Youth Critical Engagements with Machine Learning: Insights from a Future Archetypes Workshop

Eva Durall Gazulla
eva.durallgazulla@oulu.fi
University of Oulu
Oulu, Finland

Najme Babai
najme.babai@oulu.fi
University of Oulu
Oulu, Finland

Netta Iivari
netta.iivari@oulu.fi
University of Oulu
Oulu, Finland

Kylie Peppler
kpeppler@uci.edu
University of California Irvine
Irvine, United States

Santiago Ojeda-Ramirez
sojedara@uci.edu
University of California Irvine
Irvine, United States

Marianne Kinnula
marianne.kinnula@oulu.fi
University of Oulu
Oulu, Finland

Abstract

This exploratory study examines how youth critically engage with artificial intelligence (AI)/machine learning (ML) and sustainability through an activity inspired by future scenario archetypes. We analyze a workshop in which 20 teens (aged 15–17) collaboratively designed speculative AI/ML-based services related to love and friendship. Our analysis focuses on interactional dynamics and the technology- and sustainability-related discourses that emerged during the design process. The findings show how speculative, playful, and cooperative engagement supported youth in situating AI/ML within broader sociotechnical worlds, questioning data dependencies, and developing critical perspectives on sustainability and social justice. Based on these insights, we propose recommendations for child-computer interaction (CCI) activities, including fostering awareness of data construction, encouraging reflection on the interdependence of AI/ML and data, situating designs in complex sociotechnical systems, and using metaphors to explore various futures.

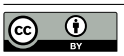
CCS Concepts

• **Human-centered computing**; • **Computing methodologies** → **Machine learning**; **Learning paradigms**; **Supervised learning**; **Unsupervised learning**; **Reinforcement learning**;

Keywords

Youth, Speculative design, Future archetypes, World-building, Sociotechnical futures, Critical AI literacy, Machine learning, Critical Data Literacy

ACM Reference Format:
Eva Durall Gazulla, Najme Babai, Netta Iivari, Kylie Peppler, Santiago Ojeda-Ramirez, and Marianne Kinnula. 2026. Youth Critical Engagements with Machine Learning: Insights from a Future Archetypes Workshop. In *Proceedings of the 14th Nordic Conference on Human-Computer Interaction (NordiCHI '26), October 03–07, 2026, Vaasa, Finland*. ACM, New York, NY, USA, 17 pages. https://doi.org/10.1145/3829807.3829902



NordiCHI '26, Vaasa, Finland

ACM ISBN 979-8-4007-2373-5/2026/10
https://doi.org/10.1145/3829807.3829902

1 Introduction

Artificial Intelligence (AI)/Machine learning (ML)–powered services are increasingly embedded in young people's everyday lives and are expected to become even more pervasive as datafication intensifies across social, educational, and economic domains [38, 48]. These systems shape what information young people encounter, how their actions and performance are assessed, and what futures appear possible or desirable, with implications for issues ranging from opportunity and participation to sustainability and social justice. As a result, researchers and educators have raised growing concerns about how to support young people's understanding of AI and ML [23, 59, 98], leading to a range of initiatives aimed at promoting critical AI/ML and data literacies among youth [3, 28, 45, 70, 98].

Within this body of work, many scholars argue that young people should have meaningful opportunities to participate in shaping the futures imagined and enacted around technology, rather than being positioned solely as end users of systems designed according to the interests and assumptions of large technology companies [39, 79]. This position aligns with the participatory design (PD) tradition, which emphasizes the right of those affected by a technology to have a genuine say in its design and in the societal arrangements it supports [15, 17, 85, 86]. PD with children and youth has extended these commitments to child–computer interaction (CCI), foregrounding children's and young people's meaningful involvement in the design of digital technologies and recognizing them as legitimate contributors to debates about technology and society [31, 41, 42].

While this work establishes a strong normative and methodological foundation, less is known about how specific design arrangements support young people's critical engagement with AI/ML, particularly when that engagement is oriented toward imagining alternative futures. If we assume that futures are not predetermined but actively constructed through present-day decisions, then critically reflecting on the kinds of futures being produced by AI/ML systems becomes an important design challenge [35, 53]. Futures research has long attended not only to the plurality of possible futures, but also to questions of who gets to imagine, articulate, and influence them [4, 30, 34, 44]. For CCI, this raises a practical question about how design activities might be structured so that young people can meaningfully explore, contest, and reimagine AI/ML-driven futures through interaction.

We take this as the high-level aim of this study: to understand what kinds of design arrangements and practices support youth critical engagement with AI/ML through interaction, particularly in enabling critical reflection on sociotechnical implications and imagined alternative futures. This aim brings together three mutually reinforcing orientations: sustainability as a concern for intergenerational futures, participatory design as a commitment to giving voice to those who will inhabit them, and speculative design as a practice for making those futures tangible and open to critique. Youth are uniquely positioned at the intersection of all three.

Recent CCI research has begun to explore speculative and futures-oriented approaches with children and youth, positioning imaginaries as a site of empowerment and critique [39, 55, 83]. At the same time, a growing body of research in CCI and the learning sciences has begun to reposition youth not only as learners of AI/ML, but as designers, evaluators, and critical interlocutors in its development. For example, [48]'s discussion on K–12 students creating AI/ML applications frames design as a pedagogical space where youth grapple with data practices, model behavior, bias, and accountability. [67] similarly show how adolescents, acting as peer auditors of AI/ML applications, learned to identify algorithmic bias and connect it to dataset and model design choices. However, much of this work remains oriented toward present-day systems and system-level critique (see e.g. [13, 74, 84]. Less attention has been paid to the design arrangements that might support young people in engaging with AI/ML as a future-oriented sociotechnical phenomenon—one entangled with broader questions of social relations, sustainability, and justice (e.g. [5, 58]). Emerging work on speculative design with youth suggests that practices such as worldbuilding and futures imagination can help learners interrogate dominant assumptions about AI and ML and articulate competing visions of what could be otherwise ([52, 53, 91, 92]. Yet, we still lack detailed accounts of the interactional dynamics through which such critical futures-oriented engagement unfolds in practice.

The present study contributes to this line of work by exploring young people's critical and situated understandings of AI/ML through interaction in a speculative design activity. We examined a design activity in which 15–17-year-olds were invited to envision sociotechnical futures of love and friendship using futures archetypes, with AI/ML algorithms as a central mediating technology. Our analytic focus is on the interactional dynamics that emerged during the activity—that is, the social and material processes through which participants collaboratively made sense of AI/ML futures, negotiated values, and developed critiques in interaction with peers and workshop materials—and on how these dynamics supported critical engagement with AI/ML in relation to social and environmental futures. We adopt a situated perspective, viewing the learning activities and materials as introduced into, and intermingled with, a complex fabric of practices, relations, discourses, and histories that shape interaction moment by moment (e.g., [80]).We see this exploratory study as foundational for examining youths' critical engagements with AI/ML, with implications for the design of activities that support informed and justice-oriented AI/ML literacy.

As our research question we ask: *What kinds of interactions emerge when youth engage in futures archetype–based AI/ML design activities, and how do these interactions support their critical thinking about sociotechnical futures, including sustainability and social justice?* Based on our analysis, we articulate recommendations to guide CCI research and practice aimed at fostering children's and young people's empowerment with AI/ML through futures-oriented design.

2 Related Background

2.1 Youth as Designers of Machine Learning Systems

In response to growing calls to support youth AI/ML literacies, curricula and pedagogical designs aimed at fostering understanding of AI concepts, such as ML, have proliferated [12, 78, 95]. Across this work, a recurring challenge is that young people often struggle to grasp abstract AI and ML concepts despite initial curiosity [36]. To address this, researchers have increasingly engaged youth as active designers of AI/ML systems, using hands-on design experiences to scaffold conceptual understanding. Most of these efforts, however, focus on how AI/ML systems work in the present—how they classify, predict, and potentially discriminate—rather than on how they may shape future social and material worlds.

Research positioning youth as designers of AI/ML systems has explored diverse design activities through which young people engage with data-driven technologies [14, 60, 69, 76, 82]. These studies situate learning within hands-on design practices, allowing youth to engage with data, models, and system outputs. They consistently highlight the close relationship between data literacy and algorithmic literacy [47, 57], showing that understanding AI/ML systems requires attention not only to algorithms but also to the origins, structure, and representativeness of datasets [29, 72]. Yet a systematic review of 34 AI/ML education tools found that most simplify this complexity, showcasing AI/ML capabilities while obscuring key design decisions such as data curation and trade-offs between performance and fairness [12]. Similarly, a review of 43 K-12 AI/ML education studies found that while pedagogical development dominates the field, critical reflection on the societal and ethical implications of AI/ML remains limited [78].

To encourage ethical reflection, some initiatives have engaged youth in design, evaluation, and reflection activities [9, 13, 77, 96]. Others have involved youth in evaluating and auditing AI/ML systems [67, 71], demonstrating how evaluation can function as a design practice that helps participants connect data, models, and social consequences. Such capacities have been framed as part of AI/ML agency, emphasizing young people's ability to critically use, examine, and produce ML applications [2].

Youth have also been positioned as co-designers of AI/ML education programs, helping shape curricula through cooperative inquiry [9, 95]. Across K-12 contexts, studies show students collecting and labeling data, training models, and iteratively refining performance [48]. These activities often generated unexpected outcomes, such as biased or inaccurate predictions, prompting participants to revisit design decisions and consider issues of fairness and accountability [48]. However, such engagements remain largely grounded in existing technologies and contemporary problems. AI/ML is typically treated as a present-day artifact to be understood or critiqued, leaving underexplored how future trajectories might be imagined,

redirected, or evaluated in terms of the values and social conditions they may reproduce.

A smaller body of work has combined AI/ML education with speculative approaches. Speculative design grounded in young people's lived experiences has positioned youth as co-designers and generated social justice-oriented design principles [90]. Likewise, [5] used design fiction to introduce teenagers to ML, showing how speculative methods can support both learning and critical reflection. Yet even these approaches remain largely focused on existing systems and their current shortcomings.

Despite growing interest in critical and speculative AI/ML design [53], explicitly exploring AI/ML as sociotechnical systems that could unfold differently across possible futures remains underdeveloped [10]. This points to a need for youth-as-designers approaches that support speculative and future-oriented engagement with AI/ML, attending to how materials, prompts, and social configurations can scaffold critical reflection and collaborative sense-making.

2.2 Speculation and Worldbuilding for Youth Engagement with Sociotechnical Futures

Over the past two decades, HCI research has increasingly engaged with speculative design [33]. As a futures-oriented practice, speculation complements participatory design with youth: if participatory design establishes the right of those affected by technology to shape it, speculative design provides means to exercise that right in relation to the futures technology produces. Approaches such as design fiction, critical design, and design futuring seek to question dominant assumptions, imagine alternatives, and provoke reflection on sociotechnical trajectories [11, 16, 32, 33, 56, 75, 87]. Within HCI, speculative design has been used to probe and envision digital futures through artifacts, prototypes, and scenarios that make possible worlds tangible and discussable.

Research characterizes effective speculative design as both speculative—fictional, critical, and sociopolitical—and discursive, inviting reflection and dialogue [75]. Although speculative design may incorporate critical perspectives, it does not necessarily address injustice, power, or oppression directly [99]. Recent work also emphasizes process qualities such as grounding, participation, and playfulness [75]. While speculative design was not historically participatory [99], contemporary practice increasingly involves stakeholders as active contributors [75], aligning well with long-standing CCI concerns.

CCI research has only recently begun to draw explicitly on speculative design to support young people's engagement with digital futures [43]. Existing studies show its value for enabling youth to probe, debate, and imagine sociotechnical futures [40, 81, 99]. This includes work addressing AI/ML systems as sociotechnical phenomena rather than purely technical artifacts [53, 62, 89]. For example, [53] found that combining speculative design with AI/ML learning enabled adolescents to imagine AI/ML in alternative sociocultural contexts and examine its political, ecological, and cultural implications. Such activities encouraged participants to consider diverse stakeholder perspectives and challenge dominant AI/ML narratives [53, 89].

Overall, speculative approaches show strong potential for supporting critical engagement with power relations and the societal implications of intelligent technologies [40, 53, 62]. They may also encourage youth to move beyond critique toward imagining interventions and alternatives [40, 53, 62]. However, questions remain about how speculative activities can be designed to support engagement with complex sociotechnical entanglements. Closest to our study, [53] used materials such as cards and dioramas to support worldbuilding but left open how youth can be systematically scaffolded to examine AI/ML interrelations with social, political, and ecological dimensions. Our study addresses this gap by combining non-digital AI/ML materials with worldbuilding techniques, specifically futures archetypes, to support explorations of AI/ML futures and their societal implications.

To conceptualize this approach, we draw on worlding, or worldbuilding, a concept that has received limited explicit attention in CCI despite its promise. Worlding refers to an active, productive practice of co-creating worlds (e.g., futures, imaginary realms, lived realities, and ways of being), acknowledging that no single, given world exists [24, 25, 88, 100]. In this framing, young people are speculative designers who create imaginative projections of alternative presents and futures [100]. Unlike linear narratives, worldbuilding creates coherent worlds that surface possibilities, tensions, desires, and uncertainties [25, 100]. It can also support critical agendas by foregrounding the power relations, assumptions, and ideologies embedded in imagined worlds, thereby opening space for reflection and challenge [24]. While speculative design and worldbuilding have been explored with youth, little research has examined how worlding practices shape engagement with complex sociotechnical relations, particularly in the context of AI/ML systems.

2.3 Sustainability and Social Justice in the Design of Digital Technologies

Sustainability has become an important concern in CCI, understood not only as an environmental goal but as a sociotechnical design concern. It provides the normative horizon connecting the two preceding threads: if youth are legitimate designers of AI/ML systems and speculative design enables alternative futures, sustainability concerns whether those futures will be liveable and just for future generations.

Although the Brundtland Report defines sustainability as meeting present needs without compromising those of future generations [51], later work has expanded it to include environmental, social, economic, political, cultural, and technological dimensions [65]. In CCI, sustainability has thus shifted from a policy objective to a design-oriented concern that shapes how futures are imagined and enacted through sociotechnical practices [18, 27]. It focuses not only on technological performance or ecological outcomes, but also on how design practices structure relations among humans, non-humans, data, and infrastructures, and how benefits, risks, and responsibilities are distributed [93].

A growing body of CCI research has explored sustainability through participatory design with young people. Vasalou and Gauthier's review shows that environmental sustainability research, while still limited, increasingly seeks to make hidden environmental

processes visible, support understanding of cause–effect relationships, and foster affective connections with ecological systems [97]. Similarly, [26] highlight the need to strengthen research on how digital technologies support youth learning and participation in sustainability-oriented action.

Beyond environmental sustainability, CCI research has also addressed social and economic dimensions, emphasizing healthy habits, community cohesion, ethical reflection, and equitable access to resources and opportunities for long-term well-being [1, 8, 63, 79, 94]. More recently, researchers have called for systemic approaches that recognize the interdependence of ecological, social, and economic concerns [54]. Despite these advances, sustainability in CCI remains predominantly framed through designed outputs rather than as a quality of the design and participation processes themselves. In this regard, Kinnula et al. argue that sustainability with AI requires systems thinking, whereby young people develop holistic understandings when interdependencies, power relations, and long-term consequences are foregrounded [54]. Similarly, Om et al. emphasize the need for methods that support young people in reasoning about environmental issues as systemic and relational [73].

Another limitation is the limited attention to justice and power. Although sustainability is often associated with ethics and responsibility, few studies explicitly address social inequities, structural power, or whose futures are prioritized. Albar et al. show that many sustainability-oriented games overlook connections between environmental, social, and economic dimensions, marginalizing questions of equity [6]. Vasalou and Gauthier similarly argue that sustainability is often treated as an application domain rather than a critical lens for examining broader sociopolitical relations [97]. Consequently, sustainability risks becoming technocentric, emphasizing individual behavior over questions of justice and the distribution of risks and benefits.

This study extends this literature by conceptualizing sustainability as a processual, justice-oriented sociotechnical practice enacted through young people's participation in designing with and about AI/ML. It examines how sustainability emerges through speculative, collaborative, and critical design processes in which youth negotiate environmental concerns alongside questions of power, responsibility, and future-making. As the primary inhabitants of the futures being designed, young people are central participants in this conversation.

3 Methodology

3.1 Research Setting and Participants

This exploratory study focuses on a two-hour workshop entitled *Questioning the future: Futures archetypes* conducted as part of the AI Ambassador program (see Table 1 for the full program overview), a two-week paid summer traineeship organized by university researchers in collaboration with city-level partners at a Nordic university in 2025. The purpose of the study was to explore salient learning processes and forms of critical engagement that emerged during the activity and to generate qualitative insights to inform the design of future CCI activities.

The call was advertised through a range of channels, including the city website, university channels, local education networks, and teachers connected to the researchers, who distributed the information to their students. Following this campaign, 20 teenagers aged 15–17 who applied voluntarily were selected to take part in the program. The researchers facilitating the activities selected participants to ensure a diverse group in terms of gender, cultural background, and socioeconomic circumstances. The participant group included nine boys and eleven girls, representing African, Asian, American, and European backgrounds. All resided in the city where the program took place; several had migration experiences. Positioned as a form of non-formal education, the program framed youth as active contributors and "ambassadors" rather than passive learners. The program was delivered in English, and all participants demonstrated sufficient proficiency to participate fully in discussions and collaborative activities.

Before data collection, participants completed an informed consent process in which they learned about the research purpose, the personal data collected, and their rights as research participants. They were clearly informed that they could withdraw from the study at any time without consequences. Information on data privacy and consent forms were provided in both English and one of the official national languages. The research adhered to ethical guidelines for research with children formulated by the National Board for Research Integrity, and the informed consent process was overseen by an external ethics advisory board composed of internationally recognized experts in research ethics involving youth.

Prior to the workshop, participants filled a background questionnaire probing their motivations, expectations, and prior experiences with AI and ML, and digital technologies. Responses indicated a broad range of interests: some expressed interest in technical aspects such as model training, testing, and real-world applications, while others voiced concern about AI/ML's societal impacts, including environmental sustainability and everyday life implications. Most participants reported spending several hours per day on digital media platforms, particularly TikTok, YouTube, Instagram, and online games. While many recognized common AI and ML-driven systems (e.g., chatbots, recommendation algorithms, and ChatGPT) this awareness often stemmed from peer discussion or media exposure rather than direct hands-on use. Besides ChatGPT, no other large language models were explicitly mentioned in their open-ended responses. Participants' definitions of AI and ML ranged from technical ("a program that learns from data and algorithms") to anthropomorphic ("a smart computer that thinks for you") and critical ("the most dangerous technology because it mimics humans and can steal identity"), reflecting varied and sometimes conflicting mental models.

3.2 Workshop Outline and Supporting Materials

At the workshop, participants were invited to collaboratively envision the futures of love and friendship through AI/ML speculative designs in five groups of four, pre-formed with them in prior sessions to balance self-identified skills and participation styles considering a range of competencies, including research skills, teamwork, creativity, and technical skills. Each group designed a ML-driven service set in the year 2065, situating their designs within a possible future and considering social, economic, and environmental implications alongside technical choices. The focus on love and

Table 1: Overview of the AI Youth Ambassador programme activities

Day	Focus	Activities
D1	Grounding	Programme welcome and orientation (2 h); team formation and community building (1 h); introduction to ML concepts and algorithms (1 h 15 min); AI in media and society (1 h).
D2	Exploring and questioning	Energizer (15 min); imagining ML futures with futures archetypes (2 h); digital storytelling and video production workshop (2 h 30 min); reflection and social media engagement (30 min).
D3	Questioning	Energizer (15 min); hands-on ML with Teachable Machine (2 h); AI, democracy, and decision-making (2 h 30 min); reflection and social media engagement (30 min).
D4	Analysis	Energizer (15 min); understanding ML through decision trees I (2 h 15 min); exploring societal consequences of AI through the futures wheel (2 h 15 min); reflection and social media engagement (30 min).
D5	Analysis and envisioning	Energizer (15 min); understanding ML through decision trees II (2 h 15 min); exploring plausible AI futures with the futures triangle (2 h 15 min); reflection and social media engagement (30 min).
D6	Envisioning	Energizer (15 min); video production: Voices of Youth on AI (2 h 15 min); design fiction ideation and brainstorming (2 h 15 min); reflection and social media engagement (30 min).
D7	Envisioning	Design fiction storytelling and world-building (2 h 15 min); creating design fiction artefacts (2 h 30 min); reflection and social media engagement (30 min).
D8	Envisioning	Producing design fiction artefacts (2 h 15 min); creating short films based on future AI scenarios (2 h 30 min); reflection and social media engagement (30 min).
D9	Sharing and advocacy	Editing and refining design fiction films (2 h 15 min); developing social media stories and final presentations (2 h 30 min); reflection and social media engagement (30 min).
D10	Sharing and advocacy	Finalising social media stories and presentations (2 h 15 min); showcasing outcomes and exchanging insights with fellow AI Ambassadors and broader audience (3 h).

friendship was selected by the researchers facilitating the program based on their knowledge of topics relevant to this age group and insights from previous workshops. The theme was also timely, given recent claims by Meta's CEO that AI companions could increasingly play a role in people's social relationships, making it a potentially provocative topic for youth.

To scaffold this process, participants were provided with three sets of cards. The first set visualized four futures archetypes—growth, collapse, constraint, and transformation—drawing on Dator's typology [30] and serving as metaphoric lenses for imagining diverse socioeconomic and environmental conditions. Growth refers to futures driven by continued economic and technological expansion; collapse depicts breakdowns of systems due to crises or unsustainable trajectories; constraint envisions futures shaped by limits, where resources and behaviors are regulated to maintain stability; and transformation represents radical systemic change, often involving shifts in values, technologies, or social organization (Figure 1).

The second set introduced three broad AI/ML paradigms: supervised (learning from labeled data), unsupervised (discovering patterns in unlabeled data), and reinforcement learning (learning through trial and error based on rewards and penalties), supporting participants' discussions about technical functionality (Figure 2). The rationale for using current AI/ML paradigms in the design of future services stemmed from the view that design is a future-making activity, through which possible futures are imagined, explored, and materialized [64].

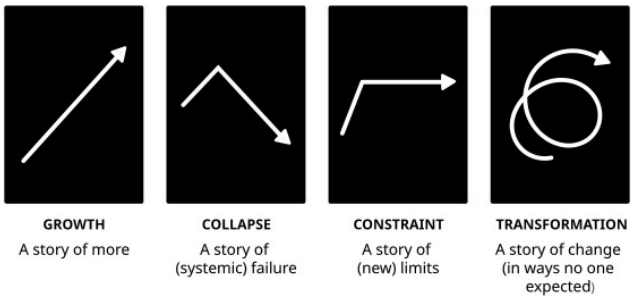


Figure 1: Cards visualizing the four future archetypes consisting in growth, collapse, constraint and transformation.

The third set consisted of dataset cards describing potential training data, including scope, biases, and relevant metadata such as geographical origin, language, and user engagement metrics (see Figure 3). These materials were inspired by prior kits such as "I Love Algorithms" [22], designed to demystify AI/ML and to support hands-on ideation without requiring computational infrastructure. At the beginning of the workshop, participants were introduced to different types of futures and AI/ML paradigms. As participants reported, this information was new to them. Following this introduction, they engaged in a one-minute-per-idea brainstorming activity using the Crazy 8's method, generating sketches of multiple potential futures for each archetype. Participants then shared their sketches within their groups and selected ideas to develop a future

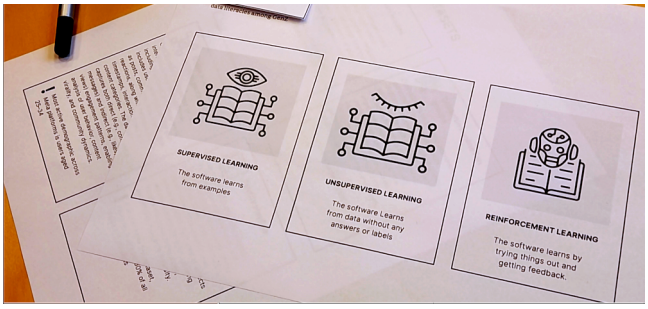


Figure 2: Cards representing visually the AI/ML paradigms the participants could select: supervised learning, unsupervised learning and reinforcement learning

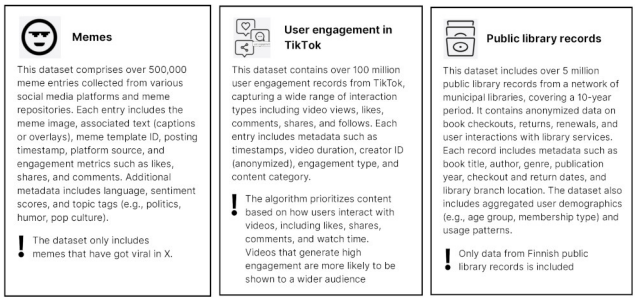


Figure 3: Dataset cards containing information about the metadata included and one embedded bias.

scenario aligned with their assigned archetype. During the group design phase, a future archetype was randomly assigned to each group to ensure diversity in the scenarios developed and encourage participants to imagine futures beyond their initial assumptions and preferences. With this basis, participants chose one AI/ML paradigm and two dataset cards to inform their speculative design. The workshop concluded with short group presentations, followed by peer questions and a whole-group discussion (Figure 4).

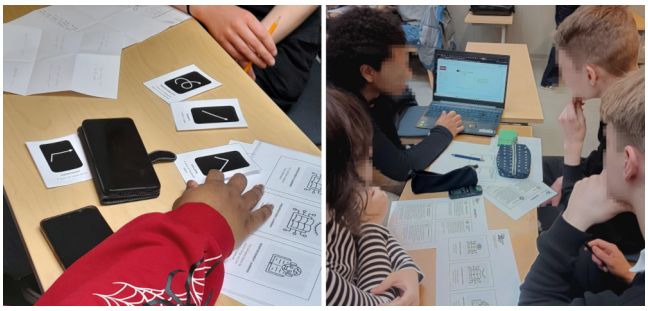


Figure 4: Participants working on their future AI/ML services for supporting love and friendship in a particular future

3.3 Data Collection and Analysis

Data for the current study included background questionnaires, audio and video recordings from the workshop, three facilitators' observation notes, and the materials and artifacts produced during the activities. Participants' feedback on the impact of the activity was collected through video-recorded focus groups conducted one week after the workshop and reflective essays submitted five weeks after completion of the AI Ambassador program. The focus groups explored participants' experiences of the workshop process, their expectations, and the relevance of learning about AI algorithms and decision-making. Using a visual timeline activity, participants reflected on their engagement across sessions, identified the most and least meaningful activities, and discussed the development of key competencies, changes in their understanding of AI, and the impact of the workshops on their attitudes toward technology and future interactions with AI. Finally, participants shared key takeaways and suggestions for improving future iterations of the programme. This approach aimed to capture participants' insights by allowing time for reflection and supporting diverse forms of expression, producing a dataset that captured both immediate and retrospective perspectives (see Table 2).

The analysis of the data followed Braun and Clarke's six-phase framework for reflexive thematic analysis (RTA) [19, 20] to qualitatively identify salient learning processes and discursive patterns in youths' engagement with AI/ML and sustainability (see Table 3). Through four cycles of collaborative discussion with co-authors, the first two authors compared, refined, and expanded interpretations, aiming to deepen insights via reflexive dialogue [20] (see Table 3).

The two first-authors began by familiarizing themselves with the dataset and inductive, semantic coding of the full data. This initial stage generated preliminary codes such as "future archetypes", "dataset choice", "bias and limitations", "ethical concerns", "sustainability", and "emotional or relational dimensions" which were discussed with all co-authors. During the second and third round of analysis the initial codes were iteratively clustered and arranged around the categories of "futures", "social", "technical" and "sociotechnical" aspects that captured recurring and conceptually significant patterns. During the fourth iteration, the collaborative analysis examined patterns in interaction and discourses about sustainability and justice embedded in participants' work, leading to the generation of themes such as "ethical tensions with AI/ML" and "AI/ML threats to sustainability" (see Table 4 for examples). The process also incorporated interpretive triangulation by connecting insights across the different data sources. Subsequent phases involved developing, reviewing, and defining themes to construct a coherent narrative, which concluded in writing up the analysis in relation to relevant literature. The insights gained in each cycle of analysis were discussed with all co-authors which informed the next round of analysis.

4 Findings

To address our research question, we report findings on how youth engaged in futures archetype-based AI/ML design activities and how these interactions supported critical thinking about sociotechnical futures, sustainability, and social justice. Our analysis distinguishes between critiques prompted by workshop scaffolds (e.g.,

Table 2: Overview of data collection activities in chronological order

Activity	Description	Research data
Background questionnaire	Distributed at the beginning of the program to gather participant motivations, prior knowledge of AI and ML, and digital habits	Questionnaire responses (n=20)
Group design activity	Five groups of four spent 60 min developing AI/ML service concepts using future archetypes, algorithms, and datasets	Verbatim transcription of the group discussions' audio recordings (5 h), Crazy 8s scenarios (n=14), observation notes from three facilitators
Final presentations	Each group presented and discussed their futures scenarios and speculative designs of AI/ML services with the whole group	Verbatim transcription of the presentations video recordings (30 min), presentation slides (five sets)
Focus groups	Two groups (one of 10 and another of nine) reflected on their experience across the program	Audio recordings of focus group discussions (four hours)
PechaKucha presentation	Each group presented their takeaways (20 slides, 20 s each) at the end of the program	Video recordings of presentations (35 min), presentation slides (five sets)
Reflection essays	Individual reflective essays about their experience during the AI Ambassador Program (500-700 words)	Written essays (n=19)

Table 3: Summary of the data analysis process based on Braun and Clarke's six-phase framework for reflexive thematic analysis

Phase N.	Phase name	Description
1	Familiarising yourself with the dataset	The two first authors reviewed the data.
2	Coding	The two first authors coded a portion of the data.
3	Generating initial themes	Tentative codes were discussed with all co-authors, after which the two first authors coded the remaining data.
4	Developing and reviewing themes	Discussion of themes with co-authors. Two first authors conducted further analysis rounds.
5	Refining, defining and naming themes	Collaborative refinement, definition, and naming of fourth-cycle themes with all co-authors.
6	Writing up	The final themes were presented in findings section, and all co-authors reviewed them.

Table 4: Examples of the subcodes and quotes identified during the fourth cycle of analysis

Code	Identified theme (example)	Quote (example)
Interaction pattern	Ethical tensions with AI/ML	*"Isn't having such a device in your home kind of inhumane?"*
Interaction pattern	Critical engagement with algorithms and data	*"I think also the problem with algorithm is like they always like try to be like nice so that means they might give you the information they like over exaggerated"*
Discourse	AI/ML threats to sustainability	*"And it [AI/ML] uses or it takes really much resources from our nature, because like, like for example, chatGPT, it uses really much water, every time you use it."*
Discourse	Value-driven AI/ML design for sustainability	*"Hence, these [human] needs can be used as axioms for building a framework of niceness."*

future archetypes, dataset cards, and AI/ML model cards) and those that emerged as participants expanded or challenged these prompts through discussion. We begin with an overview of the future scenarios created by participants, before examining how critical speculation emerged and how sustainability, social justice, and AI/ML became intertwined in the futures they imagined.

4.1 Overview of the Futures Scenarios Created by Youth

When ideating their services for supporting love and friendship, participants were randomly assigned one of the four future archetypes. This led to two groups focusing on collapse, two on constraint, and one on growth future scenarios. Each group was tasked with not

only imagining a future world but also designing a digital service in that context, powered by AI/ML to recommend social connections.

Within one hour, all five groups successfully produced both a scenario aligned with their assigned archetype and a conceptual AI/ML-powered service. While all services shared the broad aim of fostering love and friendship, they varied significantly in terms of context, target audiences and geographic scope. Three were designed for a global reach, while two focused on national or local levels. The imagined social contexts reflected major challenges such as societal polarization, misinformation and propaganda, climate change, as well as hate speech and violence (see table 5 for a summary).

A notable feature across groups was the way youth envisioned AI/ML models. The algorithms they described integrated diverse metadata from multiple datasets—ranging from memes, library records, and hobby club lists to social media activity and multilingual interactions—to train the models powering their digital services. Participants were free to choose the AI/ML paradigm they considered most appropriate for their service. Most groups opted for supervised learning, while one group selected reinforcement learning.

In the following section, we present our findings on the interaction dynamics that emerged during the activity, as well as the key discourses on social justice and sustainability reflected in participants' conversations. To illustrate these themes, we include de-identified excerpts using a group–participant notation (e.g., G1P4), where *G* indicates the group number and *P* the individual participant within that group.

4.2 Future Archetypes as Catalysts for Critical Speculations

4.2.1 Ethical Engagements with Sociotechnical Systems through Collaborative Negotiation. Across groups, interactional dynamics were shaped by ethical tensions that surfaced as participants engaged with the future archetypes and imagined sociotechnical futures. While the archetypes provided broad conditions for speculation, many of the ethical concerns emerged through participants' discussions of how AI/ML systems might shape social relations and everyday experiences in those futures.

Instead of avoiding or quickly resolving disagreement, moments of hesitation, discomfort, and critique became productive sites for examining how data and AI/ML might reshape social relations. Participants used these exchanges to question assumptions, weigh trade-offs, and recalibrate proposed designs. Ethical reflection thus unfolded through collaborative negotiation grounded in concrete decisions about data collection, representation, and use.

A common pattern in three groups involved negotiating ethical disagreements, especially when AI/ML features clashed with values such as privacy or social justice. In Group 4, for instance, a proposal for continuous behavioral monitoring immediately raised discomfort: *"It just monitors like everything… it's kind of dystopian"* G4P4. Rather than discarding the idea, participants reframed surveillance as purposeful and socially beneficial: *"It can connect to their page and see what they're interested in"* G4P3; *"It… monitors your online footprint… and then it can match people with similar interests"* G4P3. Here, ethical tension prompted justification, with participants acknowledging potential harm while articulating why such practices might be acceptable within their imagined future.

Similar negotiations emerged when participants discussed systems that could reveal personal traits. When one idea involved an AI/ML companion disclosing characteristics during an interview, the suggestion prompted immediate resistance. As one participant put it, *"Imagine… your robot just tells the person how you are"* G4P3. Another followed by questioning the loss of control this would entail, noting, *"That's a bit weird. Cause what if you wanna hide something?"* G4P4. Rather than dismissing the feature outright, the group worked through contextual boundaries for disclosure, showing how critique surfaced through friction and the need to define conditions and limits. Participants also explicitly raised concerns about AI/ML fallibility when imagining high-stakes scenarios such as police use.

In other groups, ethical tensions focused on visibility and exclusion in data-driven systems. Group 3 initially proposed filtering out *"Swifties"*, but peers challenged the implications of excluding an entire fan community, prompting reconsideration of how algorithmic moderation could undermine the goal of fostering friendship. Group 2 encountered similar tensions around inclusivity. Using hobby-related data seemed promising until participants recognized it would exclude older adults or people who do not engage in sports. Exploring alternatives, such as library datasets, revealed further limitations, as participants noted that relying on public library records would exclude those who do not read books or engage with such services. Ethical critique unfolded through iterative refinement as participants debated what inclusive data practices might look like.

4.2.2 Critical Explorations of Data and AI/ML. Across groups, participants relied on the material and conceptual scaffolds of the activity to design their ML service ideas, especially the dataset and AI/ML model cards. These materials frequently prompted discussion about data sources, model types, and their implications. At the same time, participants often extended these prompts through their own reasoning, introducing concerns about representation, consent, legitimacy, and exclusion that were not explicitly specified in the materials. Together, the workshop scaffolds and participants' interpretations shaped their envisionings of future sociotechnical systems.

During the group work, participants used the dataset cards as a starting point for evaluating which datasets would be meaningful for training their AI/ML models. While the cards drew attention to different forms of data, participants themselves expanded the discussion to questions of bias, representation, and exclusion. Rather than treating datasets as neutral technical inputs, youth questioned which practices and populations they made visible or invisible. In Group 2, a discussion about using a hobby-based platform (*"My Club"*) raised concerns about inclusivity: *"old people don't really go to clubs, right?"* G2P1. This quickly shifted once participants imagined themselves in 2065: *"We are the old ones, then?"* G2P4. Repositioning themselves within the future helped them see that present-day data infrastructures may not endure, prompting attempts to broaden representation. As one participant noted, limiting data to sports or reading was *"not inclusive enough"* G2P3, while another (G2P4)

Table 5: Summary of participants' envisioned future scenarios and the services shaped by specific future archetypes.

Group and Archetype	Future scenario description	Service ideated
G1: Collapse	A future United States where democratic institutions have collapsed, giving rise to a fascist regime marked by authoritarian control, mass surveillance, and extreme ideological polarization. A global economic crisis fuels a new Cold War, with society pushed toward civil war at home.	**MAFA:** A dual-use app functioning as a conventional platform for romantic connections, and to facilitate discreet networking for organizing a revolutionary movement. Its matching AI/ML algorithm integrates metadata from memes shared across social media and online celebrity fan groups to connect individuals discretely. By using people's digital traces, the system allows individuals to hide their involvement while helping resistance networks coordinate in a divided society.
G2: Constraint	Climate change and finite resources have forced society to embrace limits. Long-distance travel is minimized, local consumption is prioritized, and technology is harnessed for ecological resilience while human agency over AI/ML systems remains central.	**LokaLove:** a sustainability-driven mobile app that helps local communities connect while promoting sustainable living. It employs a reinforcement learning model trained on metadata from public library borrowing records and online leisure group activities to recommend nearby relationships and shared initiatives. By encouraging nearby relationships and reducing long-distance travel, the app helps lowering carbon emissions. The technology behind the app runs on clean energy and uses efficient systems to minimize environmental impact.
G3: Collapse	Society is characterized by the erosion of trust and the spread of discourses based on fear. Uncontrolled misinformation, amplified by platforms like TikTok have severely damaged societal fabric. The prevalence of confusion and manipulation makes it extremely challenging to build authentic human connections.	**Propaganda-free platform:** an AI/ML-powered friendship and dating app that aims to restore trust through supervised learning trained on TikTok activity metadata and fan community interactions. The model actively detects and filters propaganda signals, enabling authentic and safe human connections to flourish in a landscape otherwise saturated by disinformation.
G4: Growth	Society evolves under pervasive AI/ML surveillance, where facial recognition and emotional tracking are an intrinsic part of daily life. AI/ML systems monitor interactions and guide decisions, even within institutions like law enforcement. As AI/ML becomes embedded in personal relationships, trust and connection are increasingly mediated—and in some cases, replaced—by intelligent systems that learn and adapt over time.	**FACE (Facial Recognition and Cue Evaluation):** A dual-purpose companion robot and institutional tool that uses facial recognition and social media data to analyze behavior and provide real-time feedback on social cues. Its supervised learning model integrates multimodal metadata—facial recognition inputs, TikTok and Netflix behavioral logs, and police interrogation recordings—To deliver real-time social feedback.
G5: Constraint	After a period characterized by social division, society embraces empathy and mutual understanding. Guided by kindness and reflection on past mistakes, communities strive to cultivate relations based on respect and cooperation. Though inequalities persist, authentic human connection and collective care are the center of this future society.	**NICELINGO:** an AI/ML-powered app designed to help people connect through compassion. Using supervised learning on multilingual datasets containing social interactions metadata from Meta platforms such as Instagram, Nicelingo filters and teaches kind, culturally respectful phrases across languages—empowering users to be nice to more people, wherever they go.

suggested that library data might capture a wider range of activities, such as borrowing movies.

As participants looked for alternatives to the existing datasets, their engagement with data shifted from selecting among the provided options to critically examining the broader conditions under which data could be collected and used. Notably, concerns about consent did not stem directly from the dataset cards themselves but emerged through participants' own attempts to imagine how new datasets might be created. For instance, when Group 2 proposed creating their own dataset, concerns about legitimacy were immediately raised: *"How can you… collect data without like [consent]?"* G2P3. Drawing on the dataset cards, participants discussed the institutional arrangements and responsibilities involved in making such systems possible. These discussions became sites of speculative design in which participants articulated and negotiated ethical considerations about the acceptability of data-driven futures.

Group 4 similarly used the dataset cards as a stepping stone for debating the appropriateness of different data sources. A dismissal of Spotify—*"Definitely not Spotify"* G4P3—elicited a challenge that reframed music as socially expressive: *"Music actually makes a pretty big part of your personality"* G4P4. While the cards foregrounded potential datasets, participants extended the discussion by reasoning about how metadata such as location histories, fitness tracking, and bodily indicators, might be used by AI/ML systems to infer social traits through aggregated traces of everyday activity. Such reasoning reflects a form of critical engagement with data practices, where participants recognized how intimate aspects of life could become legible, categorized, and acted upon by future sociotechnical systems.

As with the dataset cards, participants used the AI/ML model cards to think through possible service designs. When participants struggled to align imaginative ideas with technical models, interaction slowed and became more reflective. For instance, in Group 1, participants showed uncertainty as they revisited algorithm definitions: *"Guys, we might have this all wrong. Wait, what does it say? Unsupervised learning, supervised learning, reinforcement learning"* G1P4. Such pauses marked moments where speculative storytelling gave way to re-engagement with algorithmic logic.

Through engagement with AI/ML materials participants also reflected on the limits of automated classification. Group 1 initially proposed training a model on memes to detect extremist content—*"If you like these fascist memes you can't enroll" G1P4—but quickly recognized problems of interpretation: "They use very complex dog whistle[s]"* G1P1. Concluding *"We don't use the memes"* G1P4, the group abandoned the dataset entirely. This moment shows how participants questioned the interpretability of data and deliberated about what AI/ML can reliably parse—and where those capacities break down.

Taken together, these examples illustrate how participants engaged critically with data and AI/ML algorithms in their discussions. Through their interactions, they raised questions about how data-driven sociotechnical systems operate and considered the ways such systems could contribute to social exclusion and injustice.

4.2.3 Playful Interactions Supporting Critical Speculations. Humor and exaggeration emerged as central interactional resources, enabling participants to explore sociotechnical future worlds in ways that were playful, critical, and grounded in familiar digital experiences. Rather than undermining the task, these modes of engagement helped sustain participation, manage uncertainty, and open space for imaginative speculation. Humor thus functioned as a form of youth agency: participants appropriated the design brief and reshaped it into futures that were meaningful to them, while still engaging with contemporary sociotechnical concerns.

In Group 3, exaggerated scenarios about platforms like TikTok became tools for reasoning about systemic risk and information manipulation. While discussing a future dominated by algorithmic control, participants joked about moderator collapse and platform takeover. One participant suggested: *"So, systemic failure… the TikTok mods just… go to sleep"* G3P4, prompting another to playfully escalate, *"They take over. Oh my god"* G3P3. This humorous framing continued as they imagined misinformation sparking global conflict: *"All the, like, big world powers… create a really convincing narrative that they're all fighting each other"* G3P4. Although delivered jokingly, these exchanges revealed awareness of propaganda, platform governance, and algorithmic amplification, with humor serving as a low-stakes way to surface shared concerns.

Humor also enabled reflection on participants' own digital dependencies. While imagining a scenario involving North Korean propaganda and restricted access to platforms, one participant of Group 3 joked, *"Imagine not to be addicted to TikTok!"* G3P4. The comment, though lighthearted, surfaced tensions between state-controlled information environments and everyday media habits.

In Group 4, exaggeration became a mechanism for probing surveillance and control. Ideas escalated quickly toward total monitoring, including an app that *"monitors just like everything you do"* G4P4 and *"takes over the whole world"* G4P4. Other members of the group expanded on this with visions of invisible scanning and lie detection through facial recognition. These extreme scenarios allowed participants to test ethical limits and consider trade-offs between safety, privacy, and autonomy by pushing concepts far beyond realistic boundaries. Group 1 relied heavily on humor to navigate the social implications of algorithmic matching. Jokes about a dating app based on shared interests included imagining users meeting only to realize, *"Damn, you're ugly!"* G1P4, followed by the quip, *"Love is blind!"* G1P4. Humor later took a political turn as participants riffed on potential branding: *"I wouldn't be surprised if some MAGA had developed (...) a dating app called Making Dating Great Again"* G1P4.

Across groups, humor and exaggeration formed a distinctive interactional mode marked by imaginative freedom and critical reflection. By amplifying trends such as surveillance, misinformation, and algorithmic governance, participants used playful exaggeration to engage seriously with sociotechnical futures while maintaining ownership of the process.

4.3 Interweaving Sustainability, Social Justice, and AI/ML in Future Worlds

4.3.1 AI/ML Challenges to Social Sustainability. Participants' discussions about the negative impacts of AI and ML focused on how these systems can undermine democratic governance and erode individuals' fundamental rights. In particular, challenges related to privacy, the right to information, and freedom of expression were actively discussed when participants reflected on aspects connected to social sustainability.

Two groups envisioned future societies in which surveillance enabled by data-centric technologies had become a normalized social practice (see G1 and G4 in Table 5). In both cases, participants highlighted continuous monitoring and data extraction as defining features of technologically advanced societies where individuals' right to privacy had been sacrificed. While in Group 1 participants described a totalitarian regime using mass-surveillance technologies to suppress dissent, in Group 4 surveillance performed by data-driven systems was framed as serving the common good - *"It doesn't necessarily (...) spy on people. It's just trying to (...) help you maintain friendships"* G4P4.

Narratives about misinformation and societal polarization also appeared in future scenarios of the participants (see G1, G3, and G5 in Table 5), where these dynamics were framed as threats to social

sustainability because they eroded trust and social relations. In Group 3 for example, societal polarization is fueled by AI/ML algorithms supporting the spread of propaganda and misinformation in social media platforms such as TikTok. Yet the discourse underlying these scenarios also revealed a complex relationship to data itself, as participants recognized that it is not straightforward to determine which data should be labeled as misinformation or propaganda. As one participant noted, *"propaganda is in our everyday life, in the news and everything"* G3P2, followed by agreement from peers. Although the group chose not to follow this thread further to avoid getting *"stuck in debates about what is propaganda and what is not"* G3P4, the moment nonetheless reflected a nuanced understanding: the criteria for deciding what data is valid, trustworthy, or harmful are not fixed, and may be contested.

AI/ML ability to identify and filter data was linked to discourses about technology-enabled societal control (see G1, G3, and G5 in Table 5). These future scenarios explored different facets of technology-enhanced moderation and its impacts on freedom of expression and social sustainability. For instance, Group 1 members envisioned a unique encoding system designed to bypass conventional moderation mechanisms, allowing individuals critical of the autocratic regime to express themselves and connect with like-minded peers. In contrast, Groups 3 and 5 described AI/ML-assisted moderation as a valuable strategy for managing negative phenomena such as misinformation and hate speech spreading in social media: *"We're focused on banning like all forms of propaganda, so we'd like train the AI to recognize propaganda and recognize the characteristics of propaganda, so it doesn't allow it to say"* G3P4 (during their presentation to the whole group). These speculative designs sparked debates about freedom of speech, the challenges of misinformation, and the extent to which technologies could contribute to more sustainable societies.

4.3.2 Integrating Sustainability into AI/ML Processes. Participants' reflections on sustainability extended beyond the end-impacts of AI/ML and data-driven technologies on societal development. As shown in their group discussions and future scenarios, they also considered qualities related to the processes through which these technological outcomes were produced. For instance, groups focusing on social sustainability (see G4, G5 in Table 5), i.e., technology supporting an inclusive and equal society where wellbeing of citizens is taken care of, envisioned designs that foster hybrid intelligence. In these cases, the AI/ML algorithms powering the services generated models designed to guide human interactions while still preserving human agency. In both scenarios, the emphasis was on providing real-time feedback to support conflict-free and respectful communication between humans, with agency distributed between the AI/ML algorithms and the humans involved.

While participants acknowledged the role of AI/ML algorithms in mediating and shaping social relations, they emphasized the importance of preserving human agency. According to participants, human agency could be maintained by ensuring that humans had the final say in decisions - see for instance G4 member clarification: *"So, the robot is (...) not giving absolute solution, it's giving just suggestions"* G4P3 - and by involving them in the process of creating the models on which the AI/ML systems based their decisions - see G1 member explanation: *"So we could keep a close eye on it because it would be, it's very important that (...) that the AI doesn't (...) match people incorrectly"* G1P4. Four of the groups (G1, G3, G4, G5) pointed at human supervision in the data labelling process. They justified their choice through rationales such as *"The model type, we decided to use supervised learning, since in supervised learning, we can tell it what kind of language to filter out, and thus it can teach the nice language to the users"* G5P3, *"We were gonna supervise it... We were gonna give it (...) different human emotions. So we're gonna say, (...) this picture means happy, this picture means sad"* G4P4. In the single instance where participants used reinforcement learning to train their service's AI/ML algorithm, they also underscored how human input shaped the process: *"And we're using the reinforcement learning method because (...) then people can give (...) feedback like if they liked the person, the app matches you with. So it learns by that"* G2P4. In all these cases, participants pointed at human oversight and the possibility to intervene in the training of the AI/ML algorithm as crucial for ensuring the social sustainability of their service design. They associated social sustainability with designing services that support fairness, inclusion, accountability, and the well-being of the people interacting with them.

Participants' understanding of sustainability as a concern with underlying processes also surfaced in discussions about the environmental cost of training the AI/ML models that supported their services. Environmental sustainability was a central theme in Group 2 envisioning, where participants highlighted the challenges of accessing reliable information about the carbon footprint of processing large datasets and emphasized the need to prioritize smaller but more relevant datasets for model training: *"It could be, like, a sustainable nature-friendly app that collects only (...) local information"* G2P3.

Taken together, these reflections show that participants didn't view sustainability as a fixed end goal. Instead, they saw it as continually shaped through the everyday practices, choices, and interactions involved in developing AI/ML-driven services. They emphasized the importance of human judgment, transparent and inclusive ways of building models, and an awareness of the environmental costs tied to data and computation. In doing so, they treated sustainability not as a box to tick at the end of a project, but as an ongoing responsibility woven into every stage of creating technology.

4.3.3 Future Worlds Guided by Values of Social Justice, Care, and Sustainability. In their future scenarios, youth recognized the tensions and frictions between economic expansion, environmental degradation, and social well-being. The relationalities they envisioned - tighter friendships, increased reliance on community, or new patterns of interaction - were shaped directly by these intersecting pressures, which they perceived as challenges to sustainability (see Table 5 for a summary of participants' future scenarios and service designs).

Participants' sensitivity to how different pressures shape everyday life also informed how they approached sustainability more broadly. A central feature of participants' future scenarios was the way they brought together social, economic, and environmental concerns. Such multi-dimensional understanding of sustainability was characterized by the intertwinement of ecological, social, and

economic conditions shaping everyday life and interpersonal relations. A member of G4, working within a growth-driven scenario, explained this in detail: *"We got growth, and you can see how growth is going today, it's not going very well. The more we try and grow our economy in society (…) the more it destroys, it continues global warming and stuff like that. So, the society we're thinking of (…) it's become pretty toxic outside. It's dangerous to stay outside for too long […]. So (…) friendships are just more important"* G4P4. This systemic way of thinking was evident in worlds where social justice required economic reorganization, or where environmental limits reshaped political and cultural practices.

Many groups challenged mainstream assumptions about economic growth, resource use, or technological progress, experimenting with futures that not only felt possible and plausible but also pushed against dominant narratives. As Group 5 members argued, imagining alternative futures required questioning long-standing global relationships and considering more just, if imperfect, ways of relating: *"So perhaps there might not be equality in all of the world, but as for the colonial relationships, [...] in the interest of being nice, there might be some other way"* G5P4. In this case, the rejection of worldviews reproducing exploitative relations was expressed through a conscious choice of alternative sustainability-oriented values such as fairness, care, and mutual respect even when discussing large-scale political or historical dynamics. Another example of critiquing the continuous-growth paradigm—and highlighting the need to embrace limits—can be seen in Group 2 proposal for a digital service that supports local love and friendship by using smaller but highly relevant datasets to reduce the environmental footprint of large-scale data processing.

Across all scenario archetypes—whether framed around collapse, constraint, or growth—the workshop materials prompted participants to consider broad sustainability challenges. Participants then elaborated these conditions in different ways, connecting them to issues such as colonial relationships, fairness, democratic participation, and care, themes that were not explicitly prescribed by the materials. Even when futures appeared difficult or undesirable, they did not treat these conditions as fixed. Instead, each archetype became an opening for rethinking and transforming unsustainable dynamics. Their narratives frequently positioned technology as a means to foster more democratic practices, protect fundamental rights, and support more caring relationships among individuals, communities, and the environment.

Through collectively envisioning future worlds, participants reflected on how social, environmental, and economic forces could reshape everyday life while exploring how these trajectories might be redirected. Even when their visions highlighted difficult circumstances, they approached them as contexts in which change remained possible. Rather than accepting unsustainable dynamics as inevitable, they considered alternative arrangements of social relations and technology aimed at fostering more equitable and caring futures. Their discussions therefore emphasized not only potential risks but also the roles of collective action, creativity, and everyday care in shaping more sustainable ways of living.

5 Discussion

In this study we inquired what kinds of interactions emerge when youth engage in futures archetype–based AI/ML design activities, and how these interactions support their critical thinking about sociotechnical futures, including sustainability and social justice. Through this exploratory study, we show that as youth imagined future scenarios from the perspective of a particular archetype, they collaboratively engaged in rich, critical, and ethically attuned conversations about the futures they envisioned. These interactions were characterized by collaborative negotiation, as participants articulated, compared, and revised ideas in response to one another. Through these exchanges, youth considered differing viewpoints while examining how sociotechnical systems shape—and are shaped by—human values and social dynamics. They probed the roles of data and AI/ML in their imagined worlds, questioning how data are produced, how models function, and how algorithmic choices influence social outcomes. Playful interactions, including joking, exaggeration, and teasing, supported fluid shifts between speculative imagination and critical reflection. Our analysis indicates that the futures archetype method fostered interactions centered on critical questioning, speculative exploration, and reflections on sustainability within sociotechnical systems. We argue that these interactional dynamics emerged both from the method and activity design and from the distinctive perspectives and contributions that youth brought to the work. Next, we discuss these findings alongside recommendations for practice in relation to recent literature in CCI research.

Our exploratory research findings align with recent CCI studies advocating critical AI and ML literacy [50, 53, 98] and the computational empowerment of young people in relation to AI/ML systems [13, 49, 53, 66, 68]. Within this literature, critical and ethical aspects of AI/ML systems have been explored with children and youth, including agency, responsibility, bias, transparency, and the social implications of algorithmic decision-making. We extend this work by attending to interaction dynamics that reveal youths' adoption of a protagonist role as they explore the complex relations between AI/ML, society, and possible futures. In line with prior research on youth's protagonist role in design activities [46, 61, 101], our analysis showed participants taking an active role in the design work: they took ownership of the AI/ML design brief, exercised agency through collaborative negotiation, and engaged in critical exploration by questioning assumptions and critiquing the futures they envisioned. In these accounts, the protagonist role is understood not as something youth simply possess, but as something that emerges through interaction, collaboration, and recognition. Our study contributes to this body of research by providing a case focused on AI/ML design through speculative activities.

Our findings show that a central aspect in participants' critical engagement with AI/ML was their attention to data and questions of representation, inclusion, and fairness. Their discussions frequently moved beyond AI/ML system outputs to consider the datasets that underpin such systems and the assumptions embedded within them. This finding points to the importance of supporting youth in critically examining data as part of AI/ML literacy activities. Based on this, we offer the following design recommendations:

Recommendation 1: Trigger awareness of the situated and constructed character of datasets in AI/ML-powered services. Speculative design activities with youth can be strengthened by making explicit that datasets used in AI/ML systems—and the metadata that describe, organize, and delimit them—are not neutral representations of the world but are produced within specific social, historical, and political contexts. Facilitators should emphasize that simply collecting more data does not resolve issues such as bias, exclusion, or discrimination, as these problems are often embedded in how datasets are constructed, documented, and maintained, whose experiences are represented, and whose are left out. Encouraging youth to reflect on the limitations, gaps, and assumptions encoded in datasets and their metadata can surface questions of power, representation, and fairness early in the design process, shifting attention from technical performance toward the broader social consequences of data-driven systems.

In our study, this reflection was supported through dataset cards that described not only the intended use and content of each dataset, but also its limitations, omissions, and constraints. Although social justice was not explicitly named in the activity brief, participants' discussions around dataset selection frequently revolved around who might be excluded, whose data were missing or misrepresented, and how such absences could shape system behavior and reinforce existing inequities. This approach is novel in the context of speculative AI/ML design with youth, as it moves beyond treating datasets as abstract or given resources and instead invites youth to engage critically with datasets and metadata as socially situated artifacts. By foregrounding dataset limitations and documentation as sites of inquiry, the activity supported early forms of critical data literacy, enabling young people to reason about AI/ML systems in relation to fairness, accountability, and social justice rather than optimization alone.

Recommendation 2: Encourage thinking about the interdependence of AI/ML and data. Much CCI research on AI/ML literacy has focused on helping youth understand the impacts, ethics, or outcomes of AI/ML systems, while the relationships between data, learning paradigms, and model behavior often remain abstract or implicit. As a result, core questions about how AI systems come to know, classify, or predict—and how these processes are shaped by data choices—can remain invisible. Our findings suggest that making these relationships explicit is essential for supporting deeper critical understanding. Using concrete materials, such as dataset cards and AI/ML-paradigm cards, helped surface how data actively shapes model construction, interpretation, and deployment, shifting attention from AI/ML as a "black box" toward the underlying logics of learning from data.

These tangible, non-digital materials enabled youth to reason about AI/ML paradigms (e.g., supervised, unsupervised, or reinforcement learning) in relation to the available data and the assumptions those paradigms encode. By requiring participants to justify their design choices—such as why a particular AI/ML paradigm made sense given the dataset they selected—the activity encouraged reflection on representation, labeling, and agency. Together, the dataset and paradigm cards supported participants in reasoning about how data and models shape one another. This approach goes beyond many existing AI/ML literacy activities by foregrounding the interdependence of data and learning paradigms rather than treating models as neutral or universally applicable. Integrating such materials into collaborative and creative design work can therefore foster more nuanced forms of critical data literacy, helping young people articulate how AI/ML systems are shaped by data practices and modeling choices, rather than viewing outputs as purely technical or inevitable consequences of "intelligent" systems.

In addition to contributing to work supporting youth's critical reflection on AI/ML and its social impacts, our study illuminates how young people engage with speculation as a space for inquiry rather than solution-finding. While prior studies have demonstrated the promise of speculative approaches and worldbuilding for inviting children and youth to critique and envision AI- and ML-based systems as complex sociotechnical phenomena with entangled political, ecological, and cultural dimensions [53, 62, 89], our findings show how worldbuilding sessions became sites where young people grappled with sociotechnical complexity in practice—articulating contested futures, negotiating values, and making sense of uncertainty through collaborative negotiation, critical exploration, and playfulness. These findings point to the importance of supporting youth in situating technologies within broader social, political, and environmental contexts, which we express through the following recommendation.

Recommendation 3: Invite youth to situate their designs within complex sociotechnical worlds. Speculative design activities on AI/ML can be strengthened by explicitly encouraging youth to imagine not only individual technologies, but the broader worlds in which those technologies exist. Worldbuilding offers a powerful way to scaffold this thinking by prompting participants to consider how imagined AI/ML systems relate to social relations, economic arrangements, environmental constraints, cultural norms, and geopolitical conditions. By constructing future worlds rather than isolated artifacts, young people are supported in envisioning AI/ML as part of interconnected sociotechnical ecosystems—where technologies both depend on and reshape societal structures. This approach helps surface questions about what conditions make certain innovations possible, whose interests they serve, and how technological possibilities are shaped by broader material and social realities.

Within these imagined worlds, facilitators should encourage reflection on governance arrangements, power relations, and rights, asking how different futures enable or constrain individual and collective agency. For example, youth might explore how AI/ML systems operate under different types of government, or how rights related to privacy, freedom of speech, and information are upheld or eroded. While prior speculative design work has often emphasized ethical concerns or individual impacts, it has less frequently engaged youth in considering how AI/ML technologies intersect with governance arrangements and structural inequalities. Attending these dimensions through worldbuilding can support deeper critical reflection, enabling young people to think about AI/ML not only as a technical or ethical artifact, but as a force entangled with political authority, social justice, and collective futures.

In our study, the value of worldbuilding was closely tied to the use of futures archetypes as scaffolds for imagining alternative worlds. Through this study, we illustrate how participants

created rich, controversial, and debatable possible futures that escaped flat binary (utopian/dystopian) descriptions. The futures archetypes—growth, collapse, constraint, and transformation [30]—served as broad framing devices for imagining alternative socioeconomic, environmental, and cultural conditions. As youth developed speculative AI/ML designs responding to a particular futures archetype, they engaged in critical reflection on how ideas might unfold differently across societal contexts, considering how social, political, and ecological assumptions shape technological possibilities. This resonates with prior CCI scholarship emphasizing critical imagination as a collaborative practice through which young people create cohesive, plausible, and socially embedded future worlds that account for ecological, political, cultural, and technological dynamics [24, 25, 68, 88, 100]. Consistent with speculative design traditions, situated and imagined scenarios helped render abstract AI/ML concepts tangible, fostering dialogue, critique, and reflection grounded in collective sense-making [75]. The role of the archetypes in shaping participants' thinking further suggests their value as practical scaffolds for futures-oriented AI/ML design activities. We capture this insight through the following recommendation:

Recommendation 4: Use metaphors to guide consideration of various futures. When running speculative design activities on AI/ML with youth, facilitators should intentionally introduce metaphors that frame futures as plural and contingent rather than singular or inevitable. Even when youth focus on developing one future scenario, metaphors of varying futures (e.g., growth, collapse, constraint, transformation) can help open up their thinking about how social, political, ecological, and technological conditions might differently shape AI/ML systems, including questions of environmental limits, resource use, and long-term sustainability. These metaphors provide an accessible shared language that supports "what if?" reasoning, enabling youth to reflect on values, trade-offs, unintended consequences, and the sustainability implications of technological choices without requiring advanced technical knowledge.

In practice, metaphors can be introduced through short narratives, visuals, or prompts and revisited throughout the activity to encourage comparison and reflection. Facilitators might invite youth to briefly consider how their designs would change under a different future metaphor. Such reflections support critical dialogue and help prevent overly deterministic or techno-optimistic visions of AI/ML detached from sustainability concerns. While such metaphors are well established in futures research, their use as scaffolds in speculative design with youth remains underexplored. Our findings suggest that working with futures metaphors can help young people situate AI/ML within broader sociotechnical and ecological contexts, fostering critical imagination and openness to multiple possible—and more or less sustainable—futures.

The importance of these scaffolds became particularly visible in participants' discussions of sustainability. Our study highlights how youths' engagements with various futures archetypes brought sustainability-related questions to the fore in concrete and situated ways. As participants worked across futures characterized by differing ecological conditions, resource constraints, governance arrangements, and societal priorities, sustainability emerged as something they actively wrestled with rather than a predefined theme. For instance, when imagining futures based on the growth or collapse archetype, youth raised concerns about environmental limits, extraction, and long-term viability, while discussions situated in the constraint archetype leaned towards more regenerative or social justice-oriented alternatives. Across these engagements, participants came to articulate sustainability as an entanglement of environmental, social, and economic considerations, rather than as isolated dimensions. These discussions also foregrounded how AI/ML technologies—depending on the values they embody and the contexts in which they are deployed—might intensify existing inequities or contribute to more equitable and sustainable sociotechnical trajectories. Our findings resonate with [7] argument that imagining futures is a values-laden social practice that benefits from layered scaffolding to guide collective envisioning. In this vein, the futures archetypes supported participants not only in generating diverse future scenarios—much like futures-envisioning games such as *The Thing from the Future* [21]—but also in recognizing sustainability as an ongoing negotiation between technological possibilities, ecological conditions, and social justice [37].

6 Conclusion

This study contributes to CCI research by introducing a worldbuilding activity inspired by the futures archetypes, in which 15–17 years-old participants envisioned sociotechnical futures of love and friendship by designing speculative AI/ML services. Our analysis examines the interactional dynamics that emerged during the activity and how these dynamics supported youth's critical engagement with AI/ML in relation to social and environmental futures. We situate our findings within scholarly work on critical AI/ML and data literacies, speculative futures, and sustainability. Drawing on these insights, we propose several recommendations to guide future CCI design work with youth: (i) trigger awareness of the situated and constructed character of data in ML-powered services, (ii) encourage thinking about the interdependence of AI/ML and data, (iii) invite youth to situate their designs within complex sociotechnical worlds, and (iv) use metaphors to guide consideration of various futures. Future work could build on this approach by exploring its long-term impact, adapting it to other age groups and cultural contexts, and examining how similar worldbuilding activities might support sustained critical reflection and civic engagement around emerging AI/ML technologies.

Acknowledgments

We are grateful to the participants of the AI Youth Ambassador Program 2025, as well as the Samueli Academy's students and their teacher Patricia Vining, who provided feedback on an earlier version of the activity. This work has been supported by the Research Council of Finland through the projects Critical Datalit 354445, TakeOver (363594) and FutuProta (351584). The work was also supported by the European Union's Horizon Europe research and innovation programme, through the Critical Changelab project under grant agreement 101094217, the Spencer Foundation and the Erasmus+ Programme of the European Union through the Alliance for Innovation project ComeThinkAgain under grant agreement No 101139845. Views and opinions expressed are however those of the authors only and do not necessarily reflect those of the European

Union or the European Education and Culture Executive Agency (EACEA). Neither the European Union nor the granting authority can be held responsible for them.

References

[1] Tessa Aarts, Panos Markopoulos, Lars Giling, Tudor Vacaretu, and Sigrid Pillen. 2022. Snoozy: a chatbot-based sleep diary for children aged eight to twelve. In *Proceedings of the 21st Annual ACM Interaction Design and Children Conference*. ACM, 297–307. doi:10.1145/3501712.3529718

[2] Ibrahim O. Adisa. 2023. Developing Machine Learning Agency Among Youth: Investigating Youth Critical Use, Examination, and Production of Machine Learning Applications. In *Proceedings of the 22nd Annual ACM Interaction Design and Children Conference (IDC '23)*. ACM. doi:10.1145/3585088.3593929

[3] Ibrahim Oluwajoba Adisa, Ian Thompson, Tolulope Famaye, Deepika Sistla, Cinamon Sunrise Bailey, Katherine Mulholland, Alison Fecher, Caitlin Marie Lancaster, and Golnaz Arastoopour Irgens. 2023. SPOT: A Game-Based Application for Fostering Critical Machine Learning Literacy Among Children. In *Proceedings of the 22nd Annual ACM Interaction Design and Children Conference*. ACM, 507–511. doi:10.1145/3585088.3593884

[4] Toni Ahlqvist and Martin Rhisiart. 2015. Emerging pathways for critical futures research: Changing contexts and impacts of social theory. *Futures* 71 (2015), 91–104. doi:10.1016/j.futures.2015.07.012

[5] Mariana Aki Tamashiro, Maarten Van Mechelen, Marie-Monique Schaper, and Ole Sejer Iversen. 2021. Introducing teenagers to machine learning through design fiction: An exploratory case study. In *Proceedings of the 20th annual ACM interaction design and children conference*. ACM, 471–475. doi:10.1145/3459990.3465193

[6] Raghad Albar, Andrea Gauthier, and Asimina Vasalou. 2024. A Playful Path to Sustainability: Synthesizing design strategies for children's environmental sustainability learning through gameful interventions. In *Proceedings of the 23rd Annual ACM Interaction Design and Children Conference*. ACM, 201–217. doi:10.1145/3628516.3655797

[7] Corina Angheloiu, Leila Sheldrick, Mike Tennant, and Goldie Chaudhuri. 2020. Future tense: Harnessing design futures methods to facilitate young people's exploration of transformative change for sustainability. *World Futures Review* 12, 1 (2020), 104–122. doi:10.1177/1946756719844050

[8] Alissa N Antle, Yumiko Murai, Alexandra Kitson, Yves Candau, Zoe Minh-Tam Dao-Kroeker, and Azadeh Adibi. 2022. "There are a LOT of moral issues with biowearables"... Teaching Design Ethics through a Critical Making Biowearable Workshop. In *Proceedings of the 21st Annual ACM Interaction Design and Children Conference*. 327–340.

[9] Golnaz Arastoopour Irgens, Ibrahim Adisa, Cinamon Bailey, and Hazel Vega Quesada. 2022. Designing with and for Youth: A Participatory Design Research Approach for Critical Machine Learning Education. *Educational Technology & Society* 25, 4 (2022), 126–141.

[10] Osnat Atias and Areej Mawasi. 2025. Conceptualizing AI literacies for children and youth: A systematic review on the design of AI literacy educational programs. *Computers and Education: Artificial Intelligence* (2025), 100491. doi:10.1016/j.caeai.2025.100491

[11] Jeffrey Bardzell and Shaowen Bardzell. 2013. What is" critical" about critical design?. In *Proceedings of the SIGCHI conference on human factors in computing systems*. 3297–3306.

[12] Karl-Emil Kjær Bilstrup, Magnus Høholt Kaspersen, Simon Enni, Ira Assent, and Marianne Graves Petersen. 2022. From Demo to Design in Teaching Machine Learning. In *Proceedings of the 2022 ACM Conference on Fairness, Accountability, and Transparency (FAccT '22)*. ACM, 2168–2178. doi:10.1145/3531146.3534634

[13] Karl-Emil Kjær Bilstrup, Magnus H Kaspersen, and Marianne Graves Petersen. 2020. Staging reflections on ethical dilemmas in machine learning: A card-based design workshop for high school students. In *Proceedings of the 2020 ACM designing interactive systems conference*. ACM, 1211–1222. doi:10.1145/3357236.3395558

[14] Karl-Emil Kjær Bilstrup, Magnus Høholt Kaspersen, Marie-Louise S. Kjerstein Sørensen, and Marianne Graves Petersen. 2022. Opportunities and Challenges of Teaching Machine Learning as a Design Material with the micro:bit. In *Adjunct Proceedings of the 2022 Nordic Human-Computer Interaction Conference (NordiCHI Adjunct '22)*. ACM. doi:10.1145/3547522.3547689

[15] Gro Bjerknes, Pelle Ehn, and Morten Kyng. 1987. *Computers and democracy-a Scandinavian challenge*. Gower Publishing.

[16] Mark Blythe. 2014. Research through design fiction: narrative in real and imaginary abstracts. In *Proceedings of the SIGCHI conference on human factors in computing systems*. ACM, 703–712. doi:10.1145/2556288.2557098

[17] Susanne Bødker, Christian Dindler, Ole S Iversen, and Rachel C Smith. 2022. What is participatory design? In *Participatory design*. Springer, 5–13. doi:10.1007/978-3-031-02235-7_2

[18] Sophia Brady, Eunice Kang, Emanuel Louime, Samantha Naples, Andrew Katz, and Avneet Hira. 2021. Emerging learning technologies for education on sustainability topics. In *REES AAEE 2021 conference: Engineering Education Research Capability Development: Engineering Education Research Capability Development*. Engineers Australia Perth, WA, 844–856. doi:10.52202/066488-0023

[19] Virginia Braun and Victoria Clarke. 2019. Reflecting on reflexive thematic analysis. *Qualitative research in sport, exercise and health* 11, 4 (2019), 589–597. doi:10.1080/2159676X.2019.1628806

[20] Virginia Braun and Victoria Clarke. 2021. One size fits all? What counts as quality practice in (reflexive) thematic analysis? *Qualitative research in psychology* 18, 3 (2021), 328–352. doi:10.1080/14780887.2020.1769238

[21] Stuart Candy. 2018. Gaming futures literacy: The thing from the future. In *Transforming the future*. Routledge, 233–246. doi:10.4324/9781351048002

[22] Carissa Carter, Megan Stariha, Mark Grundberg, Ariam Mogos, Kelly Schmutte, Seamus Harte, and Asha Lamanque. [n. d.]. *I Love Algorithms: A Machineless Machine Learning Creation Kit*. Retrieved September 11, 2025 from https://dschool.stanford.edu/tools/i-love-algorithms-machineless-machine-learning

[23] Lorena Casal-Otero, Alejandro Catala, Carmen Fernández-Morante, Maria Taboada, Beatriz Cebreiro, and Senén Barro. 2023. AI literacy in K-12: a systematic literature review. *International Journal of STEM Education* 10, 1 (2023), 29. doi:10.3102/ip.25.2192179

[24] Michael Alan Chang and Thomas M Philip. 2025. Caught between expansive world-building and the status quo: Using figured worlds to understand world-building in AI for educational co-design contexts. *Journal of the Learning Sciences* (2025), 1–49. doi:10.1080/10508406.2025.2490514

[25] Paul Coulton, Joseph Lindley, Miriam Sturdee, and Michael Stead. 2017. Design fiction as world building. In *Proceedings of the 3nd Biennial Research Through Design Conference*. ACM, 1–16. doi:10.1145/3240167.3240270

[26] Bronwyn Cumbo, Andrea Gauthier, and Asimina Vasalou. 2024. In the era of climate change: Exploring research at the intersection of children, emerging technology, and environmental sustainability. In *Proceedings of the 23rd Annual ACM Interaction Design and Children Conference*. ACM, 1023–1025. doi:10.1145/3628516.3661153

[27] Bronwyn J Cumbo and Ole Sejer Iversen. 2020. CCI in the wild: designing for environmental stewardship through children's nature-play. In *Proceedings of the interaction design and children conference*. 335–348.

[28] Aayushi Dangol, Robert Wolfe, Akeiylah Dewitt, Ben Chickadel, Julie Kientz, and Sayamindu Dasgupta. 2025. Reading AI and reading the world: Using an interactive AI system to promote Children's understanding of AI bias. *ACM Transactions on Computer-Human Interaction* 32, 6 (2025), 1–30. doi:10.1145/3762807

[29] Sayamindu Dasgupta and Benjamin Mako Hill. 2021. Designing for critical algorithmic literacies. *Algorithmic rights and protections for children* (2021). doi:10.7551/mitpress/13654.003.0010

[30] James Dator. 1979. The futures of cultures and cultures of the future. *Perspectives on Cross Cultural Psychology, Academic Press, New York, NY* (1979), 369–88.

[31] Christian Dindler, Rachel Smith, and Ole Sejer Iversen. 2020. Computational empowerment: participatory design in education. *CoDesign* 16, 1 (2020), 66–80. doi:10.1080/15710882.2020.1722173

[32] Anthony Dunne and Fiona Raby. 2021. *Design noir: The secret life of electronic objects*. Vol. 2. Bloomsbury Publishing. doi:10.5040/9781350070660.ch-005

[33] Anthony Dunne and Fiona Raby. 2024. *Speculative Everything, With a new preface by the authors: Design, Fiction, and Social Dreaming*. MIT press.

[34] Kwamou Eva Feukeu. 2024. A pluriversal definition of Futures Studies: Critical futures studies from margin to centre. In *Handbook of futures studies*. Edward Elgar Publishing, 80–97. doi:10.4337/9781035301607.00013

[35] Iosif Gidiotis. 2026. Speculative futures of artificial intelligence in education: a causal layered analysis of education fiction. *Futures* (2026), 103762. doi:10.1016/j.futures.2026.103762

[36] Eric Greenwald, Maxyn Leitner, and Ning Wang. 2021. Learning artificial intelligence: Insights into how youth encounter and build understanding of AI concepts. In *Proceedings of the AAAI Conference on Artificial Intelligence*, Vol. 35. 15526–15533. doi:10.1609/aaai.v35i17.17828

[37] Katrin Grossmann, James JT Connolly, Małgorzata Dereniowska, Giulio Mattioli, Luca Nitschke, Nicola Thomas, and Anaïs Varo. 2022. From sustainable development to social-ecological justice: Addressing taboos and naturalizations in order to shift perspective. *Environment and Planning E: Nature and Space* 5, 3 (2022), 1405–1427. doi:10.1177/25148486211029427

[38] Amelia Hassoun, Ian Beacock, Sunny Consolvo, Beth Goldberg, Patrick Gage Kelley, and Daniel M Russell. 2023. Practicing information sensibility: How Gen Z engages with online information. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. ACM, 1–17. doi:10.1145/3544548.3581328

[39] Netta Iivari. 2020. Empowering children to make and shape our digital futures–from adults creating technologies to children transforming cultures. *The International Journal of Information and Learning Technology* 37, 5 (2020), 279–293. doi:10.1108/ijilt-03-2020-0023

[40] Netta Iivari, Heidi Hartikainen, Leena Ventä-Olkkonen, Sumita Sharma, Essi Lehto, and Jenni Holappa. 2024. Exploring Digital Futures with Children—Implications for a Future-Oriented IS Research. In *ECIS 2024 Proceedings*. 1–16.

[41] Netta Iivari and Marianne Kinnula. 2016. Inclusive or inflexible: A critical analysis of the school context in supporting children's genuine participation. In *Proceedings of the 9th nordic conference on human-computer interaction*. 1–10.

[42] Netta Iivari, Marianne Kinnula, and Leena Kuure. 2015. With best intentions: A Foucauldian examination on children's genuine participation in ICT design. *Information Technology & People* 28, 2 (2015), 246–280. doi:10.1108/itp-12-2013-0223

[43] Netta Iivari, Sumita Sharma, Leena Ventä-Olkkonen, Tonja Molin-Juustila, Kari Kuutti, Jenni Holappa, and Essi Kinnunen. 2022. Critical agenda driving child–computer interaction research—Taking a stock of the past and envisioning the future. *International Journal of Child-Computer Interaction* 32 (2022), 100408. doi:10.1016/j.ijcci.2021.100408

[44] Sohail Inayatullah. 1990. Deconstructing and reconstructing the future: Predictive, cultural and critical epistemologies. *Futures* 22, 2 (1990), 115–141. doi:10.1016/0016-3287(90)90077-u

[45] Golnaz Arastoopour Irgens, Hazel Vega, Ibrahim Adisa, and Cinamon Bailey. 2022. Characterizing children's conceptual knowledge and computational practices in a critical machine learning educational program. *International Journal of Child-Computer Interaction* 34 (2022), 100541. doi:10.1016/j.ijcci.2022.100541

[46] Ole Sejer Iversen, Rachel Charlotte Smith, and Christian Dindler. 2017. Child as protagonist: Expanding the role of children in participatory design. In *Proceedings of the 2017 conference on interaction design and children*. ACM, 27–37. doi:10.1145/3078072.3079725

[47] Shiyan Jiang, Hengtao Tang, Cansu Tatar, Carolyn P Rosé, and Jie Chao. 2023. High school students' data modeling practices and processes: from modeling unstructured data to evaluating automated decisions. *Learning, Media and Technology* 48, 2 (2023), 350–368. doi:10.1080/17439884.2023.2189735

[48] Yasmin Kafai, R Benjamin Shapiro, Franz Jetzinger, Tilman Michaeli, Matti Tedre, Henriikka Vartiainen, Netta Iivari, Line Have Musaeus, Ole Sejer Iversen, Safinah Ali, et al. 2025. Youth as Designers of Artificial Intelligence and Machine Learning Technologies: What Do We Know About the Opportunities and Challenges of K-12 Students Creating Their Own Applications?. In *Proceedings of the 19th International Conference of the Learning Sciences-ICLS 2025, pp. 2260-2268*. International Society of the Learning Sciences. doi:10.22318/icls2025.362276

[49] Magnus Høholt Kaspersen, Karl-Emil Kjær Bilstrup, Maarten Van Mechelen, Arthur Hjort, Niels Olof Bouvin, and Marianne Graves Petersen. 2022. High school students exploring machine learning and its societal implications: Opportunities and challenges. *International Journal of Child-Computer Interaction* 34 (2022), 100539. doi:10.1016/j.ijcci.2022.100539

[50] Magnus Høholt Kaspersen, Line Have Musaeus, Karl-Emil Kjær Bilstrup, Marianne Graves Petersen, Ole Sejer Iversen, Christian Dindler, and Peter Dalsgaard. 2024. From primary education to premium workforce: Drawing on K-12 approaches for developing AI literacy. In *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems*. ACM, 1–16. doi:10.1145/3613904.3642607

[51] Brian R Keeble. 1988. The Brundtland report:'Our common future'. *Medicine and war* 4, 1 (1988), 17–25. doi:10.1080/07488008808408783

[52] Sadhbh Kenny and Alissa N Antle. 2024. Reimagining AI: Exploring Speculative Design Workshops for Supporting BIPOC Youth Critical AI Literacies. *arXiv preprint arXiv:2407.08740* (2024). doi:10.48550/arXiv.2407.08740

[53] Sadhbh Kenny, Alissa N Antle, Gillian Russell, Alexandra Kitson, and Annemiek Veldhuis. 2025. Beyond the Algorithm: Speculative Approaches to Critical AI Literacies with Diverse Youth. In *Proceedings of the 24th Interaction Design and Children*. ACM, 89–105. doi:10.1145/3713043.3727052

[54] Marianne Kinnula, Eva Durall Gazulla, Noora Hirvonen, Jonna Malmberg, and Lotta Haukipuro. 2024. Nurturing systems thinking among young people by developing business ideas on sustainable AI. *International Journal of Child-Computer Interaction* 40 (2024), 100656. doi:10.1016/j.ijcci.2024.100656

[55] Marianne Kinnula, Netta Iivari, Sirkku Kotilainen, Jussi Okkonen, and Sumita Sharma. 2020. Researchers' toolbox for the future: designing the future of technology with and for children. In *Proceedings of the 2020 ACM Interaction Design and Children Conference: Extended Abstracts*. ACM, 62–68. doi:10.1145/3397617.3398064

[56] Sandjar Kozubaev, Chris Elsden, Noura Howell, Marie Louise Juul Søndergaard, Nick Merrill, Britta Schulte, and Richmond Y Wong. 2020. Expanding modes of reflection in design futuring. In *Proceedings of the 2020 CHI conference on human factors in computing systems*. ACM, 1–15. doi:10.1145/3313831.3376526

[57] Irene Lee, Safinah Ali, Helen Zhang, Daniella DiPaola, and Cynthia Breazeal. 2021. Developing middle school students' AI literacy. In *Proceedings of the 52nd ACM technical symposium on computer science education*. ACM, 191–197. doi:10.1145/3408877.3432513

[58] Li Li. 2024. Advancing Pedagogical Innovation: Integrating Design Fiction into AI Education. *Journal of Digital Life and Learning* 4, 1 (2024), 64–85. doi:10.51357/jdll.v4i1.278

[59] Duri Long and Brian Magerko. 2020. What is AI literacy? Competencies and design considerations. In *Proceedings of the 2020 CHI conference on human factors in computing systems*. ACM, 1–16. doi:10.1145/3313831.3376727

[60] Ruizhe Ma, Ismaila Temitayo Sanusi, Vaishali Mahipal, Joseph E Gonzales, and Fred G Martin. 2023. Developing machine learning algorithm literacy with novel plugged and unplugged approaches. In *Proceedings of the 54th ACM Technical Symposium on Computer Science Education V. 1*. ACM, 298–304. doi:10.1145/3545945.3569772

[61] Matin Mahboob Kanafi, Marianne Kinnula, and Netta Iivari. 2021. Re-defining characteristics of a Design Protagonist: elements of children's design capital. In *Proceedings of the BCS 34th British HCI Conference 2021, UK*. British Computer Society. doi:10.14236/ewic/hci2021.24

[62] Laura Malinverni, Marie-Monique Schaper, Elisa Rubegni, and Mariana Aki Tamashiro. 2025. Scaffolding Children's critical reflection on intelligent technologies: Opportunities from speculative fiction. *International Journal of Child-Computer Interaction* 43 (2025), 100727. doi:10.1016/j.ijcci.2025.100727

[63] Annapaola Marconi, Gianluca Schiavo, Paolo Massa, Eleonora Mencarini, and Giulia Deppieri. 2021. From sustainable mobility to good deeds: Supporting school participation during COVID-19 emergency through a playful education platform. In *Proceedings of the 20th Annual ACM Interaction Design and Children Conference*. ACM, 80–86. doi:10.1145/3459990.3460718

[64] Ramia Mazé. 2020. Design and the Future: Temporal politics of 'making a difference'. In *Design anthropological futures*. Routledge, 37–54. doi:10.4324/9781003085188

[65] Jorge Membrillo-Hernández, Vianney Lara-Prieto, and Patricia Caratozzolo. 2021. Sustainability: A public policy, a concept, or a competence? efforts on the implementation of sustainability as a transversal competence throughout higher education programs. *sustainability* 13, 24 (2021), 13989. doi:10.3390/su132413989

[66] Luis Morales-Navarro. 2025. Investigating Youth's Technical and Ethical Understanding of Generative Language Models When Engaging in Construction and Deconstruction Activities. In *Proceedings of the 24th Interaction Design and Children*. ACM, 1172–1175. doi:10.1145/3713043.3731602

[67] Luis Morales-Navarro, Yasmin Kafai, Vedya Konda, and Danaë Metaxa. 2024. Youth as peer auditors: Engaging teenagers with algorithm auditing of machine learning applications. In *Proceedings of the 23rd Annual ACM Interaction Design and Children Conference*. ACM, 560–573. doi:10.1145/3628516.3655752

[68] Luis Morales-Navarro and Yasmin B Kafai. 2023. Conceptualizing approaches to critical computing education: Inquiry, design, and reimagination. In *Past, present and future of computing education research: A global perspective*. Springer, 521–538. doi:10.1007/978-3-031-25336-2_21

[69] Luis Morales-Navarro and Yasmin B Kafai. 2024. Unpacking approaches to learning and teaching machine learning in K-12 Education: Transparency, ethics, and design activities. In *Proceedings of the 19th WiPSCE Conference on Primary and Secondary Computing Education Research*. 1–10. doi:10.1145/3677619.3678117

[70] Luis Morales-Navarro, Daniel J Noh, and Yasmin B Kafai. 2025. High school students building babyGPTs: Engaging in data practices and addressing ethical issues through the construction of generative language models. *International Journal of Child-Computer Interaction* (2025), 100769. doi:10.35542/osf.io/vesa9_v1

[71] Daniel J Noh, Deborah A Fields, Luis Morales-Navarro, Alexis M Cabrera-Sutch, Yasmin B Kafai, and Danaé Metaxa. 2025. Youth as Advisors in Participatory Design: Situating Teens' Expertise in Everyday Algorithm Auditing with Teachers and Researchers. In *Proceedings of the 24th Interaction Design and Children*. ACM, 415–428. doi:10.1145/3713043.3728849

[72] Viktoriya Olari and Ralf Romeike. 2024. Data-related concepts for artificial intelligence education in K-12. *Computers and Education Open* 7 (2024), 100196.

[73] Chimi Om, Kellie Vella, Tshering Dema, Bernd Ploderer, Karen Murcia, and Margot Brereton. 2025. Storying-Roleplay: A Co-design Method Supporting Young Children to Understand Environmental Problems. In *Proceedings of the 24th Interaction Design and Children*. ACM, 465–478. doi:10.1145/3713043.3728842

[74] Isabella Pu, Prerna Ravi, Linh Dieu Dinh, Chelsea Joe, Caitlin Ogoe, Zixuan Li, Cynthia Breazeal, and Anastasia K Ostrowski. 2025. " How can we learn and use AI at the same time?": Participatory Design of GenAI with High School Students. In *Proceedings of the 24th Interaction Design and Children*. ACM, 204–220. doi:10.1145/3713043.3727057

[75] Ronda Ringfort-Felner, Judith Dörrenbächer, and Marc Hassenzahl. 2025. The Quality of Speculation–A Scoping Review. In *Proceedings of the 2025 ACM Designing Interactive Systems Conference*. ACM, 2373–2394. doi:10.1145/3715336.3735794

[76] Ismaila Temitayo Sanusi, Fred Martin, Ruizhe Ma, Joseph E Gonzales, Vaishali Mahipal, Solomon Sunday Oyelere, Jarkko Suhonen, and Markku Tukiainen. 2024. AI MyData: fostering middle school students' engagement with machine learning through an ethics-infused AI curriculum. *ACM Transactions on Computing Education* 24, 4 (2024), 1–37. doi:10.1145/3702242

[77] Ismaila Temitayo Sanusi and Solomon Sunday Oyelere. 2020. Pedagogies of machine learning in K-12 context. In *2020 IEEE frontiers in education conference (FIE)*. IEEE, 1–8. doi:10.1109/fie44824.2020.9274129

[78] Ismaila Temitayo Sanusi, Solomon Sunday Oyelere, Henriikka Vartiainen, Jarkko Suhonen, and Markku Tukiainen. 2023. A systematic review of teaching and learning machine learning in K-12 education. *Education and Information Technologies* 28, 5 (2023), 5967–5997. doi:10.1007/s10639-022-11416-7

[79] Marie-Monique Schaper, Rachel Charlotte Smith, Mariana Aki Tamashiro, Maarten Van Mechelen, Mille Skovhus Lunding, Karl-Emil Kjær Bilstrup, Magnus Høholt Kaspersen, Kasper Løvborg Jensen, Marianne Graves Petersen, and Ole Sejer Iversen. 2022. Computational empowerment in practice: Scaffolding teenagers' learning about emerging technologies and their ethical and societal impact. *International Journal of Child-Computer Interaction* 34 (2022), 100537. doi:10.1016/j.ijcci.2022.100537

[80] Suzie Wong Scollon. 2004. *Nexus analysis: Discourse and the emerging internet*. Routledge. doi:10.4324/9780203694343

[81] Sumita Sharma, Heidi Hartikainen, Leena Ventä-Olkkonen, Grace Eden, Netta Iivari, Essi Kinnunen, Jenni Holappa, Marianne Kinnula, Tonja Molin-Juustila, Jussi Okkonen, et al. 2022. In pursuit of inclusive and diverse digital futures: exploring the potential of design fiction in education of children. *Interaction Design and Architecture (s)* 51 (2022), 219–248. doi:10.55612/s-5002-051-010

[82] Sumita Sharma, Noura Howell, Leena Ventä-Olkkonen, Netta Iivari, Grace Eden, Heidi Hartikainen, Marianne Kinnula, Eva Durall, Michael Nitsche, Jussi Okkonen, et al. 2024. Promoting criticality with design futuring with young children. In *Proceedings of the 13th nordic conference on human-computer interaction*. ACM, 1–15. doi:10.1145/3679318.3685335

[83] Sumita Sharma, Pauli Klemettilä, and Junko Tanaka. 2025. A robot teacher" is very good for learning, but not for human relationships": Japanese Children's Critical Perspectives Towards Ethical AI Futures. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems*. ACM, 1–20. doi:10.1145/3706598.3713204

[84] Austin Shouli, Ankur Barthwal, Molly Campbell, and Ajay Kumar Shrestha. 2025. Unpacking youth privacy management in AI systems: A privacy calculus model analysis. *IEEE Access* (2025). doi:10.1109/access.2025.3585635/mm1

[85] Jesper Simonsen and Toni Robertson. 2012. *Routledge international handbook of participatory design*. Vol. 711. Routledge New York. doi:10.4324/9780203108543

[86] Rachel Charlotte Smith, Daria Loi, Heike Winschiers-Theophilus, Liesbeth Huybrechts, and Jesper Simonsen. 2025. *Routledge International Handbook of Contemporary Participatory Design*. Taylor & Francis. doi:10.4324/9781003334330

[87] Bruce Sterling. 2009. Cover story design fiction. *interactions* 16, 3 (2009), 20–24. doi:10.1145/1516016.1516021

[88] Amy Stornaiuolo and Erin Hope Whitney. 2018. Writing as worldmaking. *Language Arts* 95, 4 (2018), 205–217. doi:10.58680/la201829525

[89] Mariana A Tamashiro, Marie-Monique A Schaper, Ole Iversen, Maarten Van Mechelen, and Rachel C Smith. 2025. How fiction can support the teaching of societal implications of emerging technologies? A case study on facial recognition activities with K-9 students. *International Journal of Child-Computer Interaction* (2025), 100768. doi:10.1016/j.ijcci.2025.100768

[90] Tiera Tanksley, Angela D. R. Smith, Saloni Sharma, and Earl W. Huff Jr. 2025. "Ethics is not neutral": Understanding Ethical and Responsible AI Design from the Lenses of Black Youth. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems (CHI '25)*. ACM. doi:10.1145/3706598.3713510

[91] Tiera Chante Tanksley. 2024. "We're changing the system with this one": Black students using critical race algorithmic literacies to subvert and survive AI-mediated racism in school. *English Teaching: Practice & Critique* 23, 1 (2024), 36–56. doi:10.1108/etpc-08-2023-0102

[92] Alexandra Thrall, T Philip Nichols, and Kevin R Magill. 2024. Speculative frictions: writing civic futures after AI. *English Teaching: Practice & Critique* 23, 1 (2024), 67–82. doi:10.1108/etpc-08-2023-0095

[93] Lorraine Underwood, Karen Smith, Elisa Rubegni, and Joe Finney. 2022. Energy in schools: Empowering children to deliver behavioural change for sustainability. In *Proceedings of the 21st Annual ACM Interaction Design and Children Conference*. ACM, 308–314. doi:10.1145/3501712.3529744

[94] Sveva Valguarnera and Leandro Soares Guedes. 2020. Two in a Pod: promoting sustainability and healthy eating in children through smart gardening. In *Proceedings of the 2020 ACM Interaction Design and Children Conference: Extended Abstracts*. ACM, 241–246. doi:10.1145/3397617.3398022

[95] Henriikka Vartiainen and Matti Tedre. 2025. The CEDE Model: A Learning-Sciences Based Approach for Critical and Transformative K–12 AI Education. In *Proceedings of the 25th Koli Calling International Conference on Computing Education Research (Koli Calling '25)*. ACM. doi:10.1145/3769994.3770026

[96] Henriikka Vartiainen, Tapani Toivonen, Ilkka Jormanainen, Juho Kahila, Matti Tedre, and Teemu Valtonen. 2021. Machine learning for middle schoolers: Learning through data-driven design. *International Journal of Child-Computer Interaction* 29 (2021), 100281. doi:10.1016/j.ijcci.2021.100281

[97] Asimina Vasalou and Andrea Gauthier. 2023. The role of CCI in supporting children's engagement with environmental sustainability at a time of climate crisis. *International Journal of Child-Computer Interaction* 38 (2023), 100605. doi:10.1016/j.ijcci.2023.100605

[98] Annemiek Veldhuis, Priscilla Y Lo, Sadhbh Kenny, and Alissa N Antle. 2025. Critical Artificial Intelligence literacy: A scoping review and framework synthesis. *International Journal of Child-Computer Interaction* 43 (2025), 100708. doi:10.1016/j.ijcci.2024.100708

[99] Leena Ventä-Olkkonen, Netta Iivari, Sumita Sharma, Tonja Molin-Juustila, Kari Kuutti, Nina Juustila-Cevirel, Essi Kinnunen, and Jenni Holappa. 2021. Nowhere to now-here: empowering children to reimagine bully prevention at schools using critical design fiction: exploring the potential of participatory, empowering design fiction in collaboration with children. In *Proceedings of the 2021 ACM Designing Interactive Systems Conference*. ACM, 734–748. doi:10.1145/3461778.3462044

[100] Jon M Wargo and Jasmine Alvarado. 2020. Making as worlding: young children composing change through speculative design. *Literacy* 54, 2 (2020), 13–21. doi:10.1111/lit.12209

[101] Petra F Weixelbraun, Barbara Göbl, Matthias Steinböck, Mirjam Duvivié, and Fares Kayali. 2024. Discussing the Protagonist Role of Students in Game-Based Learning. *Proceedings of the ACM on Human-Computer Interaction* 8, CHI PLAY (2024), 1–24. doi:10.1145/3677065