Critical Tools for Machine Learning

Critical Tools for Machine Learning: Working with Intersectional Critical Concepts in Machine Learning Systems Design

Goda Klumbytė
Faculty of Electrical Engineering and Computer Science, University of Kassel
goda.klumbyte@uni-kassel.de

Claude Draude
Faculty of Electrical Engineering and Computer Science, University of Kassel
claude.draude@uni-kassel.de

Alex S. Taylor
Centre for Human-Computer Interaction Design, City, University of London
alex.taylor@city.ac.uk

ABSTRACT

This paper investigates how intersectional critical theoretical concepts from social sciences and humanities research can be worked with in machine learning systems design. It does so by presenting a case study of a series of speculative design workshops, conducted in 2021. These workshops drew on intersectional feminist methodologies to construct interdisciplinary interventions in the design of machine learning systems, towards more inclusive, accountable, and contextualized systems design. The concepts of "situating/situated knowledges", "figuration", "diffraction", and "critical fabulation/speculation" were taken up as theoretical and methodological tools for concept-led design workshops. This paper presents the design framework of the workshops and highlights tensions and possibilities with regards to interdisciplinary machine learning systems design towards more inclusive, contextualized, and accountable systems. It discusses the role that critical theoretical concepts can play in a design process and shows how such concepts can work as methodological tools that nonetheless require an open-ended experimental space to function. It presents insights and discussion points regarding what it means to work with critical intersectional knowledge that is inextricably connected to its historical and socio-political roots, and how this reframes what it might mean to design fair and accountable systems.

CCS CONCEPTS

• **Human-centered computing** → Interaction design; Interaction design theory, concepts and paradigms; • **Computing methodologies** → Machine learning.

KEYWORDS

Machine learning systems design, Intersectionality, Feminist epistemologies, Experimental practice, Interdisciplinary methodologies

ACM Reference Format:
Goda Klumbytė, Claude Draude, and Alex S. Taylor. 2022. Critical Tools for Machine Learning: Working with Intersectional Critical Concepts in Machine Learning Systems Design. In *2022 ACM Conference on Fairness, Accountability, and Transparency (FAccT '22), June 21–24, 2022, Seoul, Republic of Korea.* ACM, New York, NY, USA, 14 pages. https://doi.org/10.1145/3531146.3533207

FAccT '22, June 21–24, 2022, Seoul, Republic of Korea

ACM ISBN 978-1-4503-9352-2/22/06.
https://doi.org/10.1145/3531146.3533207

1 INTRODUCTION

There is by now a plethora of resources that address fairness and bias in machine learning (ML) systems, ranging from more technical de-biasing solutions [13, 22], implementation of criteria such as demographic parity [47, 68] and equal opportunity [47], to structural approaches that build fairness on transparency, explainability and understandability [68, 69, 74, 88]. Suggestions to pay close attention to data have also been expressed [93], and politics and processes surrounding the production of algorithmic systems have been called to attention [21, 82].

Furthermore, algorithmic and design justice scholars have highlighted the need for accountable and contextualized systems that go beyond fairness towards equity and justice [11, 26, 58] and espoused algorithmic harm rather than only bias as an organizing concept [28, 34, 71]. Fairness, accountability and transparency (FAccT) researchers have also suggested that in order to address ML systems in context and orient them towards justice, equity and ethics, it is important to pay attention to whose knowledge is valued and included [76], both in terms of the kind of constituencies and political actors that get to play a role in systems design, as well as the kind of disciplinary knowledges that are brought to the discussion [89]. Calls for intersectional and feminist approaches to ML/AI research and systems design exemplify this need for inter- and trans-disciplinary openness and collaboration [24, 32, 36, 66, 67]. The FAccT community in particular has showcased research that brings together approaches, concepts and methods from social sciences and humanities together with ML research and design, or that explicitly starts from humanities and social sciences perspectives when analyzing implications of ML systems and proposing new development approaches [38, 48, 56, 57, 59].

This contribution is positioned within the debates and existing research on inter- and trans-disciplinary approaches to fairness, accountability, and justice in ML/AI systems. It contributes to the calls for intersectional computing and design [31, 33, 63, 80], and post-colonial, decolonial, and race-critical perspectives in computing and AI/ML [2, 55, 72, 75, 81]. We offer one example of *how* to *work with* critical concepts from social sciences and humanities, specifically intersectional feminist thought, towards more inclusive, contextualized, and accountable ML systems design. We describe a process organized through a series of workshops where participants

were invited to experiment and work with four concepts: "situated knowledges" [39, 54], "figurations" [4, 15, 41], "diffraction" [6, 7] and "critical speculation/fabulation" [49, 52, 78]. Participants used these as guiding principles for ML systems design and applied them to speculative ML design scenarios. These concepts stem from critical feminist, anti-racist, anti-colonial, intersectional research and have been fundamental to feminist epistemology. In the workshops they were presented as tools to help envision ML systems design as a more contextualized, reflexive, and interdisciplinary process.

Reflecting on the workshop design and the accompanying tensions and possibilities, we discuss the role that critical theoretical concepts can play in a design process. We argue that critical concepts can serve as orienting and concretizing moments in design – as methodological tools that help guide the design towards inclusion, contextualization, and accountability, provided that the design participants are motivated to follow such direction. We show that working with critical concepts as design tools requires an open-ended experimental space that resists formalization, thus creating both constraints and possibilities. We close the paper with some insights and discussion points regarding what it means to work with critical intersectional knowledge that is inextricably connected to its historical and socio-political roots, and how this changes what it might mean to design fair and accountable systems.

2 THEORETICAL AND METHODOLOGICAL BACKGROUND

2.1 Critical Concepts

The premise of the work we present is that while computer science has developed sophisticated technical tools in the field of ML/AI, challenges remain when it comes to systemic socio-cultural bias. Critical theories, particularly intersectional feminist, postcolonial, and anti-racist theories show that such bias is embedded in sociotechnical systems. These theoretical perspectives also explain how systems not only reproduce but can give rise to new inequalities. Crucially, this scholarship has led to conceptual understandings and tools that respond to bias and inequality.

It is this systemic, sociotechnical understanding of bias and inequalities that provides the foundation to our research. In particular, we draw on feminist science and technology studies that see knowledge production as a process that is socially and culturally contingent and shifting, and involves many actors (both human and technological) [10, 42, 64, 86]. Feminist epistemologies highlight that technical tools are not objective and neutral. They are part of larger sociotechnical entanglements and thus give rise to knowledge and practices that are partial, situated, embedded and embodied, as well as implicated in a broader nexus of power relations [39, 44, 45]. It was through efforts to build on this knowledge and theorizing, that our workshops took up the concepts of "situated knowledges" ("situating"), "figuration" ("figuring"), "diffraction" and "critical fabulation/speculation". These were applied as theoretical concepts or tools to help re-imagine machine learning systems and their design as part of more contextualized, interdisciplinary processes.

"Situating" as a methodological approach emerged from Black feminist scholarship [54], and feminist epistemology [39, 44, 44]. Closely resonating with "standpoint theory" [46], it argues that scientific and technological knowledge is partial and situated in specific disciplinary, cultural and political contexts. Science (and technological knowledge production) therefore generates not objective and neutral but *partial* perspectives [39, 40]. This partiality should be embraced as the basis of accountability and a way to recognize that knowledge can both help tackle and reproduce intersecting oppressions. For example, Patricia Hill Collins has shown that Black feminist thought emerges from specific standpoints or situated perspectives of Black women and provides a specific angle of vision, characterized, among other things, by attention to collectivity, "outsider within" perspective, embracing of different sites and means of knowledge production (such as literature and other non-academic sites), and prioritization of experiential and dialogical knowledge [25, 54]. Situating as a method in ML systems design [30, 57] allows contextual factors to be taken into account in a more systemic way and also for technologies to be developed that are more accountable for their own specific partial ways of generating knowledge. Situating is akin here to "contextualization" but in ways that considers multiple contexts and remains aware of the partiality of perspectives generated through them.

"Figurations" (also sometimes called "conceptual personae") have been crucial to feminist theorizing and politics. Figurations are mappings [15–17] that address particular historical, political and material locations [14]. They are figures that bring together or embody meanings, practices, histories, power relations, and political controversies [43]. For example, a cyborg has been a potent figuration [41] that feminist theories have used to conceptualize relations between nature/culture, body/technology, and other binaries. While having originated in cybernetic and Cold War military research, the cyborg has given inspiration to cyberfeminism and a kind of "feminist cyborg politics" that highlight that there are no "innocent" or "pure" political positions (just like the cyborg is not pure human or pure machine) and that technological development is inextricably entangled with politics. Technologies can be seen as materialized figurations that bring together both technological objects and clusters of meaning surrounding them [85]. As a methodology, figurations and figuring allow for imagining and developing new design perspectives that are sensitive to power hierarchies and contextual specificity [61]. In computing it resonates with the method of persona development [35], however it expands the persona to include fictional and non-human figures and necessitates taking into account historical and cultural context [61].

"Diffraction" as a feminist methodological principle [6, 7, 39] provides a model for trans-disciplinary research and practice. It suggests "reading" diverse theoretical approaches "through" one another as an experimental mode of knowledge. For instance, Karen Barad's diffractive work on using insights from quantum physics to advance feminist theory of difference and using feminist epistemology to re-interpret the role of the apparatus in Bohr's quantum mechanics is one such example of diffractive research [7]. Furthermore, Barad shows that diffraction as a theoretical concept highlights the role of technologies, disciplinary perspectives and tools in the production of knowledge, or in other words: how these tools and technologies act as "diffraction apparati" [7]. This means that technological solutions, as well as disciplinary perspectives and methods are not neutral but actively affect the "objects" that they investigate and the phenomena or problems that they are

trying to solve. Used in this way, designing with diffraction necessitates attention to the agency of technology and the role of different disciplinary knowledges [9].

"Critical fabulation" and "critical speculation" brings together Black critical theoretical thought [49, 52] and critical and speculative design scholarship [18, 27, 78]. Critical fabulation entails speculative thought and imagination – techniques and faculties crucial for the design of new technologies – and engages these activities to address absent imaginaries and power hierarchies. In particular, it is a methodological tool to address *missing* imaginaries and missing perspectives that provide an alternative to "white prototypicality" [3, 20, 60] and the solipsistic "I-methodology" [77]. For example, historian Saidiya Hartman throughout her work uses critical fabulation as a method to work with archives and the stories that are missing from them. She has explored, among other sources, archives of transatlantic slavery that exclude the voices and stories of enslaved people [52] and the personal histories of "wayward" black women in the late 19th and early 20th century that are so often reduced to objects of criminality [50]. This approach has been taken up, in more and less explicitly political ways, in critical design perspectives [8] and speculative design as a broader practice.

These four concepts, then, allow for the analysis and interventions into sociopolitical and sociotechnical relations that ML systems are part of. They also endorse the generative potential of technologies, recognizing that as hybrid (i.e. both social and technical) objects they open up the possibilities for new relations and interconnections. Building on these concepts in our work, we approach fairness and accountability in ML as irreducible to their formal definitions and dependent on the societal and technical, cultural and informational, contextual and formal-abstract factors (in line with work such as [11, 23, 58, 82] and much of the work presented at the CRAFT sessions of ACM FAccT). This in turn means that inclusive, accountable, and contextualized design requires an inter- and trans-disciplinary approach spanning humanities, social sciences, and computing. Such an approach also expands debates around fairness and accountability and orients them towards responsibility and justice.

2.2 Experimental Practice, Co-Creation, Speculative Design, and Critical Technical Practice

Methodologically, this work builds on research through design [92], speculative [5] and participatory design [73], and critical technical practice [1]. Participatory design and co-creation were the foundations of the group-work during workshops. We relied on these approaches to both encourage collective thinking, exchange and shaping of responsibility, as well as to foster the practice of using participatory design as a methodology for ML systems design [59, 62]. Speculative design was employed as a way to think through ethical concerns that emerge in ML design and to open possibilities of different futures [29, 90]. Significantly, since the emphasis was on exploring critical concepts and how to work with them in ML systems design, speculative design allowed a focus on the methodological implications without the constraints of needing to deliver a working prototype.

Critical technical practice, as an approach geared towards questioning implicit design norms and premises of knowledge production in and through AI—with its disciplinary in/exclusions and rigidities—also played an important role in developing the workshop series. We aimed to create a space where procedural and disciplinary conventions—such as the standard steps of the workflow of ML design (or lifecycle) as well as division of knowledge and labor between domain experts, data scientists, operations managers and others partaking in or being affected by ML system development—can be challenged and diffracted. Critical technical practice as an approach here thus resonated with the reflexive impetus of the critical theoretical concepts selected.

Research through design was an approach that we relied on to open up space for experimentation in practice without a predetermined outcome. We wanted to see what might emerge once the constraints to optimize ML systems for profit, or even to deliver something functional, were lifted and instead space was created to closely work with critical concepts. Furthermore, research through design included building insights not only during the workshops but also necessitated paying attention to the phase of the preparation for and design of the workshop framework—part of research that revealed valuable insights on the process of bringing more interdisciplinarity to ML design methodological toolkit.

A separate note needs to be made on the process that we initially called the "translation" of concepts and their meanings between the different disciplines: intersectional critical theories and ML systems design. Our initial goal was to explore how to "use" the knowledge embedded in critical theories and to seek to "translate" this knowledge, or "adapt" it, for the purposes of ML system design. We changed this perspective from "using" critical concepts to "working with" them. As a consequence, we came to understand our project as spanning disciplines, instead of seeking to transplant knowledge from one domain to another. Nonetheless, some work of attuning—if not translation—was required in order to create the conditions for the workshop participants to develop an understanding of the critical concepts and to open the possibility for working with them in a pragmatic way.

3 FRAMEWORK OF THE WORKSHOPS

The workshops took place over the course of two weekends, four workshops in total, four hours each, all taking place online. Nine participants were recruited, primarily students from the fields of HCI and ML.[1] The workshops were framed both as a research project as well as a possibility for participants to learn, in an experimental setting, some methodologies towards more contextualized, inclusive, and accountable systems design. The key objectives of the workshops were: (1) to introduce critical intersectional methodologies as potential design intervention tools; and (2) to generate creative ways to work with such methodologies as design approaches and tools.

[1] For participant recruitment, data collection, anonymization, and storage and overall research ethics protocols, the requirements of City, University of London and University of Kassel were followed.

3.1 Workshop Pedagogy and Design

All work during the workshops was organized in two smaller groups of four and five people, in which participants worked throughout the workshop series. Each workshop was dedicated to exploring and working with one critical concept. The practical assignment was to develop a speculative prototype of an ML system. The first workshop included choosing a speculative scenario for development, and the last workshop featured final presentations of the process and the speculative systems developed by the two participant groups. The rationale of the workshops was that participants, having chosen the speculative design scenario during the first workshop, would continue developing it throughout the series.

Each workshop had a similar structure that consisted of:

1. Introduction of the concept. This included theoretical input by the facilitators explaining the concept and collective reading of selected texts. All concepts were contextualized within critical intersectional research, with readings selected from academic and literary texts.[2]
2. Exploration of the concept through hands-on exercises. These exercises were designed by the facilitators to present ways to actualize the concept through engaged activities.
3. Working with the concept through design exercises. These activities consisted of possible ways, developed by facilitators, of working with the concept as a design tool or method, specifically in the context of a speculative design scenario.
4. Group sharing, reflection and discussion.

The workshops took place using a video conferencing platform and collective whiteboard platform. Each workshop also had a key issue and the main research question that corresponded to the selected concept. The exercises were carefully drafted as partial perspectives or examples of how they can be interpreted and explored. They were explained at the beginning of each workshop after the theoretical introduction to the concept.

Groupwork took place in breakout rooms. The facilitators did not take an active part in discussions or conversations within the groups so as to minimize the influence of the facilitators' own understanding of the concept. The idea was to foster an experimental atmosphere by disrupting the traditional idea of ownership and expertise: instead of portraying the facilitators as experts on the critical concepts, we strived to create an open space for interpretation. The participants were advised that the workshops were meant to act as experimental spaces for trying out new approaches and that there was no formalized method for how to work with the concepts—instead, they were encouraged to develop their own understanding and interpretation of them through exercises and discussion. We encouraged participants to document their discussions with images and notes in their group workspaces on the shared whiteboard.

The following sections present the guiding issues and questions that framed each workshop (WS) as well as how each concept was actualized through exploratory and design exercises. For detailed exercise descriptions that were presented to participants, see the appendix.

3.2 WS1: Situated Knowledges/Situating

Issue: ML systems are often detached from the socio-political-cultural context in which they operate. This detachment, or decontextualization, happens at the levels of data collection and use but also can be reproduced during the design process. This decontextualization enables ML systems to lay claims to objectivity and impartiality. This in turn might cause issues with ascribing accountability when ML systems generate erroneous or harmful results.

Methodology/concept: Situated knowledges and situating. The concept expresses the claim that in practice "pure objectivity" gives way to partial perspectives. Such partial perspectives acknowledge that knowledge always comes from somewhere: specific context, specific bodies, specific forms and instruments that generate it. Grounding knowledge in partiality allows for more accountability since such knowledge is contextualized, i.e., situated. Situated knowledges allow for an increased response-ability: capacity to respond towards their context.

Guiding question: How can ML systems design be situated towards greater acknowledgement of context and better forms of response-ability?

Main activity: Mapping contextual positioning of oneself and constructing shared situatedness.

Additional preparation prompt: List three objects that express something about your significant connections to the world: what kind of communities are significant to you, or do you feel you belong to; what kind of disciplinary or professional background do you have; what kind of materials do you like working with or are you fond of.

Exploratory exercises for this workshop consisted of three mapping/situating exercises. The first two (the only exercises that participants did individually) were **Mapping Positionality ("Rooting")** and **Mapping Perspectives ("Shifting")**. The "rooting" exercise was intended for participants to map out their own positionality by selecting three objects that represented their significant relations to the world and the reasons for their choices. They were asked to draw a personal map that reflected places, people or communities, professional and other significant activities, and values and questions that were important to them. In the "shifting" exercise we asked participants to consider what shifts are present in their map by reflecting on what their personal positionality implies in terms of what kind of knowledge and concerns are important or accessible to them, which kind of issues and questions are less accessible, which communities or environments do they feel responsible to/for, and what would happen if they shifted their position on this map (what kind of things would then become accessible/inaccessible). The idea was that each mapping, each positionality brings with it a perspective and that it is important to address what kind of

[2]The texts selected for collective reading were the following: for WS1 on Situated Knowledges - excerpts from Patricia Hill Collins' *Black Feminist Thought: Knowledge, Consciousness and the Politics of Empowerment* [54] and Donna Haraway's essay "Situated Knowledges: The Science Question in Feminism and the Privilege of Partial Perspective" [39] For WS2 on Figurations: an excerpt from Gloria Anzalduá's *Borderlands/ La frontera: The New Mestiza* [4]. For WS3 on Diffraction: excerpt from Karen Barad's *Meeting the Universe Halfway: Quantum Physics and the Entanglement of Matter and Meaning* [7] For WS4 on Critical Fabulation/Speculation: excerpts from Saidiya Hartman's article "Venus in Two Acts" [49] and Francesca Spektor's and Sarah Fox's article "The 'Working Body': Interrogating and Reimagining the Productivist Impulses of Transhumanism through Crip-Centered Speculative Design" [83]. Additional references for suggested reading were also provided ([41, 61] for WS2, [19, 65] for WS3, [18, 51, 79] for WS4).

perspectives and responsibilities are implied or chosen in specific positionalities. The ideas of rooting and shifting were based on propositions for feminist activism and politics by feminist scholar Nira Yuval-Davis [91].

The design exercise **Situating Towards Collective Systems** followed the two exploratory exercises and invited participants to construct a shared map or shared positionality of their work group. They did that by exploring their commonalities with regards to communities, disciplines, concerns that they share and also the parts where their contexts and experiences diverge. These discussions were then summarized and visually represented in a collective map that identified shared contexts and concerns or significant communities, shared values and perspectives, and significant differences in the group. The primary idea for this exercise was to allow participants to both get to know each other but also build a shared basis that included not only their similarities but also their differences, and to help them become aware of their partial perspectives.

At the end of this workshop, participants selected speculative ML system scenarios that they wanted to work on in their group. They could either chose from a list of suggestions, come up with their own scenario, or let the Oracle of Transfeminist Technologies[3] – a creative card system presented during one of the social events of 2020 Conference on Fairness, Accountability, and Transparency – decide the parameters of the system. The pre-defined suggestions included typical examples of systems that are built using ML: a credit score system for a local government, a recommender system for a dating website, a prediction system for targeted healthcare. Participants could also mix-and match the different options. It was also important to tie these scenarios to their respective positionalities and shared values and concerns.

3.3 WS2: Figurations/Figuring

Issue: ML systems are perceived as ahistorical, disembodied, neutral tools that are devoid of power relations. Furthermore, being computational, these systems are also perceived as somewhat immaterial, operating in the abstract "digital space". The more material aspects of systems come in through user modelling. Nonetheless, here the explicit attention is usually given to the interaction between the user and the technical system where both are perceived as distinct entities.

Methodology/concept: Figurations a.k.a. conceptual personae. Figurations are "materialistic mappings of situated, i.e., embedded and embodied, social positions" [16] – in other words, material-discursive entities that account for particular historical, political, and material locations. They stitch together meanings and practices. Technologies are materialized figurations that bring together both actual physical technologies and clusters of meaning (narratives, discourses, imaginaries) surrounding them, which together form more or less stable assemblages or configurations.

Guiding question: How can ML systems design be geared towards recognizing and acting from the embodied, embedded and power-laden conditions of its position as well as the effects of ML systems?

Main activity: Building a figuration of ML system.

[3]See https://www.transfeministech.codingrights.org/.

The exploratory exercises engaged with in this WS were **Figuring Materials** and **Figuring Stories**. The former required participants to look back at the materials that they indicated in additional preparatory prompts and investigate how these materials operate as figurations: what are they made of, how they come to be made, how do they feel, how can they be used, what kind of stories would they tell and perspectives would they open up if they were animate. Due to time constraints, *Figuring Materials* exercise was not applied, and the groups worked on the second exercise, *Figuring Stories*. For that, they were tasked with weaving a figuration as a structuring metaphor or conceptual imaginary that emerges from the shared concerns or communities that they identified in their collective positionality map (see WS1). This was done in two parts: first, by looking more closely into and mapping out the shared concerns and/or communities that they had identified: the people or significant stakeholders that play a role, imaginaries or personified figures of speech that get used in those contexts, stories and narratives that are of importance, and even mythological, cultural, and literary figures that are associated with them. Then, for the second part, participants looked for a figuration that might emerge from these mappings that would be able to embody or otherwise play a role in the context of their shared concern/community. They investigated these figurations and structural metaphors by noting what role they might play and what might they signify.

This was then further applied in a design exercise, **Figuring Systems**. Participants addressed their own ML system/scenario as a figuration itself, inquiring into its material base (what kind of elements is it made of), its function, and its perspective and positionality. They drew and mapped out what kind of discursive elements it was made of: the stories and narratives that it invoked, how the system-as-figuration was related to the broader context in which it was supposed to function, and its impacts and relations to power in that context. Eventually, these exploratory and design exercises converged into an exploration of a figuration that was either embodying or acting as a structural metaphor for the speculative ML systems that the groups were developing.

Towards the end of this WS, participants concretized their ML systems by defining: their goal/task; the significant stakeholders that would take part in the design process; the main operations of the system; and a slogan or a tag line.

3.4 WS3: Diffraction

Issue: ML systems design often lacks interdisciplinarity and an in-depth understanding of the effects of its specific perspective. This prevents ML systems from attracting a wider range of expertise (or knowledge practices) and from being seen and understood in broader sociotechnical terms.

Methodology/concept: Diffracting the apparatus and tracing patterns of relations. Diffraction suggests that the apparatus of measurement and the tools of knowledge production themselves play one of the constitutive roles in generating knowledge. Diffraction generates patterns of interference, i.e., patterns of difference and relation, entangling processes of meaning-making and technology.

Guiding question: What are the effects of ML systems as complex apparati of knowledge production?

Main activity: Mapping ML system as a diffractive apparatus.

This workshop started with participants further clarifying their ML system and its structure by discussing and deciding about its tasks, optimization criteria, the data, and the type of algorithm(s) it would require (e.g. natural language processing, neural networks, classification, or regression models, etc. – these could be as concrete or abstract as groups wanted). They were also introduced to the classical ML design workflow. After further clarifying the parameters of their systems, participants then engaged in a combined design and exploratory exercise **Diffracting ML**. Pointing out that in the diffraction experiments in physics the experimental apparatus itself actively shapes the experimental situation and its outcomes, we asked participants to address their ML system as a diffractive apparatus. This consisted of three steps: identifying the societal/contextual, technical, discursive/value, and operational/logical elements of their ML-system-as-diffractive-apparatus and then diagramming the relations between those elements. Three categories of relations were suggested: constructive (one element constructs another), disruptive, and relations of interference. Participants came up with visual schematics of elements of their system and relations between them. Based on this, they reflected on broader societal effects of their systems (would it introduce new or change existing values? Intervene in the way we understand certain phenomena? Introduce new relations between communities or phenomena?). The final step was to go back to the characteristics of the system that they identified at the beginning of the WS and reflect on what they might want to change or adjust, in view of possible diffractive effects of their ML systems.

3.5 WS4: Critical Fabulation and Speculation

Issue: ML systems are thick material-discursive knots, however in design they are regarded as discourse- and story-free objects/systems. Furthermore, biases emerging in ML systems often signal not only negative impacts but exclusions of specific voices, perspectives, and histories.

Methodology/concept: Critical fabulation/speculation. Critical fabulation entails speculative thought and imagination but it positions those activities against the absent imaginaries and loci of power hierarchies. In particular, critical fabulations are methodological tools to address missing imaginaries and missing perspectives. They provide an alternative to "white prototypicality" [3] and the "I-methodology" [77] that is prevalent in technology design.

Guiding question: How can the specific grounded perspectives and forms of collective and personal accountability be addressed in ML systems design? How can these perspectives be designed with more inclusion?

Main activity: Fabulating and prototyping speculative ML systems.

The last workshop consisted of three exercises and a final presentation. First, participants engaged in **Writing Design Narratives** – short fabulated stories about their ML system and its use that included anticipated users and/or stakeholders, scenarios of use, potentially hinting at the role of the figuration that they identified in WS2, and what kind of difference the system might make for its users. Then, they spent some time on **Analyzing Narratives** from a situated and critical perspective. Participants were asked to look more closely both at who figures prominently in their narrative, from which perspective is it told, but also who or what perspectives might be *missing* or hidden from their narratives. By going through a series of questions identified in the exercise description, participants made notes and discussed what perspectives were represented in their narratives. They also discussed the broader significance of their ML system for those who might or might not be represented. Finally, they reflected on their own stakes in the design of this (type of) ML system, connecting it back to their personal positionality.

Following this, participants did a second design exercise, **Fabulating with Critical Perspectives**. Building on the introduction of the critical fabulation concept as a political tool that foregrounds missing perspectives as well as accountability for the stories told and futures envisioned, participants were asked to construct a new fabulation by incorporating the insights from their critical analysis. Specifically, we asked them to consider how their system could center less heard voices or address absences, alternative histories and practices, and collective and personal responsibility involved in its design. They were tasked to adopt a different perspective from the initial narrative and were encouraged to find collective ways of writing.

4 WORKSHOP RESULTS AND KEY INSIGHTS

Throughout the workshops participants developed two speculative ML systems: "Time Capsule" and "Therisius". "Time Capsule", with a tag line "Body Speaks, We Listen", was an assistive technology aimed at empowering people in communicating about their health, particularly to healthcare professionals. The system was designed to use natural language processing, document matching and recommendation algorithms to generate a multimedia portfolio. This may contain images, sounds and video material to complement patient's descriptions of their health. Driven by values such as body autonomy and patient empowerment among others, the system would be a continuously evolving, co-designed (by artists, patients, doctors, researchers, and other stakeholders) tool. It would take multiple knowledge sources into account and center patient's knowledge of their health. The "Therisius" (variably also spelled as "Therisuis" and "Theresias") combined the words "thesaurus" and the name of Greek mythological figure Tiresias. It was designed as a digital educational tool, a crowd-sourced "entangled dictionary" and a "suggestor" to help users learn and understand the complex histories and implications of language. The tool would allow users to create their own dictionary pages, learn about history, context, and implications of various terms, and share knowledge. For instance, the tool might suggest to a user that a particular word be considered offensive and would substantiate suggestions with information about the context and histories of the word. Motivated by the shared values of inclusion, respect, and awareness, the creators of this tool envisioned it as an aid that maintains and supports user agency, facilitates understanding of different experiences and nuances around language, and supports multiple perspectives. These speculative prototypes were presented visually as mind maps (e.g. Figure 1) that the groups created on the last day of the workshop series. Oral presentations also included the reading of speculative fabulations that accompanied them.

Overall participants spoke about the workshop positively. Notable was the openness to engage with critical concepts from a very

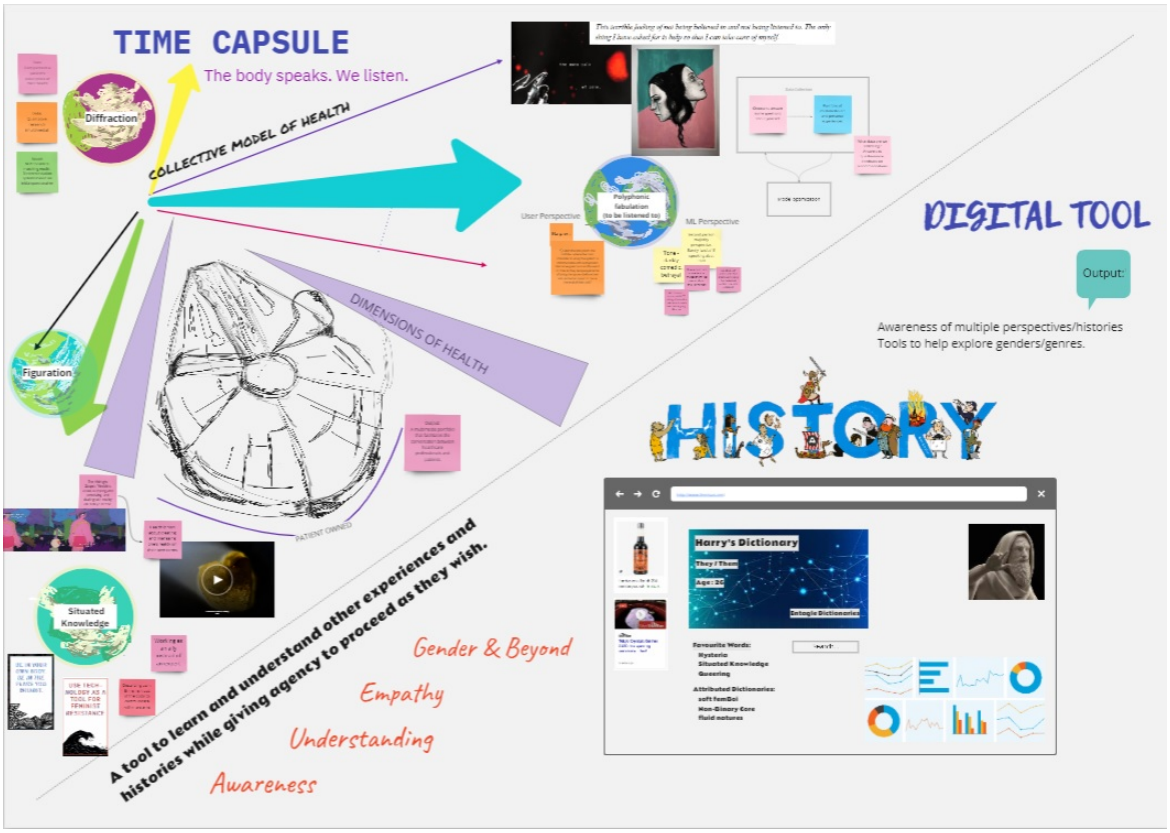

Figure 1: Visual mind maps representing the two speculative projects developed by workshop participants.

different disciplinary field than participants' own backgrounds and the growing critical engagement with the concepts and the projects they were developing. Participants found it challenging to work through and with the concepts, which we attribute not only to the use of different disciplinary language, but also to the fact that these concepts required different modes of engagement than might be normally expected in computer science and/or ML. For instance, situating oneself as a person instead of striving for a neutral, impersonal perspective, being open to failure and the lack of guaranteed results, focusing on the process and discussion as modes of engagement and critical thinking, retaining instead of simplifying complexity in design results and process, bridging and shifting between the abstract concepts and concrete implementations – all these were reactions, reflections and notes that came up as challenging points in participants' discussions and feedback. We see these challenges however also as opportunities to develop more interdisciplinary curricula and epistemically more inclusive design methodologies when it comes to ML systems.

The key workshop results that we would like to highlight here therefore pertain not so much to the speculative technology prototypes that the participants developed, but to the *process of working with critical concepts in ML systems design*, and how this process pertains to fostering ML systems design towards more inclusive, contextualized, and accountable procedures and results. This process highlighted several key tensions that were notable in the sorts of questions that the participants were struggling with, and in the work of preparing the workshop framework and actualizing the concepts through the exercises. In other words, we want to highlight the overall challenges and opportunities that emerge in the construction of more interdisciplinary ML systems design approaches. Namely, tensions emerged between:

I. Actualization and understandability of concepts: how to actualize them in ways that do justice to their foundation without overburdening participants with too much information.
II. Formalization and open-endedness: how to balance the need for formalized procedures that the disciplinary setting of computing often requires and the interpretative flexibility that is needed for critical reflection and for the work with critical theoretical concepts.
III. Direct applicability of concepts-as-methods and more incremental orientation of the design process: how to navigate the disciplinary push towards applicability and pragmatics of computing methods versus the kind of structural, slower-paced, and perspectival change that critical concepts can generate.

Each concept had a different role to play in this regard and brought different challenges and possibilities. Overall, however, we note that those tensions were generative in that they opened space for discussion and reflection on accountability and ethical

choice, collective responsibility, and the challenges that come with interdisciplinary work.

4.1 Fostering Contextualization, Inclusivity, and Accountability Through Concept-Led Design

Situated knowledges as a concept was actualized through multiple mappings that allowed participants to both position themselves, reflect on their perspectives, as well as develop a collective positionality and identify shared concerns as a group. The important aspect of this process was that participants were encouraged to account for their positionality beyond standard identity categories (that they were free to choose to disclose or not) and also reflect on how their positionality (in terms of chosen values, discipline, as well as social positioning). Through situated, partial perspectives, they were invited to examine the kinds of claims their personal and disciplinary tools and practices afforded. Furthermore, this required participants to reflect on their own accountability: who did they feel that they were accountable to? Which communities and concerns really mattered to them? We consider such reflection to be a crucial part of moving towards both context-sensitivity and accountability of and in ML systems.

In a similar vein, the workshop on diffraction was designed to invite participants to address the possible *effects* of the systems they were designing. In this framework, technologies are understood as material configurations or assemblages that are not simply inserted into the social fabric but are agentive and active in constructing, disrupting and the overall shaping of existing relations as well as introducing new forms of social and material relations. This led participants to address contextualization of ML systems and their effects through a much wider scope, including social context and institutions, values and discourses, technological and disciplinary knowledges and operations, and the overall operational logic that gets introduced into the phenomena that ML systems engage.

Working with figurations helped generate the understanding of and envision ML systems as material-discursive mappings that are rooted in specific imaginaries. The way this concept was actualized closely resonated with the idea of "structuring metaphors" [1] and the role that they play in imagining and developing new technologies. Here, the invitation was for participants not only to address the discursive level of imagination (understanding the system and its operations through a particular metaphor or figure) but also to connect it to material elements and practices as well as cultural contexts. Rooting systems in material as well as discursive contexts is crucial both for challenging the perceived disembodied nature of computing systems in general [12], as well as for tethering accountability to social, cultural, and material foundations that the systems rest on. Furthermore, as a conceptual tool actualized through inviting participants to think through the cultural heritage available to them and their communities, this workshop contributed to a more inclusive design of ML systems. Inclusive here refers to not only to identities, but also, importantly, different knowledge practices (myth, storytelling, conceptual mapping) that are not necessarily part and parcel of conventional ML systems design toolkit.

Similarly, critical fabulation and speculation also invited inclusion in terms of methods (narrative writing and analysis, narrative fabulation) and perspectives (by asking participants to fabulate from different perspectives). Furthermore, as a design approach that aims to orient towards building alliances, recuperation of silenced perspectives, interferences into dominant design narratives, and extension of capacities for accountability through design [17], critical fabulation was introduced as a method that calls to account the designers and the systems they create. Critical fabulation is a political form of materially and contextually grounded, accountable storytelling that enables to think futures, pasts, and presents together in order to both recuperate absences in dominant narratives as well as to change possible futures. In the workshops it also drew attention to how matter itself is storied and how things that are made are never just pure "tools" but are embedded in and convey stories and histories.

All these concepts provided ways to guide the design process without over-structuring it. They acted as orienting and concretizing moments by actualizing certain ideas and orienting the design towards those ideas. This, however, required a delicate balance between providing sufficient information and background material for participants to grasp the concept and guiding exercises to actualize it, and leaving enough space for them to develop their own understanding and ways of putting the concept to work in practice. For that, an experimental (in a broadest sense), open space was needed, without constraints of immediate utility or applicability. The interpretative openness was crucial for keeping with the intersectional feminist and critical anti-racist, post-colonial ethos of allowing plural interpretations, modes of doing and knowing, and avoiding foreclosure of meaning and agency. This was often at odds with the more pragmatic and application-oriented requirements that characterize the work of technology design in general, and ML systems design in particular. Nonetheless, we argue that this tension is productive in that it brings forth and encourages the questioning of disciplinary premises and highlights the tensions that are part and parcel of reflections on justice, fairness, equity, equality, oppression, and power – concepts that can be equally hard to formalize.

4.2 Re-Orienting Accountability Towards Response-Ability and Fairness Towards Justice

The workshop series proposed and enacted a way of understanding accountability in ML systems as a process rather than a fixed requirement that can be encapsulated in a single method. This is not necessarily a new idea in research on accountability of technology [23, 53, 59, 84] – the workshops, however, aimed at providing one way to *actualize* this idea. We proposed to participants to think of accountability as a process that relies on fostering the capacity to respond: response-ability [43]. This capacity depends not only on drawing official chains of accountability within organizations, but also on the deeper understanding of the histories and implications of technologies. This created palpable tensions for participants between critical theories that highlight processualism and the disciplinary conventions and pragmatic orientations of computing design that is often result-oriented. Nonetheless, such tension generated productive discussions and highlighted the idea that accountability is only possible when one is able to call one(self)

and systems to account and implicates one(self) in the histories and effects of technologies. In the context of critical feminist, anti-racist, and radical thinking, response-ability is also rooted in collective reflection and action – an aspect that was also embedded in the structure of the workshops by focusing on group work.

Collectivity is also an important dimension in re-orienting questions of fairness towards questions of justice. As has been noted in various literature on fairness and justice in ML and algorithms [37, 58, 70, 87], fairness can be a rather limiting concept, particularly as it can legitimize inequalities and has a limited scope of application [58]. The conceptual framework of these workshops encouraged participants to consider questions of exclusion, inclusion, and power, thus also steering the design process towards considerations of algorithmic justice, sensitivity to subjugated knowledges, and multiple ways of knowing. This was also done through including a broader scope of context that was taken as significant for ML systems design, as well as through prioritizing a plurality of disciplinary and other forms of knowledge (such as personal and collective experience) and design – in other words, epistemic plurality [89].

Last but not least, considerations of power differentials, including inequalities and oppressions, ran as a constant red thread throughout the workshop series. All the concepts that led the design process emerge from and are contextualized within explicitly political intersectional, feminist, anti-racist, de- and post-colonial research. This is the case both with more abstract concepts such as "diffraction", as it is entangled in questions of power and politics of knowledge and objectivity, as well as with concepts such as "critical fabulation", which was initially conceived by Saidiya Hartman in the context of historiography and critical race theory as a method of working with the archives of slavery [49, 51, 52]. Following the positioning of critical concepts, the facilitators therefore streamlined reflexive questions on power and positionality throughout the exploratory and design exercises. The focus on power and a clear invitation to align with intersectional feminist, anti-racist, postcolonial, and overall power-critical concerns is both a challenge and a necessary part of such concept-led design work towards accountability and justice.

5 CONCLUSION

In sum, this paper presented an example of working with critical intersectional concepts towards more contextualized, inclusive, and accountable ML systems design. Four concepts from intersectional feminist scholarship were selected: "situated knowledges", "figurations", "diffraction", and "critical fabulation/speculation". These were explored through a series of four workshops, actualizing the concepts through exploratory and design exercises. The concepts guided the design process, acting as orienting and concretizing moments, leading to questions of positioning, epistemic plurality, and power differentials. Apart from detailing the exercises and the workshop process, we have argued that the power of these concepts lies in working with them in an open, experimental setting and maintaining the tensions between open-endedness, interpretative flexibility, and the need for formalization. We have also argued that the critical concept-led workshops created a space where questions of fairness were re-oriented towards justice, and accountability towards response-ability. Our goal is to offer this workshop framework as one example of how critical theoretical concepts from the humanities and social sciences can be worked with in ML systems design. This is certainly not the only possible way to do this, however, we hope that it can serve as a reference point and inspiration for further research and design.

In closing, we would like to offer some questions and reflections that we find important for the fostering of fairness and accountability in ML systems. One of the questions is how much the methodology we proposed can be adapted and implemented in more mainstream computer science ethics curricula and industry or professional development settings. Since this project was experimental and small in scope, this does remain an open question yet to be answered. Our initial suggestion is that working with critical intersectional concepts is possible in mainstream and industry contexts, yet it requires not only that humanities and social sciences concepts be "attuned" to computer science, but also that computer science holds itself open to these different disciplinary ways of knowing and doing. This requires creating the space for different processes and languages and adjusting expectations of what the work process and the end results might look like.

Through discussions with workshop participants, it became clear that it was important to highlight that there is no fool-proof way to implement fairness and accountability in and through systems design. Instead, what seemed important to us, is to position accountability and fairness as inextricably dependent on the *process* of making one(self) and systems and communities accountable through deliberate research and decision-making that is attentive to both the past, the present, and the future of sociotechnical lifeworlds. This also requires deliberate willingness to invest in fairness and accountability, to become response-able to the contexts, materials, and communities that ML systems operate within and affect, particularly if we are to create not simply fair but also just technologies. Such response-ability is as much a disposition, as it is a skill, and thus not easy to formalize and encapsulate within requirements and standards. Instead, we would like to suggest that tensions between formalization and open-endedness, different ways of knowing and doing (disciplinary and otherwise) need to be held open to foster the deliberation of values and ethics – in other words, a space must be kept open for *political* ground in/of design, in the broadest sense of the word.

ACKNOWLEDGMENTS

This research was conducted during a fellowship that was supported by the German Academic Exchange Service (DAAD) stipend within the program IFI – Internationale Forschungsaufenthalte für Informatikerinnen und Informatiker (Doktoranden), 2019-2022 (57515303). We are also grateful to anonymous peer reviewers for their helpful comments and constructive feedback.

REFERENCES

[1] Phillip E. Agre. 1997. Toward a Critical Technical Practice: Lessons Learned in Trying to Reform AI. In *Social science, technical systems, and cooperative work. Beyond the great divide*, Geoffrey C. Bowker, Susan L. Star, Les Gasser and William Turner, Eds. Computers, cognition, and work. Psychology Press, New York, 131–157.

[2] Syed M. Ali. 2016. A brief introduction to decolonial computing. *XRDS* 22, 4, 16–21. DOI: https://doi.org/10.1145/2930886.

[3] Ramon Amaro. 2019. Artificial Intelligence: warped, colorful forms and their unclear geometries. In *Schemas of Uncertainty. Soothsayers and Soft AI*, Danae Io and Callum Copley, Eds. PUB/Sandberg Instituut, Amsterdam, 69–90.

[4] Gloria Anzalduá. 2012 [1987].*Borderlands/ La frontera: The new Mestiza* (4th edition). Aunt Lute Books, San Francisco.

[5] James Auger. 2013. Speculative design: crafting the speculation. *Digital Creativity* 24, 1, 11–35. DOI: https://doi.org/10.1080/14626268.2013.767276.

[6] Karen Barad. 2014. Diffracting Diffraction: Cutting Together-Apart. *Parallax* 20, 3, 168–187. DOI: https://doi.org/10.1080/13534645.2014.927623.

[7] Karen M. Barad. 2007. *Meeting the universe halfway. Quantum physics and the entanglement of matter and meaning*. Duke University Press, Durham, N.C.

[8] Shaowen Bardzell, Jeffrey Bardzell, Jodi Forlizzi, John Zimmerman, and John Antanitis. 2012. Critical design and critical theory. In *Proceedings of the Designing Interactive Systems Conference on - DIS '12*. ACM Press, New York, New York, USA, 288. DOI: https://doi.org/10.1145/2317956.2318001.

[9] Corinna Bath. 2014. Diffractive Design. In *Gender-UseIT: HCI, Usability und UX unter Gendergesichtspunkten*, Nicola Marsden, Ute Kempf, Doris Allhutter, Corinna Bath, Göde Both, Elisabeth Büllesfeld, Claude Draude, Dorothea Erharter, Leonhard Glomann, Melanie Irrgang, Veronika Kemper, Bente Knoll, Jasmin Link, Petra Lucht, Susanne Maaß, Tanja Paulitz, Bianca Prietl, Kristin Probstmeyer, Gabriele Schade, Saskia Sell, Ina Stegemöller, Jochen Süßmuth, Meinald T. Thielsch, Kamila Wajda, Elka Xharo and Christian Zagel, Eds. De Gruyter Oldenbourg, 27–36. DOI: https://doi.org/10.1515/9783110363227.27.

[10] Peter L. Berger and Thomas Luckmann. 1966. *The Social Construction of Reality*. Random House, London.

[11] Abeba Birhane. 2021. Algorithmic injustice: a relational ethics approach. *Patterns* 2, 2, 100205. DOI: https://doi.org/10.1016/j.patter.2021.100205.

[12] Jean-François Blanchette. 2011. A material history of bits. *J. Am. Soc. Inf. Sci.* 62, 6, 1042–1057. DOI: https://doi.org/10.1002/asi.21542.

[13] Tolga Bolukbasi, Kai-Wei Chang, James Y. Zou, Venkatesh Saligrama, and Adam T. Kalai. 2016. Man is to Computer Programmer as Woman is to Homemaker? Debiasing Word Embeddings. In *Advances in Neural Information Processing Systems*, 29, 1–9.

[14] Rosi Braidotti. 2001. On Conceptual Personae and Historical Narratives. A comment on History and Poststructuralist Philosophy. In *Gender, die Tücken einer Kategorie*, Claudia Honegger, Caroline Arni and Joan W. Scott, Eds. Chronos, Zürich, 65–75.

[15] Rosi Braidotti. 2011. *Nomadic subjects. Embodiment and sexual difference in contemporary feminist theory* (2nd ed.). Columbia University Press, New York, Chichester.

[16] Rosi Braidotti. 2013. *The Posthuman*. Polity, Cambridge.

[17] Rosi Braidotti. 2018. A Theoretical Framework for the Critical Posthumanities. *Theory, Culture & Society* May 2018. DOI: https://doi.org/10.1177/0263276418771486.

[18] Kirsten Bray and Christina Harrington. 2021. Speculative Blackness: Considering Afrofuturism in the Creation of Inclusive Speculative Design Probes. In *Designing Interactive Systems Conference 2021*. ACM, New York, NY, USA, 1793–1806. DOI: https://doi.org/10.1145/3461778.3462002.

[19] Loren Britton, Goda Klumbyte, and Claude Draude. 2019. Doing thinking: revisiting computing with artistic research and technofeminism. *Digital Creativity* 1, 1, 1–16. DOI: https://doi.org/10.1080/14626268.2019.1684322.

[20] Simone Browne. 2015. *Dark Matters. On the Surveillance of Blackness*. Duke University Press, Durham.

[21] Joy Buolamwini and Timnit Gebru. 2018. Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification. In *Proceedings of the 1st Conference on Fairness, Accountability and Transparency*. Proceedings of Machine Learning Research. PMLR, 77–91.

[22] Angel A. Cabrera, Will Epperson, Fred Hohman, Minsuk Kahng, Jamie Morgenstern, and Duen H. Chau. 2019. FAIRVIS: Visual Analytics for Discovering Intersectional Bias in Machine Learning. In *2019 IEEE Conference on Visual Analytics Science and Technology (VAST)*. IEEE, 46–56. DOI: https://doi.org/10.1109/VAST47406.2019.8986948.

[23] Florian Cech. 2020. Beyond Transparency. In *Companion of the 2020 ACM International Conference on Supporting Group Work*. ACM, New York, NY, USA, 11–14. DOI: https://doi.org/10.1145/3323994.3371015.

[24] Sarah Ciston. 2019. Intersectional AI Is Essential. *CITAR Journal* 11, 2, 3–8. DOI: https://doi.org/10.7559/citarj.v11i2.665.

[25] Patricia H. Collins. 2016. Black Feminist Thought as Oppositional Knowledge. *Departures in Critical Qualitative Research* 5, 3, 133–144. DOI: https://doi.org/10.1525/dcqr.2016.5.3.133.

[26] Sasha Costanza-Chock. 2020. *Design justice. Community-led practices to build the worlds we need*. Information policy. The MIT Press, Cambridge Massachusetts.

[27] Michael B. Dando, Nathan Holbert, and Isabel Correa. 2019. Remixing Wakanda: Envisioning Critical Afrofuturist Design Pedagogies. In *Proceedings of FabLearn 2019*. ACM, New York, NY, USA, 156–159. DOI: https://doi.org/10.1145/3311890.3311915.

[28] Kinjal Dave. 2019. Systemic Algorithmic Harms. Theories of "bias" alone will not enable us to engage in critiques of broader socio-technical systems. (2019). Retrieved from https://points.datasociety.net/systemic-algorithmic-harms-e00f99e72c42.

[29] Sara Diamond, Justine de Ridder, Khalid Hassan, and Bruno Emond. 2021. *Design for our times: workshop 4: speculative design, AI, machine learning and design. DOI:* https://doi.org/10.4224/40002673.

[30] Claude Draude, Goda Klumbyte, Phillip Lücking, and Pat Treusch. 2019. Situated algorithms: a sociotechnical systemic approach to bias. *OIR* 44, 2, 325–342. DOI: https://doi.org/10.1108/OIR-10-2018-0332.

[31] Sheena Erete, Aarti Israni, and Tawanna Dillahunt. 2018. An intersectional approach to designing in the margins. *Interactions* 25, 3, 66–69. DOI: https://doi.org/10.1145/3194349.

[32] James R. Foulds, Rashidul Islam, Kamrun N. Keya, and Shimei Pan. 2020. An Intersectional Definition of Fairness. In *2020 IEEE 36th International Conference on Data Engineering (ICDE)*. IEEE, 1918–1921. DOI: https://doi.org/10.1109/ICDE48307.2020.00203.

[33] Sarah Fox, Amanda Menking, Stephanie Steinhardt, Anna L. Hoffmann, and Shaowen Bardzell. 2017. Imagining Intersectional Futures: Feminist approaches in CSCW. In *Companion of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing - CSCW '17 Companion*. ACM Press, New York, New York, USA, 387–393. DOI: https://doi.org/10.1145/3022198.3022665.

[34] Maya I. Ganesh and Emanuel Moss. 2022. Resistance and refusal to algorithmic harms: Varieties of 'knowledge projects'. *Media International Australia* 183, 1, 90–106. DOI: https://doi.org/10.1177/1329878X221076288.

[35] Chu H. Goh, Narayanan Kulathuramaiyer, and Tariq Zaman. 2017. Riding Waves of Change: A Review of Personas Research Landscape Based on the Three Waves of HCI. In *Information and communication technologies for development. 14th IFIP WG 9.4 International Conference on Social Implications of Computers in Developing Countries, ICT4D 2017, Yogyakarta, Indonesia, May 22-24, 2017, proceedings / edited by Jyoti Choudrie, M. Sirajul Islam, Fathul Wahid, Julian M. Bass, Johanes Eka Priyatma*, Jyoti Choudrie, M. S. Islam, Fathul Wahid, Julian M. Bass and Johanes E. Priyatma, Eds. IFIP advances in information and communication technology, 1868-4238, 504. Springer, Cham, 605–616. DOI: https://doi.org/10.1007/978-3-319-59111-7_49.

[36] Joanne E. Gray and Alice Witt. 2021. A feminist data ethics of care framework for machine learning. The what, why, who and how. *First Monday* 26, 12 - 6 December 2021. DOI: https://doi.org/10.5210/fm.v26i12.11833.

[37] Lelia M. Hampton. 2021. Black Feminist Musings on Algorithmic Oppression. In *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*. ACM, New York, NY, USA, 1. DOI: https://doi.org/10.1145/3442188.3445929.

[38] Alex Hanna, Emily Denton, Andrew Smart, and Jamila Smith-Loud. 2020. Towards a Critical Race Methodology in Algorithmic Fairness. In *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency*. ACM, New York, NY, USA, 501–512. DOI: https://doi.org/10.1145/3351095.3372826.

[39] Donna Haraway. 1988. Situated Knowledges: The Science Question in Feminism and the Privilege of Partial Perspective. *Feminist Studies* 14, 3, 575–599. DOI: https://doi.org/10.2307/3178066.

[40] Donna Haraway. 1996. *Modest_Witness@Second_Millennium. FemaleMan©_Meets_OncoMouse™. Feminism and Technoscience*. Routledge, New York, London.

[41] Donna J. Haraway. 1991. A Cyborg Manifesto. Technology and Socialist Feminism In the Late Twentieth Century. In *Simians, Cyborgs and Women. The Reinvention of Nature*. Routledge, New York, 149–181.

[42] Donna J. Haraway. 1991. *Simians, Cyborgs and Women. The Reinvention of Nature*. Routledge, New York.

[43] Donna J. Haraway. 2016. *Staying with the trouble. Making kin in the Chthulucene*. Duke University Press, Durham.

[44] Sandra Harding. 1991. *Whose Science? Whose Knowledge? Thinking from Women's Lives*. Cornell University Press, New York.

[45] Sandra G. Harding. 1998. *Is science multicultural? Postcolonialisms, feminisms, and epistemologies* ([Nachdr.]). Race, gender, and science. Indiana Univ. Press, Bloomington, Ind.

[46] Sandra G. Harding. 2008. *Sciences from below. Feminisms, postcolonialities, and modernities*. Next wave: New directions in women's studies. Duke University Press, Durham.

[47] Moritz Hardt, Eric Price, and Nathan Srebro. 2016. *Equality of Opportunity in Supervised Learning*.

[48] Fox D. Harrell. 2010. Toward a Theory of Critical Computing: The Case of Social Identity Representation in Digital Media Applications. *CTheory* Code Drift: Essays in Critical Digital Studies, cds006.

[49] Saidiya Hartman. 2008. Venus in Two Acts. *Small Axe: A Caribbean Journal of Criticism* 12, 2, 1–14. DOI: https://doi.org/10.1215/-12-2-1.

[50] Saidiya Hartman. 2020. Wayward Lives, Beautiful Experiments: Intimate Histories of Riotous Black Girls, Troublesome Women, and Queer Radicals (First edition). W.W. Norton & Company, New York.

[51] Saidiya Hartman and Frank B. Wilderson. 2003. The Position of the Unthought. *Qui Parle* 13, 2, 183–201.

[52] Saidiya V. Hartman. 2008. *Lose your mother. A journey along the Atlantic slave route* (1. paperback edition). Farrar Straus & Giroux, New York.

[53] Jaz Hee-jeong Choi, Laura Forlano, and Denisa Kera. 2020. Situated Automation. In *Proceedings of the 16th Participatory Design Conference 2020 - Participation(s) Otherwise - Volume 2*. ACM, New York, NY, USA, 5–9. DOI: https://doi.org/10.1145/3384772.3385153.

[54] Patricia Hill Collins. 1991. *Black Feminist Thought. Knowledge, consciousness, and the politics of empowerment*. Routledge, New York.

[55] Lilly Irani, Janet Vertesi, Paul Dourish, Kavita Philip, and Rebecca E. Grinter. 2010. Postcolonial computing: A Lens on design and development. In *Proceedings of the 28th international conference on Human factors in computing systems - CHI '10*. ACM Press, New York, New York, USA, 1311–1320. DOI: https://doi.org/10.1145/1753326.1753522.

[56] Abigail Z. Jacobs and Hanna Wallach. 2021. Measurement and Fairness. In *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*. ACM, New York, NY, USA, 375–385. DOI: https://doi.org/10.1145/3442188.3445901.

[57] Christine Kaeser-Chen, Elizabeth Dubois, Friederike Schüür, and Emanuel Moss. 2020. Positionality-aware machine learning. In *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency*. ACM, New York, NY, USA, 704. DOI: https://doi.org/10.1145/3351095.3375666.

[58] Maximilian Kasy and Rediet Abebe. 2021. Fairness, Equality, and Power in Algorithmic Decision-Making. In *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*. ACM, New York, NY, USA, 576–586. DOI: https://doi.org/10.1145/3442188.3445919.

[59] Michael Katell, Meg Young, Dharma Dailey, Bernease Herman, Vivian Guetler, Aaron Tam, Corinne Bintz, Daniella Raz, and P. M. Krafft. 2020. Toward situated interventions for algorithmic equity. In *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency*. ACM, New York, NY, USA, 45–55. DOI: https://doi.org/10.1145/3351095.3372874.

[60] Yarden Katz. 2020. *Artificial whiteness. Politics and ideology in artificial intelligence*. Columbia University Press, New York.

[61] Goda Klumbyte, Claude Draude, and Loren Britton. 2020. *Re-Imagining HCI: New Materialist Philosophy and Figurations as Tool for Design*.

[62] Bogdan Kulynych, David Madras, Smitha Milli, Inioluwa D. Raji, Angela Zhou, and Richard Zemel. 2020. *Participatory Approaches to Machine Learning (Workshop)* (2020). Retrieved 01/20/2022 from https://participatoryml.github.io/.

[63] Neha Kumar and Naveena Karusala. 2019. Intersectional Computing. *Interactions* XXVI, 2, 50.

[64] Bruno Latour. 1996. On actor-network theory. A few clarifications plus more than a few complications. *Soziale Welt* 47, 369–381.

[65] Lucian Leahu. 2016. Ontological Surprises: A Relational Perspective on Machine Learning. In *Proceedings of the 2016 ACM Conference on Designing Interactive Systems - DIS '16*. ACM Press, New York, New York, USA, 182–186. DOI: https://doi.org/10.1145/2901790.2901840.

[66] Susan Leavy. 2018. Gender bias in artificial intelligence. In *Proceedings of the 1st International Workshop on Gender Equality in Software Engineering*. ACM, New York, NY, USA, 14–16. DOI: https://doi.org/10.1145/3195570.3195580.

[67] Susan Leavy, Eugenia Siapera, and Barry O'Sullivan. 2021. Ethical Data Curation for AI: An Approach based on Feminist Epistemology and Critical Theories of Race. In *Proceedings of the 2021 AAAI/ACM Conference on AI, Ethics, and Society*. ACM, New York, NY, USA, 695–703. DOI: https://doi.org/10.1145/3461702.3462598.

[68] Derek Leben. 2020. Normative Principles for Evaluating Fairness in Machine Learning. In *Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society*. ACM, New York, NY, USA, 86–92. DOI: https://doi.org/10.1145/3375627.3375808.

[69] Michelle S. A. Lee. 2019. Context-conscious fairness in using machine learning to make decisions. *AI Matters* 5, 2, 23–29. DOI: https://doi.org/10.1145/3340470.3340477.

[70] Alan Lundgard. 2020. Measuring justice in machine learning. In *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency*. ACM, New York, NY, USA, 680. DOI: https://doi.org/10.1145/3351095.3372838.

[71] Jacob Metcalf, Emanuel Moss, Elizabeth A. Watkins, Ranjit Singh, and Madeleine C. Elish. 2021. Algorithmic Impact Assessments and Accountability. In *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*. ACM, New York, NY, USA, 735–746. DOI: https://doi.org/10.1145/3442188.3445935.

[72] Mara Mills and Paula Chakravartty. 2018. Virtual Roundtable on "Decolonial Computing". *Catalyst* 4, 2, 1–4. DOI: https://doi.org/10.28968/cftt.v4i2.29588.

[73] Michael J. Muller and Allison Druin. 2003. Participatory Design: The Third Space in HCI. In *The human-computer interaction handbook. Fundamentals, evolving technologies and emerging applications*, Julie A. Jacko and Andrew Sears, Eds. Human factors and ergonomics. Erlbaum, Mahwah, NJ, 1051–1068.

[74] Chris Olah, Alexander Mordvintsev, and Ludwig Schubert. 2017. Feature Visualization. *Distill* 2, 11. DOI: https://doi.org/10.23915/distill.00007.

[75] Kavita Philip, Lilly Irani, and Paul Dourish. 2012. Postcolonial Computing. A Tactical Survey. *Science, Technology, & Human Values* 37, 1, 3–29. DOI: https://doi.org/10.1177/0162243910389594.

[76] Inioluwa D. Raji, Morgan K. Scheuerman, and Razvan Amironesei. 2021. You Can't Sit With Us: Exclusionary Pedagogy in AI Ethics Education. In *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*. ACM, New York, NY, USA, 515–525. DOI: https://doi.org/10.1145/3442188.3445914.

[77] Els Rommes. 2002. *Gender Scripts and the Internet - The Design and Use of Amsterdam's Digital City*. Twente University Press, Twente.

[78] Daniela Rosner. 2020. *Critical Fabulations. Reworking the methods and margins of design*. MIT Press, Cambridge, MA.

[79] Daniela K. Rosner. 2018. *Critical Fabulations. Reworking the Methods and Margins of Design*. MIT Press, Cambridge.

[80] Ari Schlesinger, W. K. Edwards, and Rebecca E. Grinter. 2017. Intersectional HCI. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems - CHI '17*. ACM Press, New York, New York, USA, 5412–5427. DOI: https://doi.org/10.1145/3025453.3025766.

[81] Ari Schlesinger, Kenton P. O'Hara, and Alex S. Taylor. 2018. Let's Talk About Race. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems - CHI '18*. ACM Press, New York, New York, USA, 1–14. DOI: https://doi.org/10.1145/3173574.3173889.

[82] Andrew D. Selbst, Danah Boyd, Sorelle A. Friedler, Suresh Venkatasubramanian, and Janet Vertesi. 2019. Fairness and Abstraction in Sociotechnical Systems. In *Proceedings of the Conference on Fairness, Accountability, and Transparency - FAT* '19*. ACM Press, New York, New York, USA, 59–68. DOI: https://doi.org/10.1145/3287560.3287598.

[83] Franchesca Spektor and Sarah Fox. 2020. The 'Working Body': Interrogating and Reimagining the Productivist Impulses of Transhumanism through Crip-Centered Speculative Design. *Somatechnics* 10, 3, 327–354. DOI: https://doi.org/10.3366/SOMA.2020.0326.

[84] Lucy Suchman. 2002. Located accountabilities in technology production. *Scandinavian Journal of Information Systems* 14, 2, 91–105.

[85] Lucy Suchman. 2006. *Human–Machine Reconfigurations*. Cambridge University Press, Cambridge.

[86] Ann Swindler and Jorge Arditti. 1994. The New Sociology of Knowledge. *Annual Review of Sociology* 20, 305–329.

[87] Linnet Taylor. 2017. What is data justice? The case for connecting digital rights and freedoms globally. *Big Data & Society* 4, 2, 205395171773633. DOI: https://doi.org/10.1177/2053951717736335.

[88] Sandra Wachter, Brent Mittelstadt, and Chris Russell. 2017. Counterfactual explanations without opening the black box: automated decisions and the GDPR. *Harvard Journal of Law & Technology* 3, 2, 841–887.

[89] Evelyn Wan, Aviva de Groot, Shazade Jameson, Mara Păun, Phillip Lücking, Goda Klumbyte, and Danny Lämmerhirt. 2020. Lost in translation: an interactive workshop mapping interdisciplinary translations for epistemic justice. In *FAT* '20: Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency*, 692.

[90] Jason S. Wong. 2018. Design and fiction: Imagining civic AI. *Interactions* 25, 6, 42–45. DOI: https://doi.org/10.1145/3274568.

[91] Nira Yuval-Davis. 1999. What is Transversal Politics? *soundings* 12, summer, 94–98.

[92] John Zimmerman and Jodi Forlizzi. 2014. Research Through Design in HCI. In *Ways of Knowing in HCI*, Judith S. Olson and Wendy A. Kellogg, Eds. Springer New York, New York, NY, 167–189. DOI: https://doi.org/10.1007/978-1-4939-0378-8_8.

[93] James Zou and Londa Schiebinger. 2018. AI can be sexist and racist - it's time to make it fair. *Nature* 559, 7714, 324–326. DOI: https://doi.org/10.1038/d41586-018-05707-8.

APPENDIX:

A "CRITICAL TOOLS FOR MACHINE LEARNING" WORKSHOP FRAMEWORK AND EXERCISES

A.1 WS1: Situated Knowledges/Situating

Issue: ML systems are often detached from the socio-political-cultural contexts in which they operate. This detachment, or decontextualization, happens at the levels of data collection and use but also can be reproduced during the design process. This decontextualization enables ML systems to lay claims to objectivity and impartiality. This in turn might cause issues with ascribing accountability when ML systems generate erroneous or harmful results.

Methodology/concept: Situated knowledges and situating. The concept expresses the claim that in practice "pure objectivity" gives way to partial perspectives. Such partial perspectives acknowledge that knowledge always comes from somewhere: specific context, specific bodies, specific forms and instruments that generate it. Grounding knowledge in partiality allows for more accountability since such knowledge is contextualized, i.e., situated. Situated knowledges allow for an increased response-ability: capacity to respond towards their context.

Guiding question: How can ML systems design be situated towards greater acknowledgement of context and better forms of response-ability?

Main activity: Mapping contextual positioning of oneself and constructing shared situatedness.

Additional preparation prompt: List three objects that express something about your significant connections to the world: what kind of communities are significant to you, or do you feel you belong to; what kind of disciplinary or professional background do you have; what kind of materials do you like working with or are you fond of.

A.1.1 Exploratory Exercise 1: Mapping Positionality ("Rooting"). Think about the objects that you selected that represent your relation to the world. Spend some time considering: Why did you chose these specific objects? What kind of relations do they signify for you? What kind of places, people and communities, activities do they relate to? What do they say about your background and the kind of questions and values that are important to you?

Map out your answers in the digital workspace. Your goal is to form a map that represents your rootedness and specific positionality, including:

- Places (for instance, geographic locations, but also other significant background)
- People and communities
- Professional and other significant activities
- Values or questions that are important to you

A.1.2 Exploratory Exercise 2: Mapping Perspectives ("Shifting"). Consider the following questions:

- When I am positioned on my own map, what are the phenomena, issues, situations, or concerns that I can most clearly see or attend to?
- What are the phenomena/questions/concerns that I cannot see or things that are blocked from within this map?
- Which communities/concerns/environments do I feel particularly responsible to? How do the professional or disciplinary tools that I have at my disposal help me or hinder me?
- If I am to position myself slightly differently within this map, how would this list of concerns/questions change?

Spend some time thinking about this and noting down your answers as an annotation to the map that you drew before.

A.1.3 Design Exercise: Situating Towards Collective Systems. See if you can bring your maps together and construct a shared common ground within your group. Draw a map of your own group rootedness. To do that, consider the following questions:

- What are the communities, disciplines, positionalities, or orientations that you share?
- What are the significant differences that allow for diverse perspectives to be kept in mind?
- What positions might be missing?
- Can you identify concerns or problems that are significant to all of you in some way?

Work collectively to construct a shared map. Your goal is to indicate the following on this map:

- Shared contexts and concerns or significant communities
- Shared values and perspectives
- Significant differences

A.2 WS2: Figurations/Figuring

Issue: ML systems are perceived as ahistorical, disembodied, neutral tools that are devoid of power relations. Furthermore, being computational, these systems are also perceived as somewhat immaterial, operating in the abstract "digital space". The more material aspects of systems come in through user modelling. Nonetheless, here the explicit attention is usually given to the interaction between the user and the technical system where both are perceived as distinct entities.

Methodology/concept: Figurations a.k.a. conceptual personae. Figurations are mappings of situated, i.e., embedded and embodied, social positions – in other words, material-discursive entities that account for particular historical, political, and material locations. They stitch together meanings and practices. Technologies are materialized figurations that bring together both actual physical technologies and clusters of meaning (narratives, discourses, imaginaries) surrounding them, which together form more or less stable assemblages or configurations.

Guiding question: How can ML systems design be geared towards recognizing and acting from the embodied, embedded and power-laden conditions of its position as well as the effects of ML systems?

Main activity: Building a figuration of ML system.

A.2.1 Exploratory Exercise 1: Figuring Materials. Individually, explore the material that you indicated in your preparatory prompt from the previous day. Try to think about this material as a figuration. To help you do that, consider the following questions: What is it made of? Who made it? What is its history? How does it feel to the touch/taste/smell? What is the context of its use? What kind of emotions does it invoke? What are the stories related to it? What can it be made into? What can it know? Does it have memory? Does it have agency and if so, what kind? What is its body like? If it were animate, how would it see the world? What kind of concerns would it have or point to? What are the ways in which it relates to hierarchies of power? Make notes of your answers in the workspace.

Discuss together: What new perspectives, challenges or ideas does an understanding of this material as a figuration bring you? Is there something that you learned by thinking about it as a figuration?

A.2.2 Exploratory Exercise 2: Figuring Stories. For this exercise, try to weave a figuration as a structuring metaphor or conceptual imaginary that is significant for your shared concerns and/or communities.

First, consider one of the perspectives, communities, or issues that you identified as shared in your group in the Situated Knowledges collective mapping exercise. Consider the following:

- Who are the people playing a role in the definition of this issue or community?
- Are there specific imaginaries or personified figures of speech that appear often in talking about this concern or community?
- What concepts, stories and narratives are associated with this concern or community?
- What mythological, cultural, literary, or other figures are significant for this concern/issue or community?

Note the answers to these questions in the digital workspace: draw, use post-its, images – whatever media you can think of.

Once you have this rough map, try to see if there is a figuration – a structuring metaphor or a character – that emerges from these mappings as a significant concept to think with or as a figuration that embodies some of these items that you mapped out. Discuss and note down the following:

- What kind of stories does this figuration tell about your significant concern or community?
- How does this figuration illuminate, or perhaps even structure, the issue that the community is facing?

A.2.3 Design Exercise 1: Figuring Systems. The goal of this exercise it to work on the ML scenario that you selected and apply the concept of figuration to its design. Spend some time to think as a group about this ML system as a figuration.

First, consider its material base, the "stuff" it would be made of.

- What kind of immaterial and material elements would come together to form this system?
- What would its function be?
- Thinking about the system itself, what would it be able to teach you?
- What is its immediate environment within which it would make a difference?

Use sticky notes, images, digital pens etc. to note this down in the workspace.

Then, consider broader context of this system – the broader "stories" or discursive aspects:

- What historical contexts are important for the workings of this ML system?
- How does it relate to power distributions in society?
- What kind of stories would it tell?
- What kind of impact would it have?
- What would be its name?

See if you can weave a figure visually, by drawing, using images, and otherwise (in however abstract or concrete way) figuring it on the board. Pay attention to its material, discursive elements, and the relations to its context.

A.2.4 Design Exercise 2: Concretizing ML Systems. Based on the work that you did for Situated Knowledges day and Figuring Systems exercise, draw up a rough sketch – a skeleton – of how your envisioned ML system would function and what kind of design steps would be needed for its development. Consider this exercise to be a drafting of a rough outline and a design plan of your system.

Specifically, define the following:

- The goal or task of your system
- Who would be involved in its design process and what steps would the design process entail
- How would this system work
- A slogan or tag line that describes what the system is trying to do, and its name

A.3 WS3: Diffraction

Issue: ML systems design often lacks interdisciplinarity and an in-depth understanding of the effects of its specific perspective. This prevents ML systems from attracting a wider range of expertise (or knowledge practices) and from being seen and understood in broader sociotechnical terms.

Methodology/concept: Diffracting the apparatus and tracing patterns of relations. Diffraction suggests that the apparatus of measurement and the tools of knowledge production themselves play one of the constitutive roles in generating knowledge. Diffraction generates patterns of interference, i.e., patterns of difference and relation, entangling processes of meaning-making and technology.

Guiding question: What are the effects of ML systems as complex apparati of knowledge production?

Main activity: Mapping ML system as a diffractive apparatus.

A.3.1 Design Exercise 1: Clarifying ML System and Its Structure. For this part, continue working on concretizing the idea of your ML system and clarifying its parts. The goal is to start building a kind of info-sheet or portfolio and operational diagram of your ML system, or clarify it (if you have already started building it).

Define the following as clearly as possible by writing down:

1. The task of your ML system
2. Optimization criteria (what counts as good performance of your system, what are you trying to achieve)
3. The data that you would ideally use for this ML system and where would it come from
4. Operational logic of your system (what is the input, output, how would the system ideally work)
5. If you can, consider what kind of model would your system require (a model for classification or regression, for example, or perhaps you can anticipate that it would have to be a specific type of model, e.g., neural network)

A.3.2 Design and Exploratory Exercise Combined: Diffracting ML. Think about how the example of light from quantum physics might help us understand ML systems. When we say that something (e.g., a light diffraction experimental set up with two slits) acts as a diffraction apparatus, what we learned is that the apparatus itself shapes the situation and the outcome. It shapes the problem formulation and the solutions to the problem. That is, the apparatus actively intervenes in the context that it is used in. So, to think diffractively with ML is to begin with an understanding that an ML system (and everything that comes with it) from the very start plays a role in defining the situation and the outcome. This also

provides us with a really exciting possibility: what if we changed the diffraction apparatus – could we create the conditions for a different situation and outcome?

The goal of this task is to think of your ML system as a diffraction apparatus and see what can be changed by toggling with its parts. To do that, we will first identify the elements and relations that this apparatus entails and then trace the effects of this apparatus.

An apparatus has both material and immaterial parts, both technical and non-technical elements.

Step 1: To analyze an apparatus, start by identifying these elements:

- Societal/contextual elements
- Technical/disciplinary elements
- Discursive/ value elements
- Operational elements/logic

Step 2: Identify the relations between these elements:

- The relations where one element CONSTRUCTS another
- The relations where one element DISRUPTS another
- The relations where one element INTERFERES WITH another

Step 3: Trace the effects of your ML system

- Based on the elements and relations that you identified, trace what *effects* your ML system would possibly have. For example, would it introduce new values? Change existing values? Intervene into the way we understand certain phenomena? Introduce new relations between communities or phenomena?
- Discuss how you might want to change the characteristics of your system that you drafted during the first part of the day (the operational diagram of your system) based on your understanding of these effects.

A.4 WS4: Critical Fabulation and Speculation

Issue: ML systems are thick material-discursive knots, however in design they are regarded as discourse- and story-free objects/systems. Furthermore, biases emerging in ML systems often signal not only negative impacts but exclusions of specific voices, perspectives, and histories.

Methodology/concept: Critical fabulation/speculation. Critical fabulation entails speculative thought and imagination but it positions those activities against the absent imaginaries and loci of power hierarchies. In particular, critical fabulations are methodological tools to address missing imaginaries and missing perspectives. They provide an alternative to "white prototypicality" and the "I-methodology" that is prevalent in technology design.

Guiding question: How can the specific grounded perspectives and forms of collective and personal accountability be addressed in ML systems design? How can these perspectives be designed with more inclusion?

Main activity: Fabulating and prototyping speculative ML systems.

A.4.1 Design and Exploratory Exercise 1: Writing Design Narratives. For this task, the goal is to write your design narrative – a short (~200~300 words) fabulated story on what your ML system is and what it does, how it would be used and by whom.

Make sure that your narrative in some way touches the following questions:

- Who are the anticipated users and who are the other stakeholders?
- How is the system to be used?
- What role does your figuration (from WS2) play?
- What kind of difference does this system make for its users?

Write this story down in the workspace. Try to write collectively, i.e., try to make sure that you all contribute.

A.4.2 Exploratory Exercise 2: Analyzing Narratives. The narratives or fabulations that we create reflect specific situated positions. All narratives are written from somewhere and for someone and in that sense they can reveal both crucial future workings of technology as well as critical missing points.

Look at your narrative again and analyze it more closely:

- What or who are the main protagonists of this narrative? And, by extension, what or who are missing?
- What are the limitations of the figuration and imaginaries that you used in this narrative?
- Whose perspective is this narrative written from? Who does this perspective include or exclude or perhaps completely ignore?
- Does this narrative show what is the broader significance of this technology – i.e., why does it matter? If yes, who is positioned as benefitting from this technology and who might not? If no, reflect on that.
- What are your own stakes in this design – i.e., why are you designing this technology? How does it touch your personal life?

Make notes of your discussions on these questions. Don't be afraid to be critical! Analyzing your own narratives (in the context of this exercise, and more generally) is a valuable tool for understanding own positionality and the positionality of your design projects.

A.4.3 Design Exercise 2: Fabulating with Critical Perspectives. Critical fabulation as a method has a strong political grounding because it asks questions about omissions that matter, accountability for the stories we tell, and construction of the futures that we want to see. Based on the narrative analysis that you did in the previous step, fabulate a narrative with these critical questions and insights in mind. Consider the following:

- How can this system center less heard voices or address their absence?
- What alternative histories or practices can inform this technology and/or its use?
- What forms of collective and (your as designer's) personal responsibility can be introduced or demonstrated?

For this fabulation, try adopting a different perspective than you have in the first narrative (e.g., you could write it from the perspective yourself or your team, or form the perspective of the ML system itself, or perhaps from a member of a population that would be indirectly affected by your ML system). Don't edit the first story but rather write a new one.