deeper reflections. In fact, the *ML Hero* visual representation and synthesis (with the title and catching phrase) tended to take longer than suggested or to be superficial, and only in some cases, it was possible to discern attention to metaphorical details.
Accordingly, the storyboard could only scratch the surface of the project, and some naiveté in the materialization of the *ML Heroes* emerged, like in the case of a translator integrated into AR-enabled glasses. If the concept level reached in the workshops was in line with what was needed for the research, it can still be vastly improved. For instance, by requiring more work from students, as advocated by a comment, it would be possible to leverage other designerly practices, like journey mapping or prototyping, and to introduce further moments for reviews to focus on the quality and depth of the output and not just on finishing it on time. This kind of implementation

would be even more beneficial for the *responsibility model*, which, at the moment, is limited to the completion of the CBB and VDE boards, while there are plenty of opportunities to push ethical reflection further. In this regard, an extension of peer review moments (in the form of tests or user research activities) could also help.

Another element that would deserve improvement is the *ML Suitability Matrix*. This tool should support a fundamental decision in the design process: understand whether or not it is appropriate to integrate ML into a project. In the proposed form, the matrix was appraised for its functionality and adaptability and because it offered many possibilities. It was used not just to evaluate one's concept but also to correct imperfections or generate new ideas. Although it did not receive any open criticism, it was observed how the *ML Suitability Matrix* played a secondary role in the development of *ML Heroes,* and not all the groups really considered it. Unlike the other tools, it did not offer clear guidance; thus, it might benefit from new mechanics and further research to increase its usefulness and depth.

Finally, all the workshops demonstrated the **value of collaborative and interactive activities enabled by tangible tools**, which is to be counted among the strengths of the method. In particular, these very practical experiments may be the key to students' feeling of having *"learned a lot."* However, as a comment pointed out, *"It would be beautiful to really design something thanks to these tools," or rather, enabling design students to actually play with ML systems to create some concrete artifact, even if only for learning purposes, could dramatically increase their understanding. For that, facilitated prototyping tools, requiring no coding skills but leaving the freedom to test different ML capabilities, would be ideal.*

This experts' perspective box reflects on the outputs of the research. The CBB and VDE tools were introduced with the support of a video presentation, and their application, as well as the broader educational models, were explained. Generally, the interviewees appreciated the work developed and the results. The most important value they recognized is the

RELEVANCE OF THE RESEARCH SCOPE

Krogh affirmed it is *"extremely important work,"* and the main reason behind that is the effort for synthesis and comprehensiveness, as Forlizzi pointed out. In Redström's words, *"It's both about articulating an approach to this complex issue but also structuring it as a kind of learning experience where you have to take the students from A to B, and it makes sense as well."* Indeed, the relevance of the research lies in the foundational level of translation it addresses. As Graves Petersen said: *"We need tools for experimentation and design educators. We need this research."* Especially, we need to introduce design students to ML as a new material for them to play with, which is why *"getting them exposed to the space, the logics of that space [...] and planting seeds"* is so important in Sharp's view.
Other spotted strengths were the physicality of the tools, their game structure, the interactivity and discussion they enable, the reflections they support, and the overall designerly approach to it. However, as these qualities were previously discussed from students' feedback, here, it is interesting to focus on the major points of interest that the experts identified to open room for improvement and further work.

Overall, *"This is way better than magic, which is usually what you see."* Yet, as anticipated in the *In Experts' Words* box in Chapter 4, Zimmerman noticed that some aspects should be deepened in order to provide students with enough information to generate doable concepts. In particular, more reflections should be dedicated to

FEASIBILITY

Not willing to diminish the importance of ethics, he pointed out that *"there are these other barriers that are actually even harder than the ethical barriers."* Indeed, considering that *"85% of projects never get deployed, [...] you're focused a lot on the ethical issues, which really is only about those 15%, but it's not attending to the failure before that."* So, what is needed is a clear distinction between *"This is what we need to do. This is what we can do"* (Redström).

While the *ML Suitability Matrix* tackles the identification of problems that make sense to be addressed using ML systems, another important part of the picture should be clearly included:

WHO IS MAKING ML SYSTEMS?

As Zimmerman observed, the research is *"thinking about end user value, but not stakeholder value."* Much attention is dedicated to the people who are going to interact with the systems, gaining some benefit or being otherwise affected, while the needs and requirements of other stakeholders are less explicitly addressed. Indeed, he stated: *"The hard part is to make you happy in a way that also returns value back to the organization that's doing it and to learn where the co-creation of value opportunity is, where everybody wins."* With more time at disposal for the educational experiences, it would be optimal to support students in developing reasonable concepts, like in common industrial design tradition.

On the one hand, similar to any other production method, they should learn to identify what is feasible from a:

FINANCIAL PERSPECTIVE

understanding the dynamics and the limitations to define whether *"the cost to create that value is totally outweighing the value it produces"* (Zimmerman). To do so, a possibility is to provide *"examples that have proven to be commercially viable. [...] Just bringing in the examples of how has that been instantiated in the world and pulling out: what was their model performance? How did it generate value back to the service that was doing this? That's also bound up in what we mean by capability"* (Zimmerman).

On the other, the viability of an idea is also based on the

TECHNICAL PERSPECTIVE

In other terms, designers should be able to think about the manufacturability of a concept, to find the *"harmony between what can be made and what would be experienced as valuable," and acknowledge "how does it fit into the product ecologies that are in existence?"* (Zimmerman).

Hence, again with more time at disposal, the envisioned educational experience might benefit from

TAKING LEARNING-BY-DOING TO THE NEXT LEVEL

While the CBB tool provides a controlled space for the ideation of ML systems that stay at an abstract level, it would be informative for design students to see *"what happens when I press compile [...] to listen to what the system comes back [...] to put* *a stethoscope to the system"* (Krogh). *"You need to know the material, you need to be able to experiment with it, to know if I do this, what does it do?"* (Graves Petersen) and to identify the *"details that make systems fail"* (Krogh). A more direct relationship with ML systems' definition and performances could give design students a better idea of what can be reasonably done from several points of view (including technical and financial ones), possibly overcoming inappropriate expectations when they have to move out of the guided path of building the ML structure and write design specifications on the dedicated cards, a *"weak link"* in Sharp's opinion.

A crucial objective to achieve would then be to enable design students to communicate with their technical counterparts. As Graves Petersen underlined, the premises are already there: *"We have a shared tool, there are clear representations about the different elements, you can work with it iteratively, you can experiment, it's a shared reference, right? It's out there on the table. Multiple people can have discussions around this. It has a lot of very good qualities in it."*
Yet, if designers do not have the possibility to work jointly with engineers, computer, or data scientists, they should have a way to

HAND OVER TO ML EXPERTS

so that the developed ideas can find their way toward real implementation. On this matter, Krogh suggested, *"at the end of the learning program, to present the stuff to AI and ML experts. So, not just have designers discuss with designers about what they got, but actually have feedback from those people they would be communicating with, and say, «What do you see? What is actually happening here? What are the details you think could make that fail or break down? Where are the hard parts of that system? Where does it really hurt? »"*
And Redström also noted that the final outcome of the designers' ideation process *"needs to be something that we can hand over to the software team, and they know what they're going to do with it. So, it has to be narrow. And how do we deal with that as designers? I think there's more to the setup of the game than meets the eye and it has to deal with the decision-making that takes place there. [...] You are very, very close to something that could be a handover moment here. [...] Since you have this kind of really core ML stuff in there as well, it wouldn't be too difficult to interface, for instance the Suitability Matrix, with what you might hand off to the software engineering team."*
While the research is about breath, opening the space for envisioning possibilities and starting discussion, a further step for evolving and expanding the learning activities would certainly look at more depth. Therefore, one of the major issues would be understanding *"what does the boundary objects look like in the negotiation between designers and ML experts?"* (Krogh), *"and how is this documented? I mean, the game and the interaction are great, but let's add a team with data scientists to take this back to, how would they report what they did?"* (Forlizzi). In particular, thinking ahead of what is covered in the research, she reflects on the fact that design students *"will work in multidisciplinary teams, and a lot of times, designers have to have evidence or rationale for why they made the decisions they made. So, linking these to documentation forms that other people on the team could use, even a design brief, could be really interesting to help the data scientists think about the space of opportunity."*

While these are the major areas for improvement that emerged during the interviews, further suggestions, very much aligned with the actual goals and desiderata of the research, also came up. For instance, Sharp advocated the exploration of *"other shapes for the curriculum besides the two-day workshop. So, what does this look like as a one-week workshop? Or what does this look like as a semester-long course?"* Or, additionally, how could it be extended to other *"people who are designing in industry, who want to learn more about ML and AI?"* (Forlizzi). How would the educational models change?
Again, to ensure the knowledge transfer, *"What is the takeaway? What can you give them on the other side of it? Is there a checklist or something you can give them on both [...] how machine learning works and where the risks lie? Like [...] a web-based tool they can come back to that really captures a lot of this stuff I think makes sense."* (Sharp)
And possibilities multiply if the final objective is adjusted. For example, *"it could help to have a domain to work with [...] to have some kind of frame that creates a tension between the context and the project"*, like agriculture or healthcare (Redström), or *"the card sets can be really tailored toward particular applications"* like image and text manipulation (Sharp).
Ultimately, the precious contributions of the expert design researchers and educators involved confirmed that the path undertaken is valuable, but it is just the beginning and

RESEARCH MUST GO ON

In Zimmerman's words: *"You've taken a big step [...], but we can't jump to the end, right? It's one hundred little steps to get there. So, I think you've done a great job of breaking this down, pulling out some meaningful components, and pushing our community forward. I don't think it's your job to finish, but just to detail what else is missing from this. This isn't the work of one. This is the work of the community."*

6.4 Framing an interdisciplinary educational method

The illustrated experimentations represent the last action of the iterative process of the doctoral research, which here comes to a conclusion. From the broad aim of finding designerly ways to translate ML for design education, the investigation constructively explored the research questions, collecting several ingredients to build a synthesizing recipe. From contents and skills, forms and languages, tools and models, the action research identified, tested, evaluated, and made sense of the different components of a didactic translation to finally determine the qualities of an interdisciplinary educational method. To meet time constraints, this is the ultimate output of the Ph.D. research that, in line with Buchanan's argument (1999) (Muratovski, 2016), tries to find a balance and develop a proper relationship between theory and practice.
A **methodological synthesis** was inferred based on the pedagogical frameworks identified in 6.1.1 and the practical experiments in 6.3. Starting as an adaptation of Gagné's events of instruction and informed by RRI principles as well as the results of the workshops, some key moments of a desirable educational activity to merge ML, ethics, and design knowledge are portrayed in Fig. 6.29.

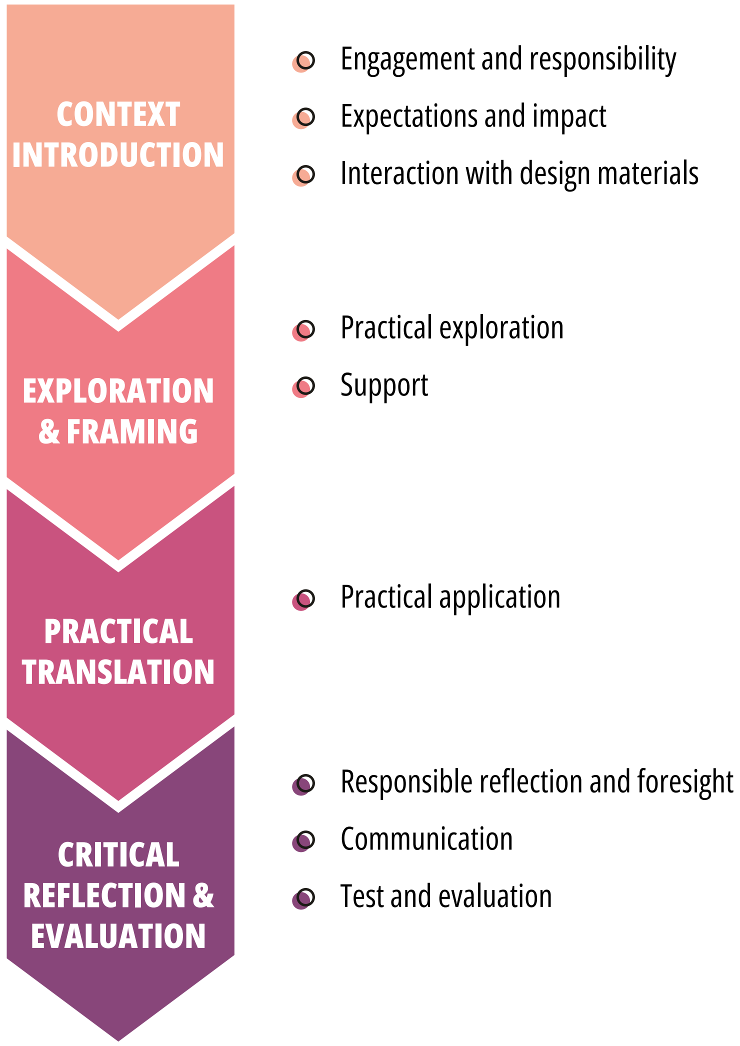


Fig. 6.29 | Educational method layout.

> 1. Engagement and responsibility. With a learner-centered approach, the first step embraces Gagné's perspective, focusing on capturing and increasing students' attention to the subject matter. Already emphasizing the holistic nature of the method, it should also move a bit further, making them understand their role and responsibility as designers. In the workshops, this was applied by problematizing the current development of ML-infused artifacts and involving the participants to make a difference (as interns of the *ML Hero Agency*) through experimentations to change technology-driven trends. Of course, the narrative can vary according to the purposes of the educational activity, but it should comprehend both practical and ethical commitment.

> 2. Expectations and impact. Gagné intends this stage to inform students of the intended learning outcomes. Still, again it might benefit from an infusion of the RRI approach broadening the expectations to include the challenges to be addressed and the motivations behind them, in addition to the formative requirements. To have fully engaged students, in fact, it is important that they feel involved and are clear about the purpose of the activity, just as *"designing for the right impact"* (von Schomberg, 2013) makes them aware of the relevance of the solution they should envision.

> 3. Interaction with design materials. As design is an applied discipline, this point aims to put an accent on the importance of an interactive approach to knowledge transfer (which proved very successful in the workshops) and on the fact that every nurtured notion or skill becomes part of designers' materials when they operate. It includes theoretical contents (from ML and ethics), tools (like the *ML Agents*, CBB, and *ML Suitability Matrix*), values, examples, and case studies. The activation of metaphors, as in Schön's perspective (1983), is also part of this step, as opposed to previous knowledge (which might not be available as the educational method intends to introduce new disciplinary knowledge for the target audience). These can be presented to students or left to be discovered independently through research, observation, experiences, or prompts. Once again, what matters is for them to be actively involved in this learning process.

> 4. Practical exploration. Then, the freshly acquired materials should be put into practice to familiarize and start making sense of the possible ways to employ and take advantage of them.

> 5. Support. As highlighted in the discussion of the results of the experimental workshops, the figure of a facilitator reassuring, giving feedback, and orienting students in their explorations with new materials is indispensable. Marked as a specific point, as Gangé does, support extends throughout all the practical activities, like in any studio format.

> 6. Practical application. Once students become acquainted with the new materials, at least a second iteration of applying the tools and knowledge gained can help designers to master them better and make them stick for future implementations. This expedient was not integrated into the tested didactic experiences but responds to the emerged necessity to make more practice with them to facilitate their internalization and prevent forgetting important aspects.

> 7. Responsible reflection and foresight. Inevitably, the generative phase must be followed by one explicitly devoted to questioning and weighing one's ideas. This entails reflecting on UX and ethical facets, steering the concepts towards desirable impacts with a value-driven approach, and anticipating possibly positive and negative outcomes to overcome or limit certain risks to happen. For this purpose, the VDE can be a helpful support.

> 8. Communication. Having the right vocabulary and means to properly communicate one's idea is essential for engaging with colleagues, users, or experts from different specializations. Working across disciplines necessarily requires discussing with different types of people to get inspiration, feedback, instructions, test, or evaluate a concept under development. Thus, to be able to properly communicate, this skill might be expressly cultivated as part of the educational method, and it permeates both the previous and subsequent steps as these activities can rarely be accomplished alone. As highlighted by the experts interviewed, this step should be reinforced by appropriate envisioned modalities for a successful handover.

> 9. Test and evaluation. Finally, as usual in design, the practical application needs to be tested and evaluated to complete the learning process and assimilate how a responsible approach can unfold. Indeed, a designerly educational process can hardly culminate in consolidation and generalization, as the knowledge and capabilities acquired and developed during the design activities might not be extendable to other contexts. However, in their uniqueness, if properly elaborated they can be useful references for future experiences.

In synthesis, the method can be framed in four main phases: **(i) an initial context introduction** (including steps 1-3), **(ii) an exploration & framing stage** (steps 4-5), **(iii) a practical translation** (step 6), and **(iv) a critical reflection and evaluation** (steps 7-9). This structure reflects the problem-based pedagogical framework (Sancassani et al., 2019) and redefines it from a more project-centered perspective.

Still imbued with the constructivist spirit, this framed method cannot be considered a final achievement. Instead, it represents the starting point for new experimentations in pursuit of further validation, improvement, and alternatives.

TO SUM UP

The chapter addresses RQ4: ***Which design education method can support the conceptualization of ML-infused solutions?***

- To test and validate tools and models, some essential components to form an educational method were identified with the support of pedagogical theories, always maintaining the twofold structure determined by the *consistency* [C] and *responsibility* [R] requirements.
  - The ILOs are more finely detailed in association with Dublin descriptors and Bloom's taxonomy parameters (in parenthesis). They can be summarized in [C1] Understand the core characteristics of ML systems *(knowledge & understanding | understand & remember)*, [C2] Understand ML capabilities *(knowledge & understanding + apply | understand, remember, apply, analyze),* [C3] Identify what problems can be solved with ML systems *(applying knowledge and understanding, making judgments | apply, evaluate),* [C4] Generate relevant, consistent, and effective design concepts including ML systems *(applying knowledge and understanding | apply + create),* [R1] Understand ML systems as socio-technical systems and their capabilities *(knowledge & understanding + apply | understand, remember, apply, analyze),* [R2] Identify and use values to drive the design of ML systems (*knowledge & understanding + apply | understand, apply)*, [R3] Identify and anticipate possible impacts of ML systems in practical, personal, social, cultural, and eco-systemic dimensions *(applying knowledge and understanding | apply),* [R4] Generate ethically acceptable, sustainable, and desirable design concepts including ML systems *(applying knowledge and understanding | apply + create)*
  - Based on Bloom's learning typologies, suitable assessment processes can be identified. Quick, informal, structured, and semi-structured formative tests (e.g., multiple choice or brief open questions) are provided for ILOs in the domain of remembering, understanding, and (partially) applying. For more reflective ones, project development is the most natural learning environment

for design students. In this regard, collective presentations, self, and peer evaluation on predefined parameters contribute to the deeper and more lasting acquisition of knowledge, skills, and competencies, including effective communication. Observation and mentoring complement these activities.

- Teaching-learning activities are shaped by a constructivist approach, Gagné's events of instruction, and a project-centered pedagogical framework.

- In parallel, an evaluation research model, envisioned as a multiple case analysis with a mixed-method strategy for data collection and processing, was selected for validation.
- Two separate and complementary educational models were formally outlined, building on the insights gathered from the previous research steps. The *consistency* and *responsibility models* were first individually piloted in two classes of the École de Design Nantes Atlantique), respectively, in 2-day and 3-hour workshops. Then, an integrated model was envisioned and tested in 3-day workshops at FH Joanneum University in Graz and Universidade da Madeira, and in a condensed 3-hour format in Politecnico di Milano. They addressed audiences from different design backgrounds, at different levels of education (bachelor and master), and with varying numbers of students involved. Among others, the findings from the application of these research outputs include:
    - All the ILOs were quite positively met, with more successful results in the workshops adhering to the educational models. The condensed one, presenting several differences, was a useful counterevidence for some of the basic assumptions.
    - The tools were helpful for their purposes. Feedback on *ML Agents* confirmed the primary role of examples in understanding the subject. Central to the design activities, the CBB was particularly appreciated for the guidance and process provided, which were inspiring but leaving the freedom to make personal decisions. The VDE successfully elicited reflection and oriented the concept development toward responsible solutions.

    - Most participants felt an increase in their knowledge about ML and underlined the effectiveness of the formative tests in supporting their learning process.
    - The tools' physical nature and playful interaction are among the most valuable features of the method, as they favor discussion and collaboration.
    - The proposed approach proved versatile to be assimilated by design students from different specializations, with only a difference in aptitude between undergraduate and graduate students, more reflexive in the second case.
    - The studio format was consistent with the educational purposes not only for the practical application of knowledge but for the preferable dimension of the class and the essential support of a facilitator.
    - Modularity worked fine, but the holistic approach, merging the disciplinary perspectives from the very beginning, expressed the full potentialities of the educational models.

- To synthesize the results of the theoretical assumptions and practical experiments, some foundational steps for an educational method merging design, ML, and ethics knowledge were outlined. Adapting Gagné's events of instruction, they include 1. Engagement and responsibility; 2. Expectation & impact; 3. Interaction with design materials; 4. Practical exploration; 5. Support; 6. Practical application; 7. Responsible reflection & foresight; 8. Communication; 9. Test & evaluation. They can be framed in four main phases: (i) an initial *context introduction* (including steps 1-3), (ii) an *exploration & framing* stage (steps 4-5), (iii) a *practical translation* (step 6), and (iv) a *critical reflection and evaluation* (steps 7-9).

7. SETTING OUT A LONG JOURNEY AHEAD

C H A P T E R S E V E N

Along its course, the research attempted to **bring order** to a still-developing field, ML, to **make it more accessible** for (future) designers, design researchers, and educators who have the potential to improve the impacts of this flourishing technology.

To conclude, an overview of the research purpose and findings is provided (7.1), and the ways for it to spread and germinate are envisioned. In fact, as requested for doctoral research, it offers theoretical and practical contributions to design and education (7.2). Additionally, it provides several opportunities for improvement and future inquiry (7.3) as appropriate for its basic, exploratory, and iterative nature.

It is patent that this research is like a grain of sand in a boundless desert, and plenty of paths have yet to be explored or discovered. The hope is that it might be a ***spark to shed light on a matter of emerging interest*** for design and ***ignite the curiosity of other researchers*** to engage in a thriving interdisciplinary journey.

7

7.1 Overview of the research aim and findings

Recognizing how a designerly intervention might be beneficial for the development of ML-infused products and services and, at large, for people and their ecosystems, the research investigates **how to translate basic ML knowledge for design students.** Indeed, designers lack the necessary understanding, language, tools, and methods to handle this technology.

Dealing with an immature subject, characterized by a conspicuous lack of theory at the intersection of design and ML, the research started from a wide exploration of other disciplinary fields to deduce the raw materials to be transferred to design. Then, in a cyclic spiraling flow, it moved from theory to practice, inducing increasingly broader assumptions from concrete experimentations. As with most qualitative and applied works, the findings might not be generalizable. However, contributing to an emergent field of research, they convey interesting insights and are replicable and customizable at will.

Summarizing, the research has produced novel theoretical and practical outputs. In particular, the first are the *ML Designerly Taxonomy* and the *Responsible Cycle for ML Design*. They respectively systematize and synthesize ML knowledge to make it applicable in practice by bridging human-centered and technical perspectives and frame a process highlighting the interconnections between the foundational elements for a responsible ML design process (namely, principles, values, risks) which have been outlined as a result of a systematic content analysis.

Starting from the most concrete, the practical outputs include the tools to transfer and operationalize knowledge in support of the design process for educational purposes (*ML Agents*, CBB, and VDE); the *consistency* and *responsibility models* detailing how teaching-learning activities can unfold (from defining ILOs to the modalities for providing contents, tools, and eliciting results), and finally an educational method outlining the essential steps and features of a didactic experience to introduce new disciplinary materials to design students in an ethical way. With these broad indications, it is also meant to inspire further applications as more design practices and education programs are needed *"to enable a culture of design in which many people contribute to bringing about new and purposeful change"* (Auernhammer & Ford, 2022), to face a continuously evolving world.

Therefore, the research outputs can contribute to the design discipline and beyond, generating different outcomes and impacts in the short, medium, and long term.

7.2 Contribution to knowledge

7.2.1 Short-term outcomes

Certainly, the most immediate outcomes can be expected to develop within design education. Becoming publicly available, the *consistency*, *responsibility*, and *integrated* educational models and the related tools (as they are) can be easily implemented in new contexts and replicated by other researchers and educators to introduce general and basic knowledge about ML with a designerly and possibly responsible approach.
As demonstrated in Chapters 5 and 6, these are also modular and flexible. Thus, by adapting the educational activity's expectations and impacts, the examples provided, and the design brief, the models would be suitable for integrating various design courses. Indeed, the *Superpowered Museums* workshop (section 5.3) was a prototypical example of how this could be possible. Understanding whether and how ML can benefit the specific context and aim of an established design course or studio, the narrative around the technology can be modified, partially used, or enriched with further contents and tools to best fit. In the cited case, ubiquitous computing, pervasive interaction theories, and science fiction prototypes helped a smoother connection.
If definite design problems within peculiar design sub-disciplines might need very specific capabilities and applications of ML systems, in contrast with the general character expressly chosen for their validation, the developed tools and models can also be customized to address these particular necessities. For instance, if voice-based interaction modalities are to be explored to make some artifacts more accessible, a deeper focus on the ways ML systems handle speech and could be exploited would be necessary. This could reflect on the selection of contents but also in the CBB tool, as it should be limited in scope and maybe enriched with more grained cards related to the topic at hand.
Moreover, by exploring **multiple possibilities within the educational method**, the limited workshop format could be extended into longer self-standing didactic experiences. Interestingly, the models could also be the foundations for the development of a course that, with more extensive timelines, could strengthen students' agency to explore (also autonomously) ML possibilities, go deeper into the definition of their ML-infused artifacts (both from a designerly and ethical point of view), and apply the more iterative learning process suggested by the educational method. Of course, with the due research implementations, it could also push the conceptualization into prototyping and collaborating with ML experts.
These kinds of applications might happen in several types of design institutions. Of course, they naturally fit more technical universities like Politecnico di Milano and FH Joanneum University of Applied Sciences (among those in which the experiments have been successful). Still, they would also be interesting for more art-based ones. In fact, the research proved that, given a foundational knowledge of the design process, the models and tools are versatile enough to address audiences with different design specializations and at different levels. Intuitively, UX, digital, and interaction design programs might be the most interested in integrating these contents. In particular, the holistic approach would be especially beneficial for master students to stimulate their critical thinking, but its introductory nature makes it suitable for undergraduates as well.
Always remaining in the educational domain, further intriguing options can be explored to exploit the research outputs. As the translation targets design students who are not knowledgeable about ML, others in this category might benefit from the simplification developed. Following the European vision promoting the dissemination of ML knowledge, the models and tools might be **transferred to undergraduates or graduates in other fields**, from disciplines that are already experimenting extensively with this technology, like medicine and the arts, to others seemingly distant, like natural sciences or human studies, to create unusual and daring connections that might trigger new meaningful applications.
Another fascinating possibility is to test the knowledge transfer at **different levels of education,** such as in high school, middle school, or elementary school, or even address the **lay public**. In all of these cases adjustments should be made to the tools and approach, which may offer new possibilities. Unless explicitly functional for the course of study, the project-based activities can be overcome to focus on disseminating ML knowledge and related ethical concerns. For instance, the CBB can be used to represent and understand ML applications that we all know and use, while the VDE might raise awareness about the subtle implications of personalized contents or automated systems.

7.2.2 Medium and long-term impacts

Based on the envisioned research outcomes and considering its objectives – i.e., enabling design students to (i) imagine meaningful, consistent, and responsible design-driven solutions integrating ML, (ii) handle ML as an asset to address challenges, and (iii) communicate with experts in different disciplines – additional medium- and long-term impacts can be outlined.
For instance, design institutions might move towards more **interdisciplinary programs**, as advocated by many experts, like (Frascara, 2020; Friedman et al., 2019; Meyer & Norman, 2020; Voûte et al., 2020) to cite a few, and future designers with **competitive curricula** that respond to changing societal needs could reframe the current paradigm of professional work involving ML systems. Unhinging the technology-driven approach, they could promote more interdisciplinary tables, playing a bridging role between multiple competencies and possibly even setting the foundations for a new disciplinary branch.
In relation to the prospects for further dissemination, instead, even broader impacts could be expected in the long term. As widely said, ML has the power to impact any area of human activities. Therefore, highly **interdisciplinary teams** could steer its development toward flourishing results, and educating potential stakeholders about ML capabilities would represent a first step in creating synergies that might have positive repercussions in research or applications that **address major challenges.**

7.3 Limitations, future steps, and research opportunities

7.3.1 *Limitations*

Punctual methodological limitations related to the qualitative nature of the exploratory research in a complex and fuzzy domain have already been pointed out along the argumentation. However, one factor can be highlighted as the study's main weakness to which others correlate: **isolation**.

Due to different factors, including the lack of access to an established network of researchers working on this emerging topic and the pandemic situation, among others, the research opened to new contexts and perspectives – within the design field – only at a late stage, mainly for validation and critical reflection on the developed work.

In fact, despite being framed as interdisciplinary research, it mainly relied on the author's capability to interpret and synthesize knowledge from different fields based on her subjective sensitivity. Not influenced by the mindset or ways of proceeding peculiar to other disciplines, the research can be considered through design. However, outside of internal reviews, it lacks an observer triangulation to increase the reliability of the results.

In the end, the **findings are highly situated** and the adopted **qualitative** methodologies, acceptable within a design context, are questionably useful to set a dialogue with more technical disciplines, which would have required more quantitative and scientific approaches.

Additionally, because of the time constraints and the possible extent of individual efforts, relevant aspects for the educational activities could not be implemented. It is the case of more practical experimentations with ML systems to complete the learning-by-doing approach and meet the students' desire for more hands-on examples, as well as the validation of one of the hypotheses of the research, putting design students and computer scientists and engineers specializing in ML in communication to test if the knowledge transfer enables a productive collaboration.

7.3.2 *Future steps*

A few more steps should be taken to crown the research efforts and spread its results. First, the materials and tools should be optimized according to feedback and reflections portrayed in Chapter 6. In addition to minor graphical adjustments and the inclusion of further specifications, it would also be relevant to unlock the possibility of constructing more complex ML-infused systems (better depicting how they really work), and to integrate elements to assess technical and financial feasibility (completing the systemic picture of the actors involved in the development and deployment of ML systems).

Once refined, careful strategies must be designed to **distribute the outputs outside of the research environment** and reach a wider community. In particular, the *ML Agents*, CBB, and VDE could be made available in their physical form and, especially, online as open-source resources to be explored and employed by design students, educators, researchers, and anyone who may find them useful for a personal consultation or to activate collaborations. For the web version, however, the author should thoroughly consider the limitations that a digital interface presents to the experience of the tools with respect to their physical counterparts. Still, she could also find ways to augment them with new interaction possibilities. Then, to increase their visibility, they could be indexed and shared through the design and AI/ML repositories that have recently risen (Kyshkan, 2020; Piet, 2020).

Closely related to this aspect, and to compensate for its major limitation, a **network** should be created for the research to resonate, be challenged, and broadened as it deals with a significant and urgent matter. Connections with design researchers, educators, and practitioners, but also with experts in the ML and computer ethics fields (within and outside the academy), could foster the undertaking of new research opportunities, which might include the application for grants. In this sense, the educational method, models, and tools could also be used to build Erasmus+ and Knowledge Alliance proposals.

Additionally, it would be worthwhile to pursue some of the research directions that have been identified since the beginning of the experiments, which could not be accomplished due to time or opportunity constraints, like those listed in the following.

7.3.3 *Research opportunities*

As can be expected from basic research, it leaves many open paths for further inquiry. Among them, two would primarily stand out, as the foundations on which to build them can be retrieved in the presented work. The first concerns one of the main reasons and objectives of translating ML knowledge to design education: **putting designers and ML experts in communication** to explore the results of such a fruitful interaction. This would entail a practical validation of the *operative knowledge* level of the *ML Designerly Taxonomy* to complement the educational experience. Taking into account the additional audience's background and skills, the educational models, tools, and requirements should evolve for both parties to exchange mutual knowledge.

Both in the case of interdisciplinary collaborations or for educational activities bounded within the design domain, another factor that could improve students' learning is the implementation of **practical experiences** to explore, interact, and exploit ML systems for design purposes. Prototyping tools and processes are one of the next frontiers in ML-related research not only to enhance learning by doing but also to offer designers the means to test and evaluate the performances of the envisioned systems faster and more iteratively. Moreover, practical experimentations would allow them to acquire a sense of what is feasible in developing ML-infused systems.

After design students have achieved a good understanding of ML capabilities and potentialities, other significant points to investigate include finding ways to stimulate and make them express their **transformative capability** as well as applying their skills and knowledge to actually **tackle relevant challenges for humanity**. Indeed, alongside the necessary comprehension of ML as a design material, design students should also be able to find new senses and uses for this technology, envisioning original, meaningful, and innovative solutions. To do so, the research indicated how metaphors can be helpful means to enhance creativity. Additionally, being exposed to

multidisciplinary teams to face real projects on sensitive matters could also put their cross-disciplinary, communication, and mediating skills to the test.
Designers' increased competencies in envisioning, handling, and prototyping ML systems that experts can develop could also give birth to interesting experimental **research projects** at the edge of disciplinary boundaries. In fact, once the conditions to create an environment where professionals from different disciplines are able to communicate and mutually understand each other are set, the focus can move to the objectives that these joint forces want to pursue.
To this end, further research opportunities lay in **third mission activities**. Still, a wide space for investigation concerns how design can orient and support ML systems' applications in the public and private sectors. Researching, working, and educating decision-makers, producers, developers, and designers toward more responsible and meaningful ML solutions could beneficially impact entire ecosystems.
Evidently, **a lot still can and needs to be done at the intersection of design, ML, and ethics.** The hope is that other researchers could benefit from this work and possibly join me in the attempt to make ML more accessible to non-experts who might steer its development for the flourishing of life on Earth.

BIBLIOGRAPHY

Bibliography

AI for Good. (n.d.). *SDG Data Catalog—AI for Good Foundation*. Retrieved 4 May 2022, from https://ai4good.org/what-we-do/sdg-data-catalog/

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

Amershi, S., Weld, D., Vorvoreanu, M., Fourney, A., Nushi, B., Collisson, P., Suh, J., Iqbal, S., Bennett, P. N., Inkpen, K., Teevan, J., Kikin-Gil, R., & Horvitz, E. (2019). Guidelines for Human-AI Interaction. *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, 1–13. https://doi.org/10.1145/3290605.3300233

Ammagamma. (2021). *Bestiario di intelligenza artificiale* (1st ed.). Franco Cosimo Panini Editore.

Annik, E. (2006). Lasciare tracce, favorire esperienze. In H. Höger (Ed.), *Design education: Studiare design attraverso lo studio per progetti ; Designausbildung konzipiert als Projektstudium* (pp. 82–89). Ed. Abitare Segesta.

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.

Biggs, J. (2003). *Teaching for Quality Learning at University* (2nd edition). SRHE and Open University Press.

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

Brey, P. A. E. (2012). Values in technology and disclosive computer ethics. In L. Floridi (Ed.), *The Cambridge handbook of information and computer ethics* (pp. 41–58). Cambridge University Press.

British Design Council. (2005). *Eleven lessons. A study of the design process*. www.designcouncil.org.uk

Buchanan, R. (1992). Wicked Problems in Design Thinking. *Design Issues*, *8*(2), 5. https://doi.org/10.2307/1511637

Buchanan, R. (2001). Design research and the new learning. *Design Issues*, *17*(4), 3–23.

Carroll, J. M. (2014). Human Computer Interaction—Brief intro. In *The Encyclopedia of Human-Computer Interaction* (2nd Ed.). https://www.interaction-design.org/literature/book/the-encyclopedia-of-human-computer-interaction-2nd-ed/human-computer-interaction-brief-intro

Christensen, B. T., & Ball, L. J. (2019). Building a discipline: Indicators of expansion, integration and consolidation in design research across four decades. *Design Studies*, *65*, 18–34. https://doi.org/10.1016/j.destud.2019.10.001

Clark, J. (Director). (2019, November 5). *AI is Your New Design Material at Amuse UX Conference*. https://www.youtube.com/watch?v=Tgzu351uDIc

Confalonieri, R., Coba, L., Wagner, B., & Besold, T. R. (2021). A historical perspective of explainable Artificial Intelligence. *WIREs Data Mining and Knowledge Discovery*, *11*(1), e1391. https://doi.org/10.1002/widm.1391

Creswell, J. W. (2014). *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches*. SAGE.

Cross, N. (1982). Designerly ways of knowing. *Design Studies*, *3*(4), 221–227. https://doi.org/10.1016/0142-694X(82)90040-0

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

Dalsgaard, P. (2017). Instruments of Inquiry: Understanding the Nature and Role of Tools in Design | Semantic Scholar. *International Journal of Design*, *11*(1), 21–33.

Dam, R. F., & Siang, T. Y. (2020). *10 Insightful Design Thinking Frameworks: A Quick Overview*. The Interaction Design Foundation. https://www.interaction-design.org/literature/article/design-thinking-a-quick-overview

Dam, R. F., & Siang, T. Y. (2021). *5 Stages in the Design Thinking Process*. The Interaction Design Foundation. https://www.interaction-design.org/literature/article/5-stages-in-the-design-thinking-process

Davis, M. (n.d.). *Introduction to Design Futures*. AIGA | the Professional Association for Design. Retrieved 2 March 2021, from https://www.aiga.org/aiga-design-futures/introduction-to-design-futures/

D'Ignazio, C. (2022). Creative data literacy: Bridging the gap between the data-haves and data-have nots. *Information Design Journal*, 6–18. https://doi.org/10.1075/idj.23.1.03dig

Dorst, K. (2011). The core of 'design thinking' and its application. *Design Studies*, *32*(6), 521–532. https://doi.org/10.1016/j.destud.2011.07.006

Dourish, P. (2001). *Where the Action Is: The Foundations of Embodied Interaction* (New Ed edition). The MIT Press.

Dourish, P., & Bell, G. (2014). “Resistance Is Futile”: Reading Science Fiction alongside Ubiquitous Computing. *Personal and Ubiquitous Computing*, *18*. https://doi.org/10.1007/s00779-013-0678-7

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*.

Findeli, A. (1994). Ethics, Aesthetics, and Design. *Design Issues*, *10*(2), 49–68. https://doi.org/10.2307/1511628

Findeli, A. (2001). Rethinking Design Education for the 21st Century: Theoretical, Methodological, and Ethical Discussion. *Design Issues*, *17*(1), 5–17. https://doi.org/10.1162/07479360152103796

Floridi, L., Cowls, J., Beltrametti, M., Chatila, R., Chazerand, P., Dignum, V., Luetge, C., Madelin, R., Pagallo, U., Rossi, F., Schafer, B., Valcke, P., & Vayena, E. (2018). AI4People—An Ethical Framework for a Good AI Society: Opportunities, Risks, Principles, and Recommendations. *Minds and Machines*, *28*(4), 689–707. https://doi.org/10.1007/s11023-018-9482-5

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, K. (2012). Models of Design: Envisioning a Future Design Education. *Visible Language*, *46*, 132-153.

Friedman, K. (2019). Design Education Today—Challenges, Opportunities, Failures. *Chatterjee Global Lecture Series*. https://www.academia.edu/40519668/Friedman._2019._Design_Education_Today_-_Challenges_Opportunities_Failures

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

Futurice. (2017). *The Intelligence Augmentation Design Toolkit*. https://futurice.com/ia-design-kit

Giaccardi, E., & Redström, J. (2020). Technology and More-Than-Human Design. *Design Issues*, *36*(4), 33–44. https://doi.org/10.1162/desi_a_00612

Glushko, R. J. (2020). *The Discipline of Organizing: 4th Professional Edition*.

Goodman, L. A. (1961). Snowball Sampling. *The Annals of Mathematical Statistics*, *32*(1), 148–170. https://doi.org/10.1214/aoms/1177705148

Google. (n.d.). *AI Experiments—Experiments with Google*. Retrieved 6 May 2022, from https://experiments.withgoogle.com/collection/ai

Google Developers. (n.d.). *Machine Learning Crash Vourse*. Retrieved 6 May 2022, from https://developers.google.com/machine-learning/crash-course

Gore, J. M., & Zeichner, K. M. (1991). Action research and reflective teaching in preservice teacher education: A case study from the United States. *Teaching and Teacher Education*, *7*(2), 119–136. https://doi.org/10.1016/0742-051X(91)90022-H

Green, W. S., & Jordan, P. W. (Eds.). (2002). *Pleasure With Products: Beyond Usability* (0 ed.). CRC Press. https://doi.org/10.1201/9780203302279

Hassenzahl, M., Burmester, M., & Koller, F. (2021). User Experience Is All There Is: Twenty Years of Designing Positive Experiences and Meaningful Technology. *I-Com*, *20*(3), 197–213. https://doi.org/10.1515/icom-2021-0034

Hassenzahl, M., Eckoldt, K., Diefenbach, S., Laschke, M., Lenz, E., & Kim, J. (2013). Designing Moments of Meaning and Pleasure. Experience Design and Happiness. *International Journal of Design*, *7*(3), 21–31.

Hebron, P. (2016). *Machine Learning for Designers*. O’Reilly Media, Inc.

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

Holbrook, J., & Lovejoy, J. (2017, August 31). *Design Is [Smart]*. Google Design. https://design.google/library/design-is-smart/

Holstein, K., Harpstead, E., Gulotta, R., & Forlizzi, J. (2020). Replay Enactments: Exploring Possible Futures through Historical Data. *Proceedings of the 2020 ACM Designing Interactive Systems Conference*, 1607–1618. https://doi.org/10.1145/3357236.3395427

Höök, K., & Löwgren, J. (2021). Characterizing Interaction Design by Its Ideals: A Discipline in Transition. *She Ji: The Journal of Design, Economics, and Innovation*, *7*(1), 24–40. https://doi.org/10.1016/j.sheji.2020.12.001

How AI can bring on a second Industrial Revolution. (2016, June). https://www.ted.com/talks/kevin_kelly_how_ai_can_bring_on_a_second_industrial_revolution

Huizinga, J. (1938). *Homo Ludens*. Routledge & Kegan Paul.

Huppatz, D. (2015). Revisiting Herbert Simon's "Science of Design". *Design Issues*, *31*, 29–40. https://doi.org/10.1162/DESI_a_00320

IBM. (2018, June 5). *Learn + Play with IBM Research AI*. IBM Research AI. https://research.ibm.com/interactive/learn-and-play/

IBM. (2019). *IBM Design for AI*. https://www.ibm.com/design/ai/

Ideo.org. (n.d.). *Design Kit*. Retrieved 12 March 2021, from https://www.designkit.org

Johnson, B. D. (2011). *Science Fiction Prototyping: Designing the Future with Science Fiction* (Vol. 3). Morgan & Claypool Publishers.

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

Kaggle. (n.d.). *Learn Intro to Machine Learning Tutorials*. Retrieved 6 May 2022, from https://www.kaggle.com/learn/intro-to-machine-learning

Keinonen, T. (2006). Fields and Acts of Art and Research. In M. Mäkelä & S. Routarinne (Eds.), *The Art of Research. Research Practices in Art and Design* (pp. 40–58). University of Art and Design Helsinki.

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

Kirschner, P. A., & Van Merriënboer, J. (2009). Ten Steps to Complex Learning A New Approach to Instruction and Instructional Design. In T. L. Good (Ed.), *21st century education: A reference handbook* (pp. 244–253). SAGE Publications. https://core.ac.uk/display/55535269?utm_source=pdf&utm_medium=banner&utm_campaign=pdf-decoration-v1

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

Krathwohl, D. R. (2002). A Revision of Bloom's Taxonomy: An Overview. *Theory Into Practice*, *41*(4), 212–218. https://doi.org/10.1207/s15430421tip4104_2

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

Kun, P. (2020). *Design Inquiry Through Data* [Delft University of Technology]. https://doi.org/10.4233/UUID:7E914DD9-2B53-4B2C-9061-86087DBB93B9

Kundisch, D., Muntermann, J., Oberländer, A. M., Rau, D., Röglinger, M., Schoormann, T., & Szopinski, D. (2022). An Update for Taxonomy Designers. *Business & Information Systems Engineering*, *64*(4), 421–439. https://doi.org/10.1007/s12599-021-00723-x

Kyshkan, V. (2020). *Machine Learning + Design*. https://www.machinelearning.design/

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

Latour, B. (2011). Love Your Monsters. *Breakthrough Journal*, *2*(11), 21–28.

Law, J. (2002). *Aircraft Stories: Decentering the Object in Technoscience*. Duke University Press. https://doi.org/10.1215/9780822383543

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

Levinson, P. (1977a). TOY, MIRROR, AND ART: The Metamorphosis of Technological Culture. *ETC: A Review of General Semantics*, *34*(2), 151–167. JSTOR.

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

Littman, M. L., Ajunwa, I., Berger, G., Boutilier, C., Currie, M., Doshi-Velez, F., Hadfield, G., Horowitz, M. C., Isbell, C., Kitano, H., Levy, K., Lyons, T., Mitchell, M., Shah, J., Sloman, S., Vallor, S., & Walsh, T. (2021). *Gathering Strength, Gathering Storms: The One Hundred Year Study on Artificial Intelligence (AI100) 2021 Study Panel Report*. Stanford University. http://ai100.stanford.edu/2021-report

Manzini, E. (2006). Il design in un mondo fluido. In H. Höger (Ed.), *Design education: Studiare design attraverso lo studio per progetti ; Designausbildung konzipiert als Projektstudium* (pp. 150–161). Ed. Abitare Segesta.

McCarthy, J., Minsky, M. L., Rochester, N., & Shannon, C. E. (1955). A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence. *AI Magazine 2006*, *27*(4), Article 4. https://doi.org/10.1609/aimag.v27i4.1904

Mekler, E. D., & Hornbæk, K. (2019). A Framework for the Experience of Meaning in Human-Computer Interaction. *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, 1–15. https://doi.org/10.1145/3290605.3300455

Merleau-Ponty, M. (1945). *Phenomenology of Perception*. Routledge.

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

Mitchell, T. M. (1997). *Machine Learning*. McGraw-Hill.

Molnar, C. (2019). *Interpretable Machine Learning*. https://christophm.github.io/interpretable-ml-book/

Muratovski, G. (2016). *Research for designers: A guide to methods and practice*. Sage Publications.

Nickerson, R. C., Varshney, U., & Muntermann, J. (2013). A method for taxonomy development and its application in information systems. *European Journal of Information Systems*, *22*(3), 336–359. https://doi.org/10.1057/ejis.2012.26

Norman, D. (n.d.). *What is Humanity-Centered Design?* The Interaction Design Foundation. Retrieved 13 December 2022, from https://www.interaction-design.org/literature/topics/humanity-centered-design

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/

Norman, D. A., & Stappers, P. J. (2015). DesignX: Complex Sociotechnical Systems. *She Ji: The Journal of Design, Economics, and Innovation*, *1*(2), 83–106. https://doi.org/10.1016/j.sheji.2016.01.002

Norman, D. A., & Verganti, R. (2014). Incremental and Radical Innovation: Design Research vs. Technology and Meaning Change. *Design Issues*, *30*(1), 78–96.

Norvig, P. (2018). *Introduction to Machine Learning | Google Developers* [ML Crash Course]. https://developers.google.com/machine-learning/crash-course/ml-intro

Norvig, P. (2022, last update). *1542 Schools Worldwide That Have Adopted AIMA*. http://aima.cs.berkeley.edu/adoptions.html

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., … Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ*, n71. https://doi.org/10.1136/bmj.n71

Perrault, R., Shoham, Y., Brynjolfsson, E., Clark, J., Etchemendy, J., Grosz, B., Lyons, T., Manyika, J., Mishra, S., & Niebles, J. C. (2019). *The AI Index 2019 Annual Report* (No. 2). AI Index Steering, Stanford Institute for Human-Centered AI. https://hai.stanford.edu/sites/default/files/ai_index_2019_report.pdf

Piet, N. (2019). *AI Meets Design Toolkit*. https://www.aixdesign.co/toolkit

Piet, N. (2020). *AIxDesign Community | aixdesign.co*. AIxDesign. https://www.aixdesign.co

R2D3. (n.d.). *A visual introduction to machine learning*. Retrieved 6 May 2022, from http://www.r2d3.us/visual-intro-to-machine-learning-part-1/

Rampino, L., & Colombo, S. (2012a). Method, Strategy or Tool? A Semantic Clarification. In L. Rampino (Ed.), *Design research: Between scientific method and project praxis notes on doctoral research in design 2012* (pp. 83–94). FrancoAngeli.

Rampino, L., & Colombo, S. (2012b). Toward a Taxonomy of Design-research Methods. In L. Rampino (Ed.), *Design research: Between scientific method and project praxis notes on doctoral research in design 2012* (pp. 109–125). FrancoAngeli.

Reaktor, & University of Helsinki. (2018). *Elements of AI*. Elements of AI. https://course.elementsofai.com/

Reaktor, & University of Helsinki. (2020). *Building AI. A free online introduction to artificial intelligence for non-experts*. Elements of AI. https://buildingai.elementsofai.com

Redström, J. (2020). Certain Uncertainties and the Design of Design Education. *She Ji: The Journal of Design, Economics, and Innovation*, *6*(1), 83–100. https://doi.org/10.1016/j.sheji.2020.02.001

Rittel, H. W. J., & Webber, M. M. (1973). Dilemmas in a general theory of planning. *Policy Sciences*, *4*(2), 155–169. https://doi.org/10.1007/BF01405730

Robson, C., & McCartan, K. (2015). *Real World Research. A Resource for Users of Social Research Methods in Applied Settings* (4th edition). 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.

Saldaña, J. (2009). *The coding manual for qualitative researchers*. Sage.

Samuel, A. L. (1959). Some Studies in Machine Learning Using the Game of Checkers. *IBM Journal of Research and Development*, *3*(3), 210–229. https://doi.org/10.1147/rd.33.0210

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., & Zavarrone, E. (2022). The qualities of AI-infused products. Reflections on emerging UX dimensions. In D. Spallazzo & M. Sciannamè (Eds.), *Embedding Intelligence. Designerly reflections on AI-infused products* (pp. 48–76). Franco Angeli Design International. https://series.francoangeli.it/index.php/oa/catalog/view/848/695/4914

Sciuto, A., Saini, A., Forlizzi, J., & Hong, J. I. (2018). "Hey Alexa, What's Up?": A Mixed-Methods Studies of In-Home Conversational Agent Usage. *Proceedings of the 2018 Designing Interactive Systems Conference*, 857–868. https://doi.org/10.1145/3196709.3196772

Shirky, C. (2008). *Here Comes Everybody: The Power Of Organizing Without Organizations*. Penguin Press.

Simon, H. A. (1996). *The sciences of the artificial* (third edition). MIT Press.

Spallazzo, D., Scianname', M., Ajovalasit, M., Ceconello, M., & Arquilla, V. (2022). Exploring AI-infused Products Qualities to Unleash AmI. *Ambient Intelligence – Software and Applications*. 12th International Symposium on Ambient Intelligence, Salamanca. https://doi.org/10.1007/978-3-031-06894-2

Stappers, P. J., & Giaccardi, E. (2017). Research through Design. In M. Soegaard & R. Friis-Dam (Eds.), *The Encyclopedia of Human-Computer Interaction, 2nd edition*. Interaction Design Foundation. https://www.interaction-design.org/literature/

book/the-encyclopedia-of-human-computer-interaction-2nd-ed/research-through-design

Stickdorn, M., Lawrence, A., Hormess, M., & Schneider, J. (Eds.). (2018). *This is Service Design Doing. Applying Service Design Thinking in the Real World*. O'Reilly.

Stoimenova, N., & Price, R. (2020). Exploring the Nuances of Designing (with/for) Artificial Intelligence. *Design Issues*, *36*(4), 45–55. https://doi.org/10.1162/desi_a_00613

Stolterman, E. (2008). *The Nature of Design Practice and Implications for Interaction Design Research*. International Journal of Dsign. http://www.ijdesign.org/index.php/IJDesign/article/view/240/148

Swann, C. (2002). Action Research and the Practice of Design. *Design Issues*, *18*(1), 49–61. https://doi.org/10.1162/07479360252756287

Thaler, R. H., & Sunstein, C. R. (2009). *Nudge: Improving decisions about health, wealth, and happiness* (Rev. and expanded ed). Penguin Books.

Tremosa, L. (2022). *The Future of Design—What's Next?* The Interaction Design Foundation. https://www.interaction-design.org/literature/article/the-future-of-design-what-s-next

Tuan, Y. (1977). *Space and Place. The Perspective of Experience.* University of Minnesota Press.

Turing, A. M. (1948). *Intelligent machinery*. National Physics Laboratory.

Umbrello, S. (2020). Imaginative Value Sensitive Design: Using Moral Imagination Theory to Inform Responsible Technology Design. *Science and Engineering Ethics*, *26*(2), 575–595. https://doi.org/10.1007/s11948-019-00104-4

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 Allen, P. (2018). Prototyping ways of prototyping AI. *Interactions*, *25*(6), 46–51. https://doi.org/10.1145/3274566

Van Allen, P. (2017, March 20). Reimagining the Goals and Methods of UX for ML/AI. *2017 AAAI Spring Symposium Series*. 2017 AAAI Spring Symposium Series. https://aaai.org/ocs/index.php/SSS/SSS17/paper/view/15338

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

Villari, B. (2012). Action Research and Design Research: Theory and Practice. In L. Rampino (Ed.), *Design research: Between scientific method and project praxis notes on doctoral research in design 2012* (pp. 127–139). FrancoAngeli.

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

Weil, D., & Mayfield, M. (2020). Tomorrow's Critical Design Competencies: Building a Course System for 21st Century Designers. *She Ji: The Journal of Design, Economics, and Innovation*, *6*(2), 157–169. https://doi.org/10.1016/j.sheji.2020.03.001

Weiser, M. (1993). Some Computer Science Issues in Ubiquitous Computing. *Communications of the ACM - Special Issue on Computer Augmented Environments: Back to the Real World*, *36*(7), 75–84.

Weiser, M. (1994). Creating the invisible interface. *ACM Symp. User Interface Software and Technology UIST'94*.

Wensveen, S. (Director). (2021, October 8). *Aesthetics of Intelligence*. https://www.youtube.com/watch?v=QyDnhibt2Bc

White, R. W. (2018). Skill Discovery in Virtual Assistants. *Communications of the ACM*, *61*(11), 106–113.

Winograd, T. (2006). Shifting viewpoints: Artificial intelligence and human–computer interaction. *Artificial Intelligence*, *170*(18), 1256–1258. https://doi.org/10.1016/j.artint.2006.10.011

Winter, M., & Jackson, P. (2020). Flatpack ML: How to Support Designers in Creating a New Generation of Customizable Machine Learning Applications. In A. Marcus & E. Rosenzweig (Eds.), *Design, User Experience, and Usability. Design for Contemporary Interactive Environments* (Vol. 12201, pp. 175–193). Springer International Publishing. https://doi.org/10.1007/978-3-030-49760-6_12

WIRED (Director). (2021, August 18). *Computer Scientist Explains Machine Learning in 5 Levels of Difficulty | WIRED*. https://www.youtube.com/watch?v=5q87K1WaoFI

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. (2018). *Machine Learning as a UX Design Material: How Can We Imagine Beyond Automation, Recommenders, and Reminders?* https://aaai.org/ocs/index.php/SSS/SSS18/paper/view/17471

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. (2017). The Role of Design in Creating Machine-Learning-Enhanced User Experience. *AAAI Spring Symposia*.

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., Sciuto, A., Zimmerman, J., & Steinfeld, A. (2018, June). Investigating How Experienced UX Designers Effectively Work with Machine Learning. *Proceedings of the 2018 Designing Interactive Systems Conference*.

Yang, Q., Steinfeld, A., Rosé, C., & Zimmerman, J. (2020). Re-examining Whether, Why, and How Human-AI Interaction Is Uniquely Difficult to Design. *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, 1–13. https://doi.org/10.1145/3313831.3376301

You, X., & Hands, D. (2019). A Reflection upon Herbert Simon's Vision of Design in The Sciences of the Artificial. *The Design Journal*, *22*(sup1), 1345–1356. https://doi.org/10.1080/14606925.2019.1594961

YouTube Originals (Director). (2019, December 18). *The Age of A.I.* https://www.youtube.com/watch?v=UwsrzCVZAb8

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

Zhang, D., Maslej, N., Brynjolfsson, E., Etchemendy, J., Lyons, T., Manyika, J., Ngo, H., Niebles, J. C., Sellitto, M., Sakhaee, E., Shoham, Y., Clark, J., & Perrault, R. (2022). *The AI Index 2022 Annual Report* (No. 5). AI Index Steering, Stanford Institute for Human-Centered AI. https://aiindex.stanford.edu/wp-content/uploads/2022/03/2022-AI-Index-Report_Master.pdf

Zwick, C. (2006). Intrecciare università e mondo del lavoro. In H. Höger (Ed.), *Design education: Studiare design attraverso lo studio per progetti ; Designausbildung konzipiert als Projektstudium* (pp. 98–106). Ed. Abitare Segesta.

INDEX OF FIGURES

Index of figures

INDEX OF TABLES

Index of tables