Diversity, Technology, Power:
An American Black Feminist Approach to Studying and Designing Diversity in Relation to Contemporary Technologies

D i s s e r t a t i o n

zur Erlangung des akademischen Grades
Doktor der Philosophie in der Philosophischen Fakultät
der Eberhard Karls Universität Tübingen

vorgelegt von Laura Schelenz
aus Köln, Deutschland

2025

Diversity, Technology, Power:
An American Black Feminist Approach to Studying and Designing Diversity in Relation to Contemporary Technologies

D i s s e r t a t i o n

zur Erlangung des akademischen Grades
Doktor der Philosophie in der Philosophischen Fakultät
der Eberhard Karls Universität Tübingen

vorgelegt von Laura Schelenz
aus Köln, Deutschland

2025

**Gedruckt mit Genehmigung der Philosophischen Fakultät
der Eberhard Karls Universität Tübingen**

Dekanin:	**Prof. Dr. Angelika Zirker**
Hauptberichterstatter:in:	**Prof. Dr. Astrid Franke**
Mitberichterstatter:in:	**Prof. Dr. Daniel Gatica-Perez**
Tag der mündlichen Prüfung:	**06.12.2023**

Universitätsbibliothek Tübingen, TOBIAS-lib

Acknowledgements

The deepest appreciation goes to the advisors to this project. First, Prof. Astrid Franke who has been a constructive force from the very beginning until the end – reliable, prudent, inspiring, and so knowledgeable, which is reflected in the good parts of this dissertation. Second, Prof. Daniel Gatica-Perez whose vast experience and expertise in computer science and social computing has made this interdisciplinary project possible. Heartfelt thanks also to Prof. Regina Ammicht Quinn and PD Dr. Jessica Heesen for an open ear when needed and for giving me the time and freedom to follow my research where it was leading me.

Many thanks to important sparring partners who were always available to discuss concepts, methods, and angles: Dr. Avi Segal, Prof. Kobi Gal, and Dr. Marcel Vondermaßen. Thanks to the participants of the colloquium in American Studies with Prof. Astrid Franke at the University of Tübingen, to colleagues at the International Center for Ethics in the Sciences and Humanities at the University of Tübingen, to peers across the pond at the University of North Carolina at Chapel Hill, to the WeNet consortium, to the Digilog consortium, and to the many brief but memorable contacts during conferences and workshops.

Abstract

This dissertation employs a critical perspective on the meaning of diversity in the design of contemporary technologies, which may be based on Artificial Intelligence (AI). Different concepts of diversity are scrutinized from a perspective that is critical of existing power relations and has a concern for social justice. The combining framework that ties individual chapters of the dissertation together is American Black feminist theory. This tradition of thought, which is built on the experiences of Black women in the USA, can render visible structural inequalities in society and technology. The theory's application to the study and design of diversity and technology in both US-American and European contexts helps discuss the social benefits and risks of contemporary 'diversity-aware' technologies and their ecosystems. By employing a critical perspective on power systems/relations inspired by American Black feminism, the dissertation picks up on a trend in computer science, where other theories of situated knowledge are increasingly applied to study and design new technologies.

The dissertation applies Black feminism to the study of diversity concepts that are embedded into the design of contemporary (AI-based) technologies, for example as a way to personalize content in computer systems that provide recommendations of content to users. Conceptualizing diversity for the design of technology is another trend in computer science. Here, diversity categories such as demographic and cultural notions of user difference are used to design for user satisfaction, fairness, or social inclusion. The dissertation scrutinizes so-called 'diversity-aware technology.' The Black feminist framework helps uncover shortcomings in the operationalization of diversity in mainstream computer science. The theory can also inspire a conceptualization of diversity that rejects the classifications of users, the use of historical data, and instead considers the lived reality of marginalized technology stakeholders. In the dissertation, mainstream notions of diversity are rethought, and diversity is (re)aligned with concerns for redistribution, civil rights, and social justice.

The dissertation expands an American Black feminist perspective to the European context. In an empirical study with Afghan refugee women in Germany, Black feminist theory helps understand the structural marginalization of Afghan women in their interaction with technology. The study finds limitations in how Afghan refugee women access and use technology and discusses the short-sighted path of a purely technical intervention for the social inclusion of refugees. Furthermore, in a series of expert interviews with designers based in Europe, the dissertation reconstructs critical perspectives on diversity and design with regard to social spaces – both virtual and physical. It compares American and European perspectives

on power and participation in technology and urban design, highlighting cultural, social, and political specificities in engaging diversity categories and stakeholders.

The dissertation is guided by the following research questions, which become relevant in the research contributions of each chapter: What are different understandings of diversity in computer science and technology development? What are the social benefits and risks of working with diversity in technology design? How can diversity be conceptualized from a perspective that is critical of existing power systems/relation? What are the benefits and limitations of an American Black feminist approach to studying and designing diversity in relation to technology? The dissertation makes scientific contributions in two major ways. First, it analyses the definitions, discourses, methods, and implications of diversity in contemporary computer science and technology development with the help of American Black feminist theory. Second, the dissertation reflects on the theory of Black feminism itself as a tool to study diversity definitions, discourses, methods, and implications of diversity in computing.

The dissertation sits at the intersection of American studies, technology studies, computer science, and gender and feminist studies. It is relevant to computer scientists and designers of computer systems who seek guidance on the meaning of diversity and how (not) to leverage diversity in their designs. It is further relevant to scholars in the humanities, social sciences, and science and technology studies for its theoretical engagement with diversity, a critique of power, and its reflection on the implications of embedding diversity in technology. Finally, the dissertation offers inspiration to practitioners of all kinds, including activists, who aspire to design more just socio-technical futures.

STRUCTURE OF THE DISSERTATION

Chapter 1: Introduction

Four years ago, I was sketching out an idea for a 'truly' diversity-aware application. The technology would leverage the diversity of users to provide educational opportunities based on the collection of data from users' socio-economic-political contexts. For example, the app would allow its users to document their experiences of advantage and disadvantage in society in the form of a diary or memoir (e.g. privilege of college education, stress about student loans, lack of low-cost qualification, language barriers, single parent). The app would then – due to its 'awareness' of users' diversity – provide those in marginalized socio-economic-political contexts with free tutoring and support to close the inequality gap. In my unspent mind, the app would be a proof of concept for designing for social justice. My key to unlocking the app's social impact was supposed to be American Black feminist theory – a critical social theory that is used to raise consciousness about societal power inequities (Collins, 2000a, 2019). After four years of working with American Black feminist theory[1] in conjunction with technology,[2] it is utterly clear that my sketched idea was flawed.

Diversity is an ambiguous concept. It can refer to the description of difference or to a normative claim towards the inclusion of different groups in society. Diversity discourses have dominated the business world since the 1990s due to increased globalization and exchange of goods and services across cultures (Bendl et al., 2016; Trompenaars & Hampden-Turner, 2012). European and US-American liberal perspectives on multiculturalism have developed with increased immigration and cultural integration since the 1950s (Chin, 2017; Dhamoon, 2010). After the Civil Rights Movement of the 1960s, African Americans raised awareness that racism continues despite the granting of formal rights and the integration of Black and White communities (Peller, 1995, 133f). Critical perspectives on diversity challenge the inclusion of minorities into mainstream institutions without actual changes to the power hierarchies in organizations (Crawley, 2006; Jack, 2016) and society at large (Berrey, 2015). Within these discourses, very different diversity notions lie at the bottom of argumentation. In fact, it has been argued that diversity is popular because it is open to interpretation (Vertovec, 2012, 2015).

[1] In this dissertation, the words black, white, and brown are capitalized when they refer to a social group or in conjunction with historical, cultural, ideological, and political concepts. The capitalization signifies the social constructedness of racial markers. It is to show that 'black,' 'white,' and 'brown' are not merely colors that may describe the appearance of skin but that Black, White, and Brown people are affected by racialization. White is racialized as a norm, whereas the racialization of Black and Brown people follows the logic of 'othering,' a practice creating boundaries between an 'us' and a 'them;' for more explanation on the term see Yuval-Davis (2011, p. 47).

[2] Technology is understood in a broad sense in this dissertation, encompassing any computerized system from large social media platforms, Artificial Intelligence, personalized recommendation systems, machine learning-based transformers like ChatGPT, to electronic datasets, individual algorithms, and small-scale applications.

Diversity discourses around culture, integration, and civil rights became relevant in the technology industry as well. Technology adoption from users around the world increased attention to the role of culture in technology design (Barber & Badre, 2006; Capurro, 2019). An influx of refugees and migrants in European and US-American societies in recent years prompted designers to use technology for the social inclusion of newcomers (AbuJarour et al., 2019; Fisher et al., 2016). With civil rights demands expanding to the technology sphere, researchers increasingly call out bias and discrimination in design, and simultaneously question the structures that enable socio-technical inequity (Buolamwini, 2019b; Buolamwini & Gebru, 2018; Eubanks, 2017). Diversity moved into mainstream discussions of technology design and development with diversity's meaning remaining as ambiguous as in broader debates.

Diversity language is employed particularly in organizational efforts and professional conduct in computing. The Association of Computing Machinery's (ACM) "Commitment to Diversity, Equity, and Inclusion in Computing" argues for the importance of diversity in the computing field with diverse teams not only generating more innovations but "lead[ing] the way in creating the change we would all like to see in society."[3] The US-based ACM is the world's largest association for education and research in computing. The organization has been criticized in the past for its lack of inclusivity of minorities (Erete et al., 2023b, 5; 10). Approaches to the recruitment of Black women have also been revealed as hypocritical in light of the ongoing discrimination of Black women in the industry and through technology (Erete et al., 2020; Noble, 2018, 64f). Plus, the burden to reduce a technology's discriminatory potential is often unfairly placed on the minority.

Although diversity work for inclusive teams is an important topic for critical inquiry, this dissertation deals with a less prominent focus: the way diversity is defined, operationalized, and embedded into so-called diversity-aware technology or used as a strategy to achieve algorithmic fairness and the social inclusion of diverse technology stakeholders.[4] Here, the ACM also has an idea of diversity: they are defining "user characteristics" in their ontology for state of the art topics in computing: "race and ethnicity, religious orientation, gender (men and women), sexual orientation, people with disabilities, geographic characteristics, cultural characteristics, and age (children, seniors, adolescents)."[5] The development of diversity-aware technology relies on these and other concepts of difference between technology stakeholders or items (content, products) that are subject to algorithmic mediation (Busso et al., 2023;

[3] https://www.acm.org/diversity-inclusion/about
[4] By technology stakeholder, I refer to users *and* those affected by a technology who are not directly using. For instance, walking by a security camera enabled with facial recognition technology does not make one a user, but one is directly impacted by the system because data is collected, and people may be subject to profiling.
[5] https://dl.acm.org/ccs

Dingemanse & Liesenfeld, 2022; Schelenz et al., 2021; Shen, 2020; Tintarev, 2017). Diversity-aware technology refers to a computer system that utilizes a) a definition of diversity to account for differences between its stakeholders and b) a normative idea related to diversity such as tolerance, inclusion, or justice. There is an implicit notion that considering diversity in the design of computer systems is beneficial (see for example Himmelsbach et al., 2019) with a potential bias that such technology must automatically be fairer, more inclusive, and more just.

After attempting to employ diversity in the design of an application myself, I have learned the following. Whether a diversity-aware technology is socially beneficial depends on how diversity is considered in its design process. Hidden in descriptive diversity concepts (such as the user characteristics by the ACM) are implicit ideas about biological or cultural differences between human beings as members of social groups. Diversity concepts then carry values and embody long-standing tensions over who belongs, who is represented, and who has access to resources. If diversity is embedded into technology, e.g. as operationalized categories in a dataset to train an algorithm, the technology becomes hyper-value-laden and, depending on its scale, may even structure social relations. Drawing on flawed, oversimplified, and historically problematic diversity concepts may compromise the goal of inclusion and justice. But even careful attention to the socio-economic-political contexts of users can increase power inequities if these contexts are documented in datasets. Working with diversity data concentrates highly sensitive information about technology stakeholders (and thus power) in the hands of technology owners. My own 'truly' diversity-aware application was abandoned as a result of these considerations.

This is not to say that working with diversity in technology design and development is wrong or that potential negative consequences always exceed benefits. Rather, diversity-aware technology design requires caution and reflection. The dissertation shows that careful attention to historical legacies of diversity concepts such as race and gender as well as broader societal structures can help envision designs for more just socio-technical futures. A perspective that enables this critical thinking can be found in American Black feminist theory. US-centric Black feminism is a critical social theory and movement that combines scholarship and activism to understand and challenge unequal power structures in society. American Black feminism guides the research in this dissertation, prompting me to study diversity and technology from a perspective that is critical of existing power systems/relations. Such a 'critique of power' allows for the review of diversity definitions and discourses around fairness or social inclusion in computer science and technology development by taking into account power relations in society and the technology industry. The theoretical and methodological approach of the dissertation is

described further below in this chapter. However, it should be pointed out up front that working with Black feminism means acknowledging that technology is not neutral. Technology interacts with society, for example, by reflecting the values of its designers and broader design environments.[6] Technical artefacts are always political as they reflect decisions of designers about the affordances of the design (Benjamin, 2019b; Costanza-Chock, 2020). With this framework of socio-technicality in mind, the dissertation explores four related research questions:

What are different understandings of diversity in computer science and technology development?[7] What are the social benefits and risks of working with diversity in technology design? How can diversity be conceptualized from a perspective that is critical of existing power systems/relation? What are the benefits and limitations of an American Black feminist approach to studying and designing diversity in relation to technology?

The first two questions focus on diversity concepts. By social benefits and risks, I refer to the implications of leveraging diversity concepts for social groups and society as a whole. Social benefits and risks relate to how diversity concepts in computer science and technology development influence broader debates around diversity and reinforce or transform existing diversity narratives. Social benefits and risks also relate to how social groups are treated by algorithm-assisted institutions or programs, and how they can access resources following the inscription of diversity concepts into the technology. The latter two research questions relate to the normative framework underlying the dissertation. The dissertation then makes two major contributions. The first is an analysis of the definitions, discourses, methods, and implications of diversity in contemporary computer science and technology design, with the help of Black feminist theory. Second, the dissertation provides a critical assessment of the theory of Black feminism itself as a tool to critically study and design diversity in relation to technology.

The dissertation is divided into three parts. Part I sheds light on the potential of Black feminist theory in studying Artificial Intelligence (AI)-based technology. AI has become an increasingly relevant area of innovation in computer science and, due to its potential for prediction and automation, has other ethical implications for technology stakeholders than

[6] The socio-technical character of computer systems is also discussed in science and technology studies or the philosophy of technology, e.g., under the terms 'algorithmic assemblage': Matzner (2019); Seyfert and Roberge (2016), 'value-sensitive design': Friedman and Hendry (2019), and 'hybrid human-AI': Peeters et al. (2021).

[7] Computer science and technology development encompasses a variety of sub-fields and practices. In the dissertation, I focus on AI-based technology such as facial recognition systems and recommendation or personalization technology but also include research from human-computer interaction (particularly research in the ACM CHI and CSCW conferences). There is also a focus on smaller-scale platforms and applications, e.g. with the WeNet platform of the WeNet project (2023), and a few examples from robotics.

'traditional' software. In part II, I employ a Black feminist lens to review contemporary diversity concepts as they are embedded in technology. This part further shows how diversity can be conceptualized from a perspective that is critical of existing power systems/relations to provide a framework for the design of critical diversity-aware technology. Part III moves away from diversity categories in technology design to explore the broader societal structures that determine the accessibility of design features for marginalized groups such as Afghan refugee women in Germany. The third part also looks at design theory and methods, and how they enable a critical diversity practice in Europe and the USA.

The dissertation is particularly relevant to computer scientists and designers of computer systems, as well as scholars in the humanities, social sciences, and science and technology studies. With diversity being such an ambiguous concept, designers of computer systems might welcome guidance on the social and cultural uses of diversity definitions. This guidance can help designers who are interested in designing for diversity, inclusion, and equity to avoid some of the mistakes that can compromise their goals. The dissertation provides examples of how to engage diversity and how not to approach diversity. At times, the dissertation provides tangible recommendations, e.g., in chapter 5. Scholars in the humanities, social sciences, and science and technology studies will find this dissertation relevant for its conceptual reflection on the meaning of diversity. Furthermore, the dissertation's insights into contemporary mainstream uses of diversity concepts in technology design are helpful to consider the long-term implications of diversity-aware technology. According to Noble (2018, p. 1), "we are only beginning to understand the long-term consequences of [AI-based] decision-making tools in both masking and deepening social inequality." The same is true for technology that leverages diversity, and science and technology studies might want to study the long-term implications of this new trend.

Diversity and Technology

The intriguing fascination with diversity and technology is that age-old questions get wrapped up in discussions about the future of humankind.[8] (AI-based) technology is a site of struggle over who has access, who is represented, and who is served by the technologies' affordances and their larger enabling ecosystems. (AI-based) technology is also a site of tension between

[8] One may argue that there are important questions to be asked about the future of planet in addition to the future of humankind. This dissertation has taken an anthropocentric view, which means that it is concerned with the effects of diversity-aware technology on humans. In future research, it is important to consider how environmental concerns can be integrated into a research program on diversity-aware technology.

opening up and closing down, crossing borders and marking territory, embracing and containing diversity. Industry, government, and educational institutions aim at including more people in a global digital society, albeit for different reasons. Industrial processes are increasingly automated to optimize work flows and increase profitability of companies. Government services are offered online, and schools and universities rely on digital means for teaching and knowledge production – a development spurred by experiences during the Covid-19 pandemic (2020-2023) that required temporary suspension of in-person services. While more people are expected to participate in the digital society, power over essential services is concentrated in the hands of a few technology companies (Borowska, 2020). There is also an aspiration of technology developers to transcend geographic boundaries and bring the world closer together, but Western economic domination makes cooperation on an equal footing impossible. This finds its expression in the dominance of the English language in technology design (S. Leavy, 2018, p. 3; Warschauer, 2004, p. 203) and the large-scale extraction of data (Coleman, 2019) and 'cheap' labor (Gray & Suri, 2019) from the Global South.

Motivated by these tensions at the intersection of diversity and technology, I start my inquiry by analyzing different interpretations of diversity in computing. Diversity is debated in computer science and technology development in the following ways: as definitions of user or item difference embedded in the design of a technology; as a strategy of diversification that helps achieve commercial goals (such as to increase the adoption of a technology); as a strategy to increase the fairness of algorithms in recommendations and personalization based on machine learning techniques; as a strategy of social inclusion that aims at supporting marginalized groups in fully participating in society; as a normative standpoint and methodological approach expressing a value for pluralism, democracy, anti-racism, and justice.

While a lot of attention has been paid to diversity in the people working in technology development (Chi et al., 2021; Dunbar-Hester, 2020; Twine, 2018), little attention has been paid to actual diversity concepts, i.e. definitions of difference, that are embedded in the design of a computer system. These definitions are descriptive in the sense that they seek to describe the difference of something. In a subfield of computer science referred to as 'recommendation and personalization systems,' there is a strong focus on item diversity. The diversity of items describes the difference between products or persons recommended to a user in an online platform like Amazon, Netflix, Facebook or LinkedIn. Item diversity implies the internal diversification of a list of products or people that are recommended via an online platform to a user.

For example, when a platform such as Netflix recommends movies, Netflix' algorithmic system is expected to recommend a diverse set of movies. On the one hand, this is a strategy to increase user satisfaction. According to Castells et al. (2015, p. 883), diversity in recommendations can "optimize the chances that at least some item pleases the user." Here, diversity is tied to a commercial interest. The more satisfied the user is with a recommendation system, the longer they will engage in the system. This generates more profit for the company owning the system. On the other hand, the system is expected to recommend movies from various directors and featuring various actors (especially in a niche-genre) to increase the fairness of the recommendations (Burke, 2017; Ekstrand et al., 2019).

Fairness in recommendations, tied to the broader research area of fairness in machine learning (Oneto & Chiappa, 2020), seeks to counteract algorithmic bias. Fairness relates to diversity in the sense that a diversification of recommended items is considered key to the equal opportunity of producers of items to be recognized by the system. Diversification seeks to make underrepresented producers of items (e.g. Black filmmakers) visible and have their products promoted by the algorithm. Race and gender (e.g. of filmmakers or actors) are used to balance the algorithmic recommendations towards achieving fairness (Ekstrand et al., 2018; Hanna et al., 2020). Here, fairness is conceptualized as equal opportunity and implies an imperative of non-discrimination. The idea of equal opportunity has been criticized for obscuring historical inequalities and the fact that members of social groups have different starting points in society (S. Ahmed, 2012, p. 8; Young, 1990, p. 12; Yuval-Davis, 2011, p. 77). While fairness may promote equality, it may not promote justice (also see Hoffmann, 2019). This will be discussed in more detail in chapter 5.

In a subfield of computer science called 'human-computer interaction' (HCI), diversity relates to the difference between users. Diversity is operationalized as the culture, skills, behavior, cognitive ability or cognitive styles, bodies, and demographics (including gender, race, education) of users engaged in a system. The operationalization of diversity categories is expected to help optimize a system for different groups of users. With increased globalization, the reach of systems has expanded and users from various backgrounds engage in the same system (Allard & Blanchard, 2011). Also due to increased sensitivity to a lack of inclusion in the industry itself, there is more interest in diversity in HCI (Himmelsbach et al., 2019).

The way that diversity concepts are defined in HCI often follows a binary logic and tends to present human differences as natural and self-evident. One example is culture. Culture

is a very difficult concept to break down into computable categories.[9] In a study on the interaction of users with picture passwords, culture has been conceptualized in a binary, oversimplified and essentialist fashion to encompass Eastern and Western orientations (Constantinides et al., 2020). In the case of the "diversity-aware" WeNet platform, culture was broken down into social practices, which describe routine behaviors and skills of users (Schelenz et al., 2021).[10] The operationalization of social practices lacks attention to stereotypes such as racialized and gendered practices (e.g., the idea that women (should) engage in care work). Additionally, people who engage in 'outlier' practices may fall through the cracks. The dissertation offers strategies to mitigate negative side effects of working with diversity concepts in computer science.

Another strategy of diversity concerns the social inclusion of marginalized communities in (the digital) society. Social inclusion is the possibility for "individuals, families, and communities" to "fully participate in society and control their own destinies" (Warschauer, 2004, p. 8). Technology developers consider the diversity of needs and assets of marginalized groups to design systems that can support their integration into society (Wong-Villacres et al., 2020; Wong-Villacres et al., 2018).[11] Social inclusion relates to diversity in the sense that marginalized groups, for example migrants and refugees, face challenges different from the majority society (e.g. language barriers). Technology is expected to assist in minimizing these challenges, e.g. by connecting refugees to volunteers and language classes (AbuJarour et al., 2019; Brown & Grinter, 2016; Schreieck, Zitzelsberger, et al., 2017; Weibert et al., 2019).

Diversity also plays an important role in the critical design community. Critical design is an area of research and practice where mainstream design practices are questioned and alternative design theories and methods are explored (J. Bardzell & Bardzell, 2013). Diversity from a critical design perspective is understood as a wanted state of broad participation of stakeholders in technology development with an emphasis on democratic values. Methods such as participatory design and methods that center the needs and assets of marginalized groups are expected to produce designs that work for 'everyone' (S. Bardzell, 2018; Harrington et al., 2019). Here, the diversity of stakeholders' experiences informs design processes (Almohamed et al., 2020; Erete et al., 2023b; Fisher et al., 2016; Tran O'Leary et al., 2019). Finally, diversity

[9] Culture is also a very difficult concept to grasp for the humanities, and culture ethicists speak of a plurality of cultures, not as different 'national' cultures but rather cultural and social practices as well as countercultures and subcultures, see Ammicht Quinn (2006); Yuval-Davis (2011, p. 41).
[10] I was a researcher in the WeNet project from 2019 to 2023. I have been responsible for the ethical guidance of the project, including a critical review of the notion of diversity in WeNet, see Schelenz et al. (2019).
[11] Designs for social inclusion are, for example, developed in HCI, service systems design, the field of 'information and communication technologies for development (ICT4D)' as well as 'AI for social good.'

in critical design relates to the situatedness of designers and researchers themselves, and how their positionality in society influences their design work (Liang et al., 2021). Chapter 8 examines critical design approaches in the USA and Europe in a comparative fashion.

Diversity-aware Technology and its Benefits

Technology that takes into account diversity as a concept in its design, as a strategy, or as a normative reference point can be referred to as "diversity-aware technology." The term has been used in the literature (Busso et al., 2023; Dingemanse & Liesenfeld, 2022; Schelenz et al., 2021; Tintarev, 2017) although no shared definition exists. Diversity-aware technology relies on a) the operationalization of diversity to make the system 'aware' of some elements of diversity, and b) a normative claim or goal related to diversity such as fairness, social inclusion, or justice (see figure 1). The terms awareness and diversity-aware should not open the door to debates over machine consciousness, i.e. whether embodied technologies have a capacity to think, feel or otherwise be 'aware.' Diversity-awareness simply means that the designers of the system have taken into account aspects of diversity.

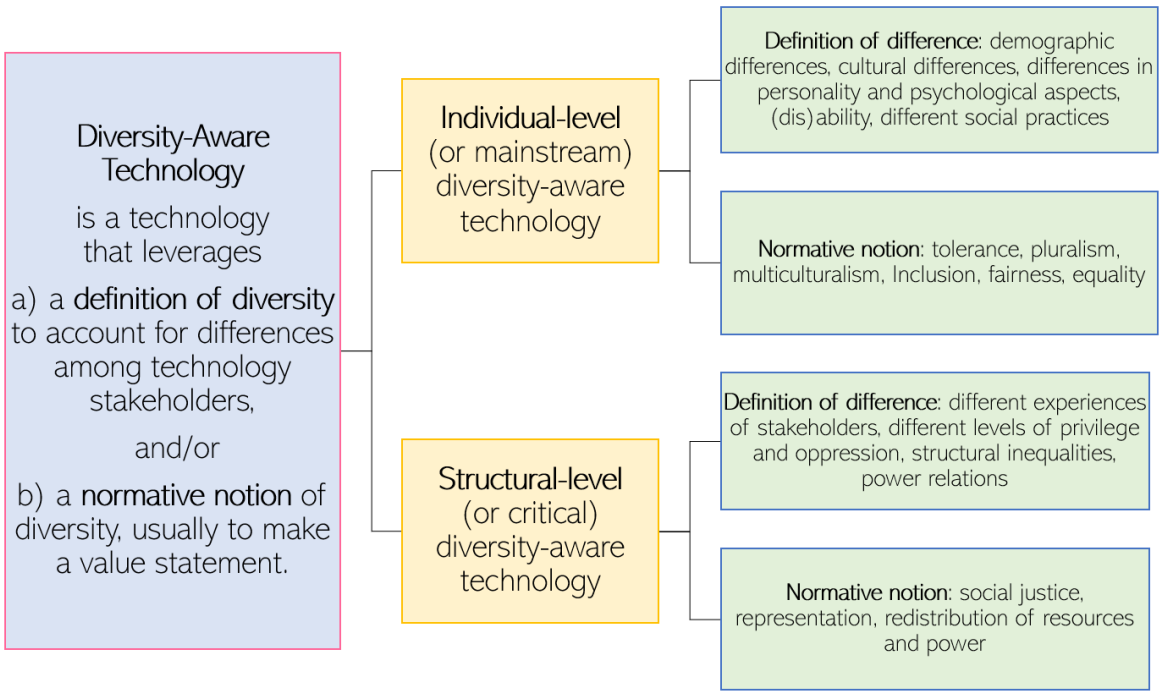

Figure 1: Visualization of the two families of diversity-aware technology

When discussing diversity-aware technology, it can be helpful to distinguish between two types: individual-level (or mainstream) diversity-aware technology and structural-level (or critical) diversity-aware technology. This distinction into two families of diversity-aware technology is developed in chapter 6. Individual-level diversity-aware technology embeds

categories and metrics such as gender, race, culture, personality in its design; these categories are understood as measures of difference at the level of an individual's user profile. The categories are usually operationalized as binary, simplified expressions of user identity and detached from considerations of larger social dynamics. The design of individual-level diversity-aware technology is usually accompanied by a language of inclusion, tolerance, and market expansion to underserved user groups.

An example of individual-level diversity-aware technology is the platform "WeNet – The Internet of Us." This platform was developed in an EU-funded research project. It seeks to bring together students, researchers, and developers to conduct studies of diversity and create new applications that leverage the diversity of its users (WeNet project, 2023). A proof of concept has been developed for an application chatbot, embedded in the Telegram platform. The chatbot "Ask for Help" connects users based on their requests for help and users' social practices (routine behaviors) that qualify them to offer help in response to the request. The diversity-aware algorithms thus work with an operationalization of diversity as social practices (more details in Schelenz et al., 2021). The technology is considered individual-level diversity-aware technology because social practices are seen as individual occupations of users. Designers thereby detach social practices from considerations of societal norms and expectations of behavior (e.g. 'women cook and men play soccer'). This becomes reflected in the data collection and algorithmic mediation, and may lead to bias (Schelenz et al., 2019).

Structural-level diversity-aware technology considers a form of structural diversity rather than individual diversity. This means that social, economic, political, and governance structures are taken into account when reflecting on differences between stakeholders of the technology. An example of difference from a structural viewpoint is the diverging experiences of discrimination that social groups are subjected to in society. Structural-level diversity-aware technology is concerned with the influence of those factors (oppression, marginalization, as well as privilege and advantage) on technology stakeholders. The idea of structural-level diversity-aware technology is to design in a way that first does not add to structural discrimination and second challenges unequal socio-technical relations. The design of structural-level diversity-aware technology is usually accompanied by a language of redistribution, non-discrimination, and a call for social justice. This implies an idea of technology design that does not prioritize profit. Rather than using diversity to scale a product, diversity helps imagine just socio-technical futures. This is reflected in the methods that typically produce structural-level diversity-aware technology: art-based and speculative design.

An example of structural-level diversity-aware technology is a virtual reality experience created by the Hyphen Labs under the term "NeuroSpeculative AfroFeminism" (Hyphen-Labs, 2019).[12] In a series of exhibitions and art festivals, the group of Black feminists expose their audience (students, general public) to a narrated fictional story of the future, where Black women are free from stereotyping, everyday racism, and structural limitations to their self-realization. The art installations reproduce a Black hair salon as a technology lab and harness the tradition of Black women's haircare as a meeting space for Black women scientists to develop new innovations. The virtual reality experience created by the Hyphen Labs can be considered structural-level diversity-aware technology because it relates to broader social contexts, juxtaposing present inequalities with imagined futures of Black women's liberation.

There are significant benefits of diversity-aware technologies. Although I see most benefits on the side of critical diversity-aware technology, there are also benefits of mainstream diversity-aware technology. Contemporary computer systems lack attention to the diverse needs of users (Wachter-Boettcher, 2017) and are prone to algorithmic bias and discrimination (Zou & Schiebinger, 2018). Improving the accuracy and fairness of technologies by drawing on individual-level user characteristics can prevent the worst discrimination. An example of a mainstream diversity-aware approach would be to collect diversity data to make datasets more representative of different groups, as datasets inform the development of a computer system. Collecting diversity data is important because most data currently originates in White, male, heterosexual, and Christian contexts (Criado Perez, 2019; Olteanu et al., 2019). A well-known example of algorithmic and data bias affecting Black women is facial recognition technology. Buolamwini and Gebru (2018) criticize that facial recognition systems exhibit higher inaccuracy for the classification of Black women's faces than White men, White women, and Black men. Providing a balanced dataset based on individual-level features (images of faces from female and male users and different skin colors) can optimize the technology's accuracy.[13]

The social benefits of critical diversity-aware technology go beyond making a computer system less prone to inaccurate classifications. Critical diversity-aware technology seeks to minimize unequal social relations. Facial recognition technology in airports, streets, and private apartment buildings has been used for surveillance and racial profiling (American Civil Liberties Union [ACLU], 2021; Kantayya, 2020; Lyon, 2003). A critical approach that looks at diversity through the prism of power, power systems, and power relations questions the improvement of facial recognition technology's accuracy because this can make the

[12] https://hyphen-labs.com/NSAF
[13] http://gendershades.org/

surveillance of marginalized groups more efficient. Such an approach (which can also be summarized under the term 'critique of power,' see below) challenges problematic designs. It allows for the abandonment of "captivating technologies," tools that extend governmental and societal practices of surveillance and carceration (Benjamin, 2019a, p. 3). The social benefit of critical diversity-aware technology is that it permits researchers and designers to question harmful technologies and rethink social relations and societal infrastructures.

Diversity and a Critique of Power Inspired by American Black Feminism

The research contributions in this dissertation draw on American Black feminist theory throughout individual chapters. US-based Black feminism is the theory and practice of understanding the lived reality of Black women from the time of American slavery until today and challenging the structural oppression that Black women continue to face. Black feminisms exist in plurality and have emerged in different regions in Africa, Latin America, the Carribean, and Europe (Emejulu & Sobande, 2019; Norwood, 2013; Rodriguez et al., 2016). I draw on US-centric Black feminism as American Black feminist works have increasingly targeted technology design and development as a site of critical inquiry (Benjamin, 2019a, 2019b; Marshall, 2023; McMillan Cottom, 2017; Noble, 2018; Noble & Tynes, 2016; Okonkwo, 2019). Furthermore, the USA is a global player in technology development (e.g. with the Silicon Valley in California) and has a long tradition of feminist and civil rights activism by Black women (Guy-Sheftall, 1996). An important analytical tool of American Black feminism (that has migrated to European contexts as well) is intersectionality; it helps render visible the co-constitution of systems of power (or systems of oppression) such as race, gender, class, sexuality, etc. (Crenshaw, 1989).

American Black feminism as a theoretical framework and intersectionality as an analytical tool enable the study of diversity in technology design and development from a perspective that is critical of existing power systems and power relations. This means that the definition and operationalization of diversity by designers, as well as discourses around diversity, fairness, and social inclusion, are scrutinized through the prism of 'power' in society and the technology industry. According to Dhamoon (2011, p. 240), "the constitutive feature of an intersectional-type research paradigm is a critique of the work of power – how it operates, its effects, and the possibilities of transformation." Attention to 'power' is key to the analysis of diversity concepts and discourses in this dissertation. Power refers to a form of domination and can be exercised in different domains: in the interpersonal domain (via the interaction between people), disciplinary domain (via rules and norms), cultural domain (via beliefs and

value systems), and structural domain (via the organization of society, relationship to government, etc., Collins & Bilge, 2016, 5ff). ‘Systems of power’ or ‘systems of oppression’ referenced in the dissertation are an expression of power and used as the vehicle by which domination is exercised.

Systems of power or systems of oppression are conglomerates of knowledge, narratives, and symbols that materialize in the creation and operation of institutions, policies, and practices. In the following, I use the terms ‘system of power’ and ‘system of oppression’ synonymously (cf. Collins & Bilge, 2016, p. 26). They are the same because a power system becomes a system of oppression for those who are not included in the hegemonic norm established as part of the power system (e.g. for race, the norm is White; for gender, the norm is male, Hull et al., 2010). The effects of the power system are thus vastly different depending on a person’s positionality, but the logic (to privilege some while disadvantaging others) is the same. Power systems are discussed in further detail when considering the benefits and limitations of American Black feminist theory, see later in this chapter und chapter 9.

Taking seriously a critique of power means that race must be considered a system of power rather than (as put forth in the ACM ontology for topics in computing) a user characteristic or demographic feature. The idea of ‘race’ entails a series of myths about the existence of different races and the natural superiority of one race over others (for example White supremacy, Bonilla-Silva, 2001). Race as a power system interacts with other systems such as gender, class, and sexuality, and produces particular experiences of oppression for individuals or social groups situated at the intersection of these systems, e.g., low-income Black women (Crenshaw, 1989, 1995). A scholarly field that investigates the employment of race narratives and how race is productive (in a bad way) in the oppression of Black people is critical race theory. According to Crenshaw et al. (1995, p. 25) “critical race theory allows us to better understand how racial power can be produced even from within a liberal discourse that is relatively autonomous from organized vectors of racial power.” This means that race as a system produces racism, but not merely in the form of violent apartheid. Race informs ideologies such as the belief that the government, courts, university admissions, and employers should be neutral as to the race of a person, an ideology called color-blindness (Bonilla-Silva, 2018). While non-discrimination (neutrality) is an important value in a society with little inequities, non-discrimination becomes a strategy of upholding inequality in a society with a history of racism. Hence, like more direct forms of racism, liberal attitudes in the form of color-blindness can produce oppressive race relations (Bonilla-Silva, 2018).

A critique of power can reveal subtle forms of domination under the guise of liberalism in (AI-based) technology. Liberal claims about the neutrality and inclusiveness of technology have been voiced for instance during the birth of the Internet. It was a moment of enthusiasm about realizing a colorblind digital society where, as long as you are a citizen of cyberspace, your race and gender do not matter (Daniels, 2015, p. 1382). Similarly, algorithmic mediation of decision-making (e.g., about loan applications) is considered neutral compared to purely human decision-making. Machines are expected to assist humans in making better (i.e. neutral or, by implication, 'fair') decisions (Benjamin, 2019b, p. 8; Roberge & Seyfert, 2017, p. 17). There are two problems that become obscured by liberal stories of neutrality in and through technology. First, the creation of algorithms relies mostly on the use of historical data. A society with a history of inequities will produce data that reflects these inequities. Hence, patterns of the past are reproduced in the present when a technology is trained on historical data.

Second, despite claims of diversity and participation in cyberspace, the Internet and widely used platforms and services have been developed in Silicon Valley and shaped by White male perspectives (Daniels, 2015, p. 1379). According to Cave and Dihal (2020, p. 2), AI design and development is dominated by a "white racial frame" (a term coined by Feagin (2013)). The White racial frame is a theoretical explanation of systemic racism, postulating that racial stereotypes and narratives, norms, or policies tracing back to White supremacist ideologies have been socialized over centuries to produce contemporary attitudes and worldviews of White people. The idea of systemic racism is that an individual does not have to be openly racist to be part of a racist system. All people are affected by this racist system; White people are affected by holding privilege that is largely invisible to them (McIntosh, 1988). Hankerson et al. (2016, p. 474) clarify how designers in HCI are affected by an overall racist system: "This is not to say that members of the HCI community are intentionally creating racist technology, as pre-existing social bias can be unintentional, one possible source is 'white privilege'." Also Cave and Dihal (2020) stress that racial bias may be unconsciously embedded in technology due to socialization rather than racist sentiment.

A critique of power helps untangle diversity definitions, narratives, and design strategies in computing, and analyze them as to their unconscious and invisible connections to larger structures of inequality, including systemic racism. To better understand hidden dynamics in the design of diversity-aware technology (such as the use of historical diversity categories, historical data, or implicit bias about human 'difference'), a critique of power necessitates a structural perspective. A structural perspective looks beyond the experiences of individuals with technology. It focuses on the experiences of social groups as 'structural identities' (Brah

& Phoenix, 2004, p. 77; Cooper, 2016, 389f). A structural perspective is relevant for computer scientists because it shifts the focus from optimizing for individual user experience to providing solutions that ensure fairness and inclusion at the societal level. The dissertation investigates the use of diversity in computer science and technology development from a structural perspective. It examines the social benefits and risks of diversity-aware technology for structural identities (social groups like Black women) and social relations (which are an expression of the relationship between different social groups).[14]

A structural perspective on diversity in society and technology helps shift the focus from individual-level demographic and cultural interpretations of diversity to social aspects of diversity. This shift is important because demographics and concepts of cultural difference dominate diversity discourses (Ålund, 1999; Dhamoon, 2010), thereby obscuring social inequalities between technology stakeholders. I do not want to deny the importance of a cultural perspective in discussing questions of belonging, identity, and social cohesion. Yet, cultural frameworks, promoted by theories of liberal multiculturalism, have become the explaining factor for differences between social groups that are in fact tied to social matters such as racism and wealth inequality. By way of example (inspired by my research in chapter 7), an Afghan refugee woman in Germany may lack safe access to technology *not* because of Islam or ethnic customs but because Afghan women in Germany lack digital literacy training and social support from the government. Hence, a structural perspective and a critique of power highlight that the different experiences (and needs, priorities, assets) of technology stakeholders must be analyzed against the stakeholders' differential treatment in society.

With a focus on the 'social,' technology design can center an overlooked element: that technology stakeholders' experiences are 'diverse' in their subjection to structural advantages and disadvantages (e.g. due to racial bias, Black women have been found to be affected most by inaccurate facial recognition technology, Buolamwini & Gebru, 2018). Anderson and Middleton (2018, p. 1) capture this notion of 'diversity of lived experiences' in their writing: "We live in a society that includes a population of people who are all similar yet 'different' in how they move through this world, experience this world, and are perceived by others in this world." In technology design and development, a focus on diverse lived experiences relates to how different social groups experience the interaction with computer systems, and how technology reinforces, maintains, or transforms existing power relations in society.

[14] Structural identities are often understood as identity groups, e.g., African Americans in the USA. In campaigns for human and civil rights, identity groups are invoked to gain leverage vis à vis the state. This, however, comes with essentializing tendencies, which means that complex identities are homogenized to stenghten the claim of an otherwise diverse group. Chapter 9 discusses in what ways essentialism can be problematic and when essentialization can be used strategically to achieve a certain goal.

A critique of power of diversity-aware technology is supported by different methods. For example, historical contextualization can reveal the influence of global value systems like modernity on diversity concepts. “Race is one of the central conceptual inventions of modernity, one that shapes our perceptions of social subjects, fellow human beings, in primarily racial terms” (Daniels, 2009, p. 20). Historically, race has been established (again, as a system) to subordinate some and privilege others in a modernist, colonial strive for progress and innovation (Desmond & Emirbayer, 2009, p. 338; Rusert, 2017). A historical contextualization thus sheds light on the legacy of diversity ‘categories.’ S. Browne (2015) offers an example for the historical contextualization of the practice of surveillance. S. Browne states that concepts like “race” are bound together with racist policies of “black luminosity” during slavery (i.e., Black people were forced to carry a light during the dark to be ‘recognizable’ by Whites, S. Browne, 2015, p. 67). Such historical knowledge can provide important insights for the contemporary analysis of facial recognition technology (cf. chapter 3). Other methods supporting the study of diversity-aware technology from a perspective that is critical of power systems/relations include conceptual and discourse analysis (chapters 4, 5, and 6) and the engagement of technology stakeholders (chapters 7 and 8).

Finally, a critique of power reveals the limitation of individual action. Designers and developers of diversity-aware technology cannot ‘solve’ structural inequalities in US-American and European societies and technologies, let alone in silos. Nor are designers and their products responsible for structural inequalities in society. Rather, a complex interplay of systems of power and infrastructures through governance, policy, socio-economic conditions, values, and social behavior generate inequality in society and technology. At the same time, individual action becomes a starting point for intervention (e.g. guidelines for designers on how to best employ diversity in technology), where individual action is a precursor to collective action (cf. Daniels et al., 2019). A critique of power then addresses both the individual level (designers, developers, students of computer science and humanities) and the structural level (technology companies, non-governmental organizations, government, higher education) by recognizing that the former make up and shape the latter (and vice versa).

Social Risks of Diversity-aware Technology

After already offering glimpses into the social benefits of diversity-aware technology, I will now address social risks that arise from working with diversity in computer science and technology development, and ways to mitigate those risks. The social risks discussed in the dissertation concern mainly individual-level diversity-aware technology but some risks may

also arise from structural-level diversity-aware technology. One social risk is that a reference to diversity might conceal potential harmful effects of a technology because diversity is often perceived as good and associated with inclusion, tolerance, and multiculturalism (Vertovec, 2012). As a result of this cognitive bias, diversity-aware technology may automatically be perceived as fairer, more just, and serving the public good. The mere term 'diversity' can thus legitimize a technology and obscure underlying social risks stemming from the operationalization of diversity and data collection. Diversity-aware technology may then escape critical inquiry. Mitigating this social risk requires constant awareness-raising. This dissertation aims to counter the idea that the blanket application of diversity is good; rather, it matters how diversity is operationalized and embedded in technology.

A set of social risks arises from the operationalization of diversity. Operationalization refers to the construction of diversity categories that are filled with information from technology stakeholders. This information is provided by the stakeholders or users themselves through self-reports or by data scientists who rely on existing and openly available data.[15] The information is stored in datasets in order to train a diversity-aware system. A reliable set of data points is required to build algorithmic systems. Data needs to be consistent, comparable, and interoperable according to gold standards in data science (Wilkinson et al., 2016). Hence, when developing mainstream diversity-aware technology, designers usually draw on commonly used categories such as demographics, culture, personality, etc. In technologies that seek to increase the fairness of algorithms, designers draw on race and gender (operationalized as skin color and gender identity, Hanna et al., 2020). This operationalization risks presenting 'difference' among stakeholders of the technology as neutral and natural, e.g., implying an alleged biological difference. However, 'difference' is produced by the social world. Race and gender are social constructs that carry values and order social relations (see above, Collins & Bilge, 2016, p. 26; Haslanger, 2012, p. 223).[16] The operationalization of race and gender as demographics obscures their origins in history, culture, religion, and governance in particular historical contexts.

Another social risk concerns the boundary-formation through the operationalization of categories in data science. It is common practice that (diversity) categories are defined in order to build technology. In this process, designers ultimately arrive at a set of categories that

[15] Openly available information must be labelled in order to correspond with the diversity categories underlying the dataset. Previous work has shed light on the exploitation of click workers in the Global South who are usually tasked with such labelling activities, see Gray and Suri (2019). This is marginally discussed in this dissertation and must be part of a future research agenda on diversity-aware technology. Seeing how information for diversity-aware datasets is labelled and who will label it is crucial to assess the social impact of diversity-aware technology.
[16] For example, a reference to human difference in terms of gender may carry unspoken views of 'proper' gender roles, hierarchies of labor, and reproductive practices.

excludes other possible sets of categories, and hence they ironically limit diversity. Defining a category means fixing boundaries. Fluidity, flexibility, and openness run counter to the logic of categorization. Users may be put in a box, stereotyped, and may be subject to stigma in an automated technology-mediated process. According to D'Ignazio and Klein (2020b, p. 10), "the process of converting life experience into data always necessarily entails a reduction of that experience – along with the historical and conceptual burdens of the term." Users are then reduced to a certain identity that, in reality, involves multifaceted experiences.

A related social risk is essentialization. This means that differences between technology stakeholders *within* a category are ignored. Both the boundaries of categories and essentialism make it especially difficult to account for technology stakeholders who are occupying continuums, are in-between, transitioning, or resisting definitions altogether. One such group is queer and transgender people (Costanza-Chock, 2020; Keyes, 2019). The risks described here pertain mostly to mainstream diversity-aware technology. The dissertation offers a design framework for critical diversity-aware technology that contextualizes these common categories and focuses on technology stakeholders' dynamic experiences rather than static labels (chapter 6). Critical diversity-aware technology is not primarily built on the categorization of people but the analysis of experiences. This said, there are essentializing tendencies even when an analysis of 'experiences' guides technology design. This is discussed in detail in chapter 9. It should be noted here, though, that essentializing effects may fuel stereotypes (or in the context of technology, cybertypes, Nakamura, 2002; Perpich, 2010). Hence, also designers of critical diversity-aware technology must reflect on the social risks of their accounts of diversity.

Another social risk that arises from the operationalization of diversity is the potential for misuse of diversity categories for the classification, surveillance, and oppression of some social groups. The practice of accounting for numbers and characteristics of people has a foul legacy. It has been used in different episodes of history to control, subordinate, and eradicate groups of people. The "Book of Negroes" details the identity of enslaved Black people who joined the British army during the American Revolutionary War and were tracked by American slave `owners´ to be `repossessed´ (S. Browne, 2015, p. 84). The Nazi government in Germany during the 1930s and 1940s systematically collected data from the German population to identify Sinti, Roma, and Jews for eradication (Luebke & Milton, 1994; United States Holocaust Memorial Museum, 2021). While the design of diversity-aware technology usually involves a normative sense of justice and fairness, some actors may coopt this language and misuse its practice for ulterior motives. In addition, visions of what is 'just' can be subject to

disagreement between political traditions. The classification of people (especially if based on notions of 'difference') should then always be seen with a certain skepticism.[17]

This leads us to mitigating design actions that can prevent the 'naturalization,' essentialization, and surveillance of social groups. One such action is the contextualization of diversity categories through examination of the lived experiences of technology stakeholders. Contextualization means that the socio-economic-political context of technology stakeholders is mapped for consideration in the design of the diversity-aware technology. Wong-Villacres et al. (2018) provide a best practice example of such a mapping. They consider the experiences of "teachers, parents, children, and other stakeholders" in low-resource educational contexts in India to design technological interventions that can support the learning experience (Wong-Villacres et al., 2018, p. 47). The authors used ethnographic methods (interviews and participant observation) to draft narratives around the factors influencing learning experiences. They state:

> "our purpose was to deductively identify interacting processes of differentiation and systems of domination, and uncover how individuals (teachers, parents, children), communities [in low-resource learning environments], and organizations (non-governmental and governmental organizations) interacted with those, to uphold or to resist them […] This […] highlighted both challenges to technology adoption and supporting structures for new technology designs" (Wong-Villacres et al., 2018, p. 48).

Ethnographic fieldwork is an effortful way to contextualize technology stakeholders' diversity. The resources for ethnographic research may not always be available to designers.

Another way to contextualization – especially when contextualizing individual-level diversity categories such as race, gender, personality, disability – is historical and cultural analysis. For instance, while personality seems utterly individualistic (and not structural or social), a cultural analysis (based on the analysis of social practices of domination) reveals the frequent use of stereotypes in describing the personality of Black women. Through long-standing media representations of Black women as strong, angry, loud, and wanton, but also lazy, unemancipated, welfare-reliant or bad mothers, Black women have been associated with certain personalities (Ladson-Billings, 2009; West, 2018). Personality remains one of the primary measures of diversity in computer science and social science. For example, the use of the Big Five personality traits (a model combining five factors of personality: openness, conscientiousness, extraversion, agreeableness, neuroticism) is common in the fields. When

[17] Even seemingly positive labels should be consumed with caution, for example when referring to Black women as 'brilliant minds' or 'strong' and 'enduring.' On top of this being essentializing, such characterization can lead to hypervisibility and unrealistic expectations, e.g. for 'exceptional' women in the academy, see Mirza (2009b, p. 235). In the technology industry, it means work overload for Black women who are perceived as being particularly resilient in light of chronic stress, according to J. O. Thomas et al. (2018).

theorizing diversity as personality, it is important to keep in mind that a person's personality may be unconsciously tied to stereotypes of the social group to which they are ascribed.

Before concluding the overview of social risks and the mitigation of these risks, I address the collection of diversity data for diversity-aware technology. Diversity data can be highly sensitive data points. Plus, the collection of data about gender, race, ethnicity, socio-economic capacity, personality, social practices, relationships, physiology, and more can generate huge datasets about users. Such information can be used to profile users, e.g., for targeted advertising. If hacked or stolen, such information can be used to extort users. Susser et al. (2019, p. 38) state: "The more that is known about each person's personality, preferences, habits, and vulnerabilities, the easier it is to construct choice environments that will guide their decision-making in the desired direction." Because of the sensitive nature of diversity data, collecting such data from users further strengthens existing power relations between technology developers and users. Technology companies working with diversity may gain unprecedented knowledge about individuals and social groups by including them into 'diversity datasets' and potentially harmful data maintenance practices (cf. Hoffmann, 2020). To mitigate this social risks, privacy and data protection regimes are important resources for designers and developers of mainstream diversity-aware technology. A critical diversity-aware technology would likely employ models of data ownership and user self-control of data.

Benefits and Limitations of Using an American Black Feminist Approach

Before getting into the benefits and limitations of using American Black feminism to study and design diversity-aware technology, I explain my decision to apply American Black feminist theory in the first place, and I clarify my position. Working with American Black feminism invites the researcher to reflect on their own positionality in society. This is because Black feminism is a theory of situated knowledge and has been developed and applied through the lens of Black women's consciousness of power and oppression. There are then limitations of a White researcher's understanding and application of Black feminism. For me, a White cis-woman and German citizen, such limitations arise. I have not experienced the everyday racism (Sue, 2010) that Black women in the USA continue to experience as sometimes small stings and sometimes loud rejections (Kallaway, 2015).

However, Black feminist writings with their attention to power and their intersectional perspective are instrumental in reflecting on diversity and difference in society. Black feminists' classic contributions as well as more recent works in science and technology studies or computer science are crucial to critically analyze diversity in relation to technology. The

position of Black feminists at the margins of society offers important epistemological advantages to consider who is missing in contemporary debates about just and 'diversity-aware' socio-technical futures. Some are critical of White people's applications of Black feminism in research, as White people tend to obscure or appropriate the achievements of Black theorists (Cooper, 2016, p. 404). While I acknowledge the limitations that my perspective brings to applying Black feminism, I believe that, if the theory is well-suited to analyze the topic at hand, a refusal to apply Black feminism due to the Whiteness of the scholar would further marginalize it. In fact, to 'respect the limitations in understanding Black women's experiences' has been used as a strategy to dismiss the theory altogether and instead promote White male frameworks of thought. Lorde (1980) calls out the behavior of White women to refuse engagement with Black feminist thought for positionality reasons:

> "All too often, the excuse given is that the literatures of women of Color can only be taught by Colored women, or that they are too difficult to understand, or that classes cannot 'get into' them because they come out of experiences that are 'too different.' I have heard this argument presented by white women of otherwise quite clear intelligence, women who seem to have no trouble at all teaching and reviewing work that comes out of the vastly different experiences of Shakespeare, Moliere, Dostoyesfsky, and Aristophanes. Surely there must be some other explanation" (Lorde, 1980, no page numbers).

It is thus a conscious decision to apply Black feminism from a White perspective, accepting the limitations that come with it. Throughout the dissertation, the knowledge gained from applying Black feminism motivates my further engagement with it.

Another important point of reflection in working with American Black feminism is that the theory has a difficult relationship with diversity. Black feminists have long criticized the politics of difference in feminist movements, but also condemned diversity work at the organizational level for being apolitical and non-transformative (S. Ahmed, 2012; Nash, 2019, p. 24). Diversity language has been used to tokenize Black women, and companies have profited from invoking Black women's labor without duly recognizing Black women or offering them more power (Crawley, 2006). Similarly, Black women have been 'promoted' to deal with diversity, equity, and inclusion at the university for no additional payment, and are silenced when they bring up racist structures in their organizations (S. Ahmed, 2012; Mirza, 2009b). Making this difficult relationship visible up front helps to avoid pitfalls and consider diversity not as an economic strategy or "happy" work (S. Ahmed, 2009, p. 43) that focuses on how 'we can all get belong' (Bonilla-Silva, 2015). Rather, diversity must be seen as a question of power relations. Throughout the dissertation, I have looked at diversity from a critical perspective that highlights social inequality as a matter of unequal power relations.

Black feminism, like any theory, has advantages and disadvantages. These also affect the study and design of diversity-aware technology. In chapter 9, I reflect on the application of American Black feminism in the context of studying and designing diversity in relation to technology. The major strength of Black feminism is the explanatory force of its power angle. Power systems such as race have already been referenced above. Power systems structure social relations by defining a norm and an 'Other,' thereby producing 'human difference,' defining hierarchies, and actuating structural experiences of privilege and oppression. If power systems are reinforced by embedding problematic diversity concepts into technology, they can invisibly hinder technology stakeholders' chances at access and safe/empowering use of technology. The dissertation does not dismiss the relevance of individual experiences in the lives of people and in interaction with technology. The dissertation focuses on structural experiences, though, to reveal underlying power dynamics that affect individuals as (externally ascribed) members of social groups.

Apart from a look at the benefits of a Black feminist approach to the study and design of diversity-aware technology, chapter 9 reflects extensively on its limitations. One limitation is the priority of race and gender as power systems in the analysis of technology stakeholders' experiences. On the one hand, the benefit of a focus on power relations/systems is that Black feminist frameworks can be applied across geographic, cultural, or social contexts. Power systems (like race, gender, and class) arguably affect people across the world.[18] Chapter 7 shows how Black feminism helps analyze the structural experiences of Afghan refugee women in Germany with technology. For instance, gender norms as an expression of power relations in Afghan society and Afghan diasporic communities shape Afghan refugee women's experiences in online social media. On the other hand, different power systems do not receive equal attention. Chapter 9 argues that power systems such as language, religion, and wealth require attention to account for the lived reality of diverse technology stakeholders. For example, religion is entangled with gender norms in the experiences of Afghan refugee women, and Afghan women users of social media require high levels of privacy and safety. Religion as a power system may shape the design of technology in the sense that secular or Christian gender norms dominate design processes. As a result, Afghan refugee women's need for gender-

[18] Norwood (2013, p. 227) shows that the significance of different power systems in structuring social relations vary across context. They cite African thinkers who criticize Western feminist movements for a focus on gender as the primary determinant of power in society: "Oyeronke Oyewumi (1997), author of *The Invention of Woman, Making an African Sense*, makes the case that in Nigeria (especially among the Oyo Yoruba), social stratification was not historically organized along a gendered dimension, but instead in relation to age. 'Seniority' was the defining element that characterized power differentials in an African [Nigerian] context" (p.227, emphasis in original).

separated and ‘female only’ online spaces, motivated by their lived reality as Muslim women, may be marginalized. These thoughts encourage the reiteration of the importance of attending to less ‘popular’ power systems, such as language, religion, and wealth to understand the full experiences of technology stakeholders in different geographic and cultural contexts.

Beyond imbalanced attention to different power systems, Black feminism exhibits further biases. One concern expressed by Black feminists themselves and critics from queer and postcolonial studies concerns essentialism. Nash (2008) and Puar (2012) argue that identity politics around race and gender reinforces problematic categories rather than dismantling them. The category of ‘Black women’ is a problematic one for Nash and Puar because it obscures diversity within the group of Black women. Nash (2008, p. 6) points out that, ironically, Black feminists have made precisely this argument towards White women in light of their own exclusion from feminist movements. Two analytical foci of intersectionality are under scrutiny in chapter 9 with regard to essentialism: the additive and structural interpretations of intersectionality. Yuval-Davis (2006) criticizes the ‘additive approach’ of intersectionality which understands race and gender as social identities and adds them up to account for Black women’s realities. This obscures complex new configurations of power that emerge at intersections (e.g., when gender is racialized). A focus on Black women’s ‘experiences’ in the structural interpretation of intersectionality (put forth in this dissertation) is seen as only slightly different. Experiences may become naturalized just like social identities, for example when oppression becomes an allegedly unpreventable fact of Black women’s lives (Nash, 2008, p. 12). The risk of essentialism can thus not be eradicated even when working with a socially sensitive theory such as Black feminism. A way to mitigate essentialism is to contextualize identities and experiences, but there may also be moments where essentialism can be used strategically to advance social justice.

Another concern about Black feminism relates to the epistemologies underlying a Black feminist critique. Some scholars working with postmodern, poststructuralist and posthumanist theories suggest that our contemporary ways of knowing are based on the violent constructs of human difference established during modern imperial expansion, colonialism, and slavery. Puar (2012, p. 54) condemns the use of the concept ‘woman’ and the practice of establishing the Black woman as a human subject in pursuit of rights. Since concepts like ‘woman’ or ‘human’ were made hugely prominent during colonization to create gendered and racial hierarchies, using the terms only cements Black women’s oppression (McKittrick, 2015; McKittrick & Wynter, 2015). Black feminism exhibits a tension between approaches using Western epistemologies to claim Black women’s rights and those abandoning the modern language of

subjectivity and human rights. This tension reflects a diversity of pathways towards social change (reform vs. a 'clean slate') and has implications for the design of socio-technical futures. Chapter 9 discusses the complexities, tensions, and necessary compromises that arise when working with Black feminism in the study and design of diversity-aware technology.

The theory-related constraints on the study and design of diversity-aware technology inspire a broader consideration of constraints on critical research and design, and how researchers and designers can deal with such constraints. Chapter 9 argues that, even though researchers and designers might envision or desire radical transformation, they can move forward within constraints by relying on careful negotiation, patience, strategic action, and cooperation. The chapter highlights the potential of Islamic feminisms[19] and the idea of 'working within constraints' (Rabaan et al., 2020, p. 16), i.e., working within larger structures (not just divine ones but also technical and societal ones). According to Islamic feminists, agency can be exercised within these structures and technology can offer pragmatic innovations closely aligned with the everyday needs of Muslim women (Rabaan & Dombrowski, 2022). Chapter 9 brings Black feminism in conversation with Islamic feminism and shows that a critical approach to the study and design of diversity-aware technology needs both aspirations to structural social change and smaller, pragmatic steps towards relief for technology stakeholders, who experience algorithmic discrimination, online violence, and data exploitation in their daily interaction with technology. Islamic feminism then complements Black feminism.

The engagement with Islamic feminism shows that American Black feminism is only one theory of many which can help make sense of diversity in society and technology. Last but not least, I wish to highlight postcolonial feminisms as another important theoretical orientation for the study and design of diversity in relation to technology. Postcolonial feminisms highlight the experiences of women of color in the Global South as well as diasporic communities. Postcolonial and Black feminisms show similarities in their critique of power and the racialized and gendered oppression that women of color face as national citizens of the UK or USA, immigrants and refugees, and transnational women (Mirza, 2009a, 2013). Whereas American Black feminism grounds historical analysis mostly in Black women's experiences during slavery, postcolonial feminisms point to the histories of settler colonialism and native resistance. A focus on colonialism and the Global South is important in the context of technology design, development, and distribution. Technology may pose a threat of digital

[19] Islamic feminisms exist in different Muslim countries and are shaped by local politics, for example in Iran, see Ahmadi (2006); Egypt, see Picchi (2017); and Afghanistan, see Choudhury (2007). In the following, I will use Islamic feminism in singular form, though, for the sake of simplicity, and to reference a broader feminist praxis of Muslim women to question established gender norms that are prescribed by male interpretations of the Qur'an.

colonization due to the predominance of Western values in design and the extraction of data and hazardous labor for the digital economy from sites in the Global South (Coleman, 2019; Gray & Suri, 2019; Schopp et al., 2019). Regarding American Black feminism, I believe it is a fruitful framework for the study and design of diversity-aware technology but can and should be complemented with elements from other feminist movements.

Reading On

By reading on, I hope that the reader will gain new insights about the complexity of diversity when embedded into technology. I further hope that the reader will come to appreciate the importance of connecting questions of diversity to broader structures of power and social relations in our societies. Engaging with diversity is not easy, and especially in the area of computer science and technology development, we are at the very beginning of tackling questions of diversity at scale. Diversity as a new mainstream design paradigm will be here to stay, so I believe, and diversity categories will increasingly be used to optimize technologies. Yet, diversity interpretations are all over the place without shared perspective or guidance to designers, and without meaningful awareness of power, social inequity, and injustice. Framed more positively, and borrowing a panel title from the 2021 ACM conference on Fairness, Accountability, and Transparency (FaccT'21), "the future is up for grabs" (Marda et al., 2021). This means that there is still much room to shape the future of diversity-aware technology.

This dissertation aims at making a contribution to shaping the future of diversity-aware technology through critical reflection of diversity and design, and by offering a constructive framework for working with diversity in a way that considers power systems/relations. This said, the dissertation is far from immune to the typical fallacies of working with diversity. Some strategic decisions about the representation of stakeholders in diversity-aware technology development were made that may have an essentializing effect. For example, I refer to "computer scientists" and "designers" as a seemingly homogenous group to describe tendencies in the field and juxtapose them with critical viewpoints from the humanities. Sometimes, I am more specific by referencing mainstream and critical approaches within computer science and design. However, it must be clear that computer science, being such a big field, employ diverse case studies, methods, and theories, and have varying relationships to the humanities.

Another strategic decision is to focus on "women" as a reference point for the treatment of diversity. This materializes both in the reference to Black women (throughout the dissertation) and Afghan women (in chapter 8). The practice of speaking of women (in line with more traditional feminist approaches) excludes non-binary, queer, and transgender persons who

are certainly affected by diversity conceptualizations in computer science. This critique is discussed in chapter 9. I hope that the reader will see the contributions in this dissertation as a starting point of engaging with diversity in computer science and technology development rather than an all-encompassing declaration on the topic. I further hope that the reader will be inspired to explore other aspects of diversity and technology.

Chapter 2 recapitulates a genealogy and core insights of an American Black feminist tradition of thought and action. It presents a synthesis of contemporary perspectives on the oppression of Black women through AI, and their resistance against AI or strategies of co-option. The goal of the chapter is to offer a first impression of the rich material that exists on the study of AI-based technology and Black women's experiences. Previous works synthesized in the chapter deal with recruitment AI, crowdsourcing work, surveillance, face recognition, hate speech, social media, visual representation, search engines, and hashtag activism. In its first part, the chapter is structured around the economic, political, and cultural oppression of Black women through AI (in line with Patricia Hill Collins's theoretical take on the oppression of Black women, Collins, 2007, p. 4). The second part presents Black feminist political activism, new frameworks for the study of AI-based technologies, and new frameworks to design technology for social justice.

Chapter 3 argues that American Black feminist theory adds new insights and raises important questions about AI-based technologies that extend work in the AI ethics community. The chapter explores the case study of facial recognition technology and surveillance through literature analysis and historical contextualization. By engaging Simone Browne's (2015) study of socio-technical control of Black people in American history, I use concepts like panoptic surveillance and 'sousveillance' to theorize Black women's oppression and resistance in relation to contemporary AI-based facial recognition technology. The chapter demonstrates that a Black feminist approach widens our perspective to ask whether a system like facial recognition (that is frequently used for racial profiling) can be designed ethically. Black feminists prompt us to reflect on the logic of surveillance itself, and, if necessary, dismiss harmful technologies. The chapter also marks Black feminist resistance and political action against facial recognition with remarkable successes in recent years.

Chapter 4 applies a perspective that is critical of power systems/relations inspired by Black feminism to the definition and use of diversity concepts in computer science. It asks how computer scientists conceptualize diversity and what the consequences are in terms of social justice. Through a literature review of 120 papers, I reconstruct trends of operationalizing 'diversity' in two subfields of computer science: recommendation systems and HCI. The

chapter finds shortcomings in the contemporary employment of diversity in computer science, especially in terms of the recognition of rights and identities of social groups. The analysis is followed by a call to (re)link diversity with concerns for civil rights. The discussion hints at a tension between neoliberal values of the digital economy and diversity as representation and redistribution.

Chapter 5 goes more into detail on diversity in the area of recommendation systems. The chapter explores computer scientists' notions of diversity in relation to fairness (fair recommendations), and focuses on an understudied aspect in recommendations: user diversity. I draw on Black feminism and critical race theory to offer a critical perspective on the notion of fairness, also falling back on Deborah Hellman's (2018) idea of 'compounding injustice.' The chapter points to the risk that leveraging individual-level, binary, and static notions of user diversity (even in a well-intended effort to increase fairness in recommendations) can compound previous injustice. A hypothetical use case around 'education' illustrates some of the mistakes that can be avoided by contextualizing diversity categories. The chapter offers concrete guidance to developers in the form of three principles: making a concept of user diversity explicit; reflection; and contextualizing a concept of user diversity.

Chapter 6 responds to the challenges laid out in the previous chapters and establishes a definition and design criteria for diversity-aware technology. I distinguish two families of diversity-aware technology: individual-level (or mainstream) and structural-level (or critical) diversity-aware technology. Diversity from a structural perspective relates to social aspects of diversity such as power inequities. Accordingly, systems that leverage structural-level diversity concepts focus on reflecting and transforming power relations through socio-technical solutions. The chapter theorizes 'difference' from a Black feminist perspective. Intersectionality helps understand how people and social groups are different in society: not because of their ascribed attributes or cultural orientation but rather because socially constructed ideas of those attributes are used to subordinate some people and privilege others. The chapter then paves the way to seeing diversity in society differently from established views in technology design and development, and consider user (lived) experience in new ways.

In chapter 7, the general direction of the dissertation shifts to incorporate broader dynamics of diversity, such as the social inclusion of refugees through technology. The geographic focus also shifts entirely to Europe. The chapter presents an empirical study with the participation of 14 Afghan women conducted in a big city in Germany. The empirical study documents the structural experiences of Afghan refugee women, who are located at the intersection of various power systems, and how these experiences influence access and use of

technology in Germany. Applying Black feminism to the case of Muslim women of color in Germany showcases the strength of the theory in understanding marginalization in different socio-technical contexts. Furthermore, the chapter presents design features preferred by participants (such as audio function and real-time assistance), which can help designers adapt services to refugees' priorities. However, as the discussion shows, technological tools are not always and not automatically the most effective way to foster social inclusion, considering digital illiteracy of Afghan women and a lack of digital literacy training in Germany.

Chapter 8 provides a comparison of US-American and Western/Northern European perspectives on critical diversity and design with a focus on the HCI community. HCI is more and more interested in improving the experiences of different populations in their interaction with technology through participatory, just designs. Emerging perspectives on diversity and design in the USA and Europe highlight the role of power in design. The chapter presents an analysis of in-depth expert interviews with 16 HCI researchers and designers based in Europe in comparison with a literature review of US-based perspectives in HCI. The juxtaposition of the two discourses reveals similarities and differences, and highlights the contributions of HCI to rethinking design.

Chapter 9, which circles back to the theory driving the analyses in this dissertation, looks in more depths at the application of Black feminism as a critical social theory. The chapter explores and discusses the possibilities and limitations of the theory to contest AI-based technology and processes such as sorting, classification, personalization, recommendation, and population-based predictions. In a way, the chapter evaluates the strong points and weaknesses of using a critical social theory to research diversity and technology. I highlight the promise of an intersectional-type analysis and a structural perspective enabling the in-depth analysis of technology stakeholders' experiences with AI. The chapter also discusses limitations of the theory when it comes to the critical study of AI. At the end, the notion of 'constraints' on critical scholarship on AI is explored.

In the concluding remarks added at the end of the dissertation, I jump off chapter 9 to offer some ways forward for the critical research and design of diversity-aware technology. I expand on the idea of working within constraints that has been raised in chapter 9. Bringing Black feminism in conversation with Islamic feminism, the concluding remarks argue that researchers and designers can move towards social justice goals within the constraints they are facing, motivated by a vision for radical transformation, and relying on strategic action, patience, solidarity, and feminist coalitions.

PART I:

BLACK FEMINISM – A THEORETICAL FRAMEWORK FOR THE STUDY OF ARTIFICIAL INTELLIGENCE-BASED TECHNOLOGIES

Chapter 2: Artificial Intelligence between Oppression and Resistance: Black Feminist Perspectives on Emerging Technologies

[This chapter was originally published in: Schelenz, Laura. 2022. "Artificial Intelligence Between Oppression and Resistance: Black Feminist Perspectives on Emerging Technologies." In *Artificial Intelligence and Its Discontents*, edited by Ariane Hanemaayer, 225–49. Social and Cultural Studies of Robots and AI. Cham: Springer International Publishing, reproduced with permission from Springer Nature]

Opening quote: "[Oppressive] systems and structures negatively impact the everyday lives of real people, and as such, the work we do aims to inspire the design and use of technology that produces positive change in the lives of those deemed marginalized, underserved, or vulnerable" (Y. Rankin & Thomas, 2019).

This chapter traces Black feminist perspectives on the design and use of artificial intelligence (AI). In recent years, Black feminist scholars and activists have raised awareness of the destructive impact of AI on the lives of Black, Indigenous, and Women of Color (BIWOC).[20] Black feminists have argued that historical interlocking forms of oppression are reproduced through AI-based technologies (Benjamin, 2019b; Noble, 2018). They further criticize that AI design is White-centered: it is created by mostly White and male system designers; it builds on data from overwhelmingly White and male people, and thus serves the White male norm best (Buolamwini & Gebru, 2018). Especially when AI is used to make decisions about college admissions, granting of loans, employment, and recidivism, built-in algorithmic biases lead to the unfair discrimination of BIWOC (O'Neil, 2016). Drawing on Black feminists' critique of AI, the first part of the chapter sheds light on three dimensions of *oppression* through AI: the economic, political, and cultural oppression of Black women.

However, Black feminists do not remain invested in uncovering oppression. Based on their analyses of the discriminatory power of AI, they have also engaged in *resistance* to White-centered technologies. This dialectic of oppression and resistance is a characteristic element of Black feminism (Collins, 2000a, p. 3). Resistance against harmful AI includes political activism, science fiction and art production, and the development of alternative social justice-

[20] In this paper, I center the scholarship and experiences of Black women particularly. The arguments may be relevant for multiple marginalized groups, including Indigenous and Women of Color. Having said that, distinct Indigenous and scholarship from the Global South attends to the particular challenges of AI for these populations, and its contribution to the study of AI should be explored in further work (Arun (2020); Mohamed et al. (2020)). Since I am a White European woman, my interpretation of the Black feminist frameworks in this paper constitutes a White perspective. My analysis is limited by my positionality: my White privilege inhibits my understanding of the actual experiences of Black women with technology.

oriented frameworks and designs for AI. Black feminists have also leveraged AI-based platforms such as Twitter to campaign for their rights (Jackson et al., 2020; Lumsden & Harmer, 2019, p. 10). The second part of the chapter maps some of the more recent resistance by Black feminists with or against AI.

A Black feminist view on Artificial Intelligence is a distinct one. It is interesting because it provides a comprehensive analysis that reveals AI's discontenting impact on users at the intersection of race, gender, class, sexuality, and more. Precisely because BIWOC have been and are structurally disadvantaged in American society, they have developed a perspective and practice that aims at dismantling *all* systems of oppression. They uncover the manifold challenges of racist and sexist AI *and* offer insights for envisioning futures of social and economic justice with or without AI. This paper applies a US-centric Black feminist lens. While sporadic perspectives from the Global South are included, most authors cited are situated in the USA. On the one hand, the United States, and especially Silicon Valley, is a leading hub for technological innovation such as AI-based products and services. On the other hand, the United States is a place of vibrant Black feminist thought due to Black women's intellectual and political work from slavery to today.

Black Feminist Theory – Brief Overview

Since this paper deals with Black feminist perspectives on AI, this section specifies the theoretical and foundational elements of Black feminism. What is a Black feminist perspective? There are numerous strands of Black feminism (Collins, 2000a, p. 22; Africana womanism, Hudson-Weems, 1989; Kawaida womanism, Karenga & Tembo, 2012; African or Afro-centric feminism, D. Lewis, 2001; womanism, Walker, 2004) and scholars frequently speak of Black feminisms in the plural. Yet, some core assumptions are held across different streams. For instance, a central argument is that Black women's lived experiences are shaped by a special kind of oppression, a convergence of racism, sexism, and classism. Based on this experience, Black women developed a knowledge and consciousness that demands tackling multiple forms of oppression simultaneously to achieve social justice (Guy-Sheftall, 1996, p. 2). In the US-American Black feminist tradition, at least two major experiences shaped the idea of multiple oppressions.

First, during slavery, Black women and men experienced the same inhumane treatment (being treated as property, heavy labor in fields, whippings, separation from family and children). However, on top of this degradation based on race, Black women suffered sexual abuse as domestic slaves based on gender (hooks, 2015a, 24ff). Second, during the Women's

Suffrage Movement, Black women were excluded due to the racism of their white colleagues. Later on, in the Civil Rights Movement, Black women were asked to subordinate their feminist and economic goals under the anti-racist cause (hooks, 2015a, pp. 3–7). Hence, Black women were oppressed in social movements either based on their race or based on their gender. This experience informed their critique of single-axis analyses of oppression (Beal, 1996; King, 1988). Although Black feminism stresses that Black women experience structural disadvantages, the theory does not claim that every single Black woman in America is oppressed (Collins, 2000a, 25f). The argument is then a *structural* one, placing more emphasis on systemic dynamics rather than individual experience.

The idea described here is known as intersectionality (Crenshaw, 1995), which holds that people are situated at the intersection of various systems of oppression. For instance, a lesbian White woman is affected by heterosexism and patriarchy at the same time. A lesbian Black woman is affected by heterosexism, patriarchy, and racism. All forms of oppression are equally constitutive of a person's lived reality and matter – whether the person is affected by racism, (hetero)sexism, classism, ableism, ageism, etc. (Lorde, 1980). Hence, intersectionality describes the experience of multifaceted structural discrimination. Simultaneously, intersectionality is a lens and, in academic contexts, a method for "understanding and analyzing the complexity in the world, in people, and in human experience" (Collins & Bilge, 2016, p. 2).

Other features of Black feminism include the balance of theory and practice and the dialectic of oppression and resistance. Concerning the former, it is crucial to recognize that Black feminist thought does not necessarily originate in academia. It is rather fed by arts, music, testimony, and other forms of expression by Black women (Collins, 2000a, p. 9). Furthermore, the Black feminist experience has been framed as a dialectic one, in which Black women are oppressed but simultaneously resist their oppression (Collins, 2000a, 3f). This makes Black feminism distinct from other theories.

> "As an historically oppressed group, U.S. Black women have produced social thought designed to oppose oppression. Not only does the form assumed by this thought diverge from standard academic theory – it can take the form of poetry, music, essays, and the like – but the *purpose* of Black women's collective thought is [...] to find ways to escape from, survive in, and/or oppose prevailing social and economic injustice" (emphasis in original, Collins, 2000a, p. 9).

These defined characteristics of Black feminism can be traced in the Black feminist perspectives on AI that are presented in the following.

AI & the Oppression of Black Women

"I believe that artificial intelligence will become a major human rights issue in the twenty-first century. We are only beginning to understand the long-term consequences of these decision-making tools in both masking and deepening social inequality" (Noble, 2018, p. 1).

According to Collins (2000a, p. 4),

> "oppression describes any unjust situation where, systematically and over a long period of time, one group denies another group access to the resources of society. Race, class, gender, sexuality, nation, age, and ethnicity among others constitute major forms of oppression in the United States."

The concrete dynamics of oppression may change over time or may significantly differ across societies. Social context matters in understanding oppression (Collins & Bilge, 2016, p. 28). In the United States, oppression is structured around schools, housing, employment, government and other social institutions (Collins, 2000a, p. 227).

Historically, BIWOC have been oppressed in three dimensions. In the economic dimension, Black women have been exploited as free or cheap labor, living in poverty and constantly struggling to survive. They have often not had the possibility to do intellectual labor and lacked social mobility. In the political dimension, BIWOC were not afforded the same rights as White people, were segregated in schools, education, and lived in poor urban or rural areas, which were framed as ghettos. In the cultural dimension, Black women were depicted in demeaning ways (particularly in sexual terms) to uphold the idea that they were inherently inferior. These images of Black women (as "Mammies," "Jezebels," and "Sapphires") were used widely in popular culture (Collins, 2000a, p. 4). Following this distinction, I will look at the economic, political, and cultural aspects of Black women's oppression through AI.

In the context of AI, Black women are affected by various unjust situations and structures. On the one hand, they are impacted by design architectures and algorithmic networks that have been found to exhibit bias against BIWOC. This bias can be traced to an overrepresentation of White male perspectives in datasets and design teams (Buolamwini & Gebru, 2018; Wachter-Boettcher, 2017; Zou & Schiebinger, 2018). Since computer models are increasingly used to make decisions that can determine subjects' paths in life, such as decisions about loans, hiring, college admission, and even parole (O'Neil, 2016), algorithmic bias can reinforce and intensify the structural oppression of BIWOC. On the other hand, the economic and political frameworks of the production and use of AI contribute to the oppression of Black women. The operation of AI-based tools relies on the recruitment (and exploitation) of low-wage, low-skill workers in the Global South (Gray & Suri, 2019). Data flows freely and is mined with little regard for data protection (Zuboff, 2019), especially stripping poor people of

their privacy rights (Bridges, 2017). This ecosystem of AI-powered services remains largely unregulated and is further legitimized by the commonly held belief that AI-based technology is neutral and fairer in its decision-making than humans (Benjamin, 2019b, p. 11 and 22).

Economic Oppression

This section highlights two trends in the digital economy that can add to the marginalization of BIWOC. Concerning the first trend, AI is increasingly used for the automation of hiring decisions and recruitment, especially by large employers. Applications and CVs are pre-scanned by a software to distinguish suitable from unsuitable candidates (Poster, 2019, p. 139). This pre-selection of candidates is done without human oversight. Amazon stopped the usage of a hiring software after the company found that it exhibited gender bias by favoring male over female applicants (Reuters, 2018). The biased algorithm was trained on data from the last 10 years of applications at Amazon. Given societal inequalities that lead to a dominance of men in STEM fields and the technology industry, the algorithms had learned that a probable candidate for Amazon is male.

Other forms of bias exist in online freelance marketplaces (e.g. TaskRabbit or Fiverr), which match employers with workers (Poster, 2019, p. 139). In these platforms, AI-based algorithms suggest a list of suitable workers to employers who want to tap the online pool of workers. Hannák et al. (2017) shows that these rankings can disadvantage Black women because of gender and racial biases. In effect, users who are perceived to be Black women are ranked lower than candidates from other social groups. They thus do not appear on the radar of the employers or they are considered less skilled. According to Hannák et al. (2017), this bias originates in the ratings and reviews from employers. Users who are read as Black and/or female get fewer and lower ratings, and written feedback contains fewer and less positive adjectives than for other groups (p. 1927).

These biases are especially damaging if we consider surging corporate calls for the inclusion of Black women in the digital economy. Big tech companies increasingly fund programs for minorities and girls that seem to address structural inequalities. These programs, of which the Google-funded *Black Girls Code* is a good example, educate and empower participants to become programmers (Noble, 2018, 64f). It is then a big irony if not insult that these *same* companies calling for Black women's inclusion produce and proliferate AI-powered recruitment tools that disadvantage Black women. It is also concerning that big tech companies target Black women for recruitment into an exploitative job market, which leads us to the second worrisome trend.

Contrary to scenarios where machines replace humans, human labor is still required for the development of AI-based technologies. In fact, the digital economy has produced a set of *new* jobs. This is what Gray and Suri (2019) call the "paradox of automation's last mile:" the more we use AI, the more we rely on humans (xxii). For instance, social media platforms rely on human intervention for "cleaning" the platform. This means that workers in low-wage jobs, working out of computer centers in the Global South, manually review images and videos to decide whether the content should be removed (Block & Riesewieck, 2018; Cherry, 2016, p. 71). From curating datasets to labelling images for an online shop, tech companies rely on "the power of the crowd" and distribute small tasks to thousands of workers via platforms such as Amazon Mechanical Turk[21] (Cherry, 2016, p. 76). These tasks help train machine learning algorithms and produce AI-based products.

Given the history of massive low-wage employment of immigrant and women of color in the technology industry (Gillard et al., 2010; Hossfeld, 1995, p. 408), these new precarious home-based jobs are likely held by the same group. They pay little taxable income and workers have little or no rights. Clients, however, can determine criteria for candidates, accept or reject the work product, and review the workers (Cherry, 2016, p. 76). Crowdsourcing jobs are still unregulated (Cahn et al., 2019, p. 8) and it remains unsolved whether workers count as "independent contractors" or "employees" (Cherry, 2016, p. 76). There are no benefits, no mentorship, and no remedy if technical matters prevent the worker from fulfilling a task (Cahn et al., 2019, p. 8; Gray & Suri, 2019, xxiv ff).

Finally, the current set-up of the digital economy follows a neoliberal capitalist framework (Zuboff, 2019), which tends to manifest global inequalities including "digital colonialism." First, the raw material (e.g., coltan) for digital technologies is extracted from conflict regions in the Congo with little consumer awareness of local human rights abuses (Noble, 2016). Second, digital technologies produced in Silicon Valley are exported to geographically and culturally different contexts without attention to local norms and values (Wakunuma & Masika, 2017). Third, workers in low-income regions of the Global South perform customer services for American consumers and are forced to fake a White identity from names to accents and greetings (Poster, 2019, p. 152). Hence, White privilege and supremacy are upheld at the expense of non-white lives and identities. Existing hierarchies become solidified. White supremacy is also at the heart of political projects of oppression.

[21] For further analysis of the Mechanical Turk, see Sinclair, this anthology [Hanemaayer, Ariane. 2022. Artificial Intelligence and Its Discontents, Springer.].

Political Oppression

AI-based innovations can lead to the political oppression of BIWOC. One source of discontent is the practice of AI-based surveillance. Historically, technologies have been used to enhance practices of "social sorting" (Lyon 2003). They help categorize and classify people in order to make decisions. In society, such classifications usually occur subconsciously and involve stereotypical ideas about other people (Goffman, 1990). We sort those we encounter into categories (male or female, White or Black) because we expect to know better whom we are dealing with. In a similar way, AI-based tools today "sort out" who is potentially dangerous (e.g., in law enforcement or criminal justice) and who is eligible for welfare benefits or a suitable candidate for college admissions, insurance, and loans (Angwin et al., 2016; O'Neil, 2016).

One prominent case study is facial recognition technology, which is used by law enforcement to identify suspects. In an intersectional study of facial recognition algorithms, Buolamwini and Gebru (2018) find that lighter and male faces are better recognized (or classified) than darker and female faces. The algorithms investigated in the study performed worse on black and brown women, often wrongly classifying them as male (Buolamwini & Gebru, 2018, p. 10). Given the use of the technology in policing, algorithmic bias may result in black and brown women being wrongfully arrested and convicted more often than whites. In the context of widespread police brutality against Black people in the United States (Hinton, 2021; Ritchie, 2017), algorithmic bias in facial recognition technology has serious implications for the free movement and safety of Black women.

Another source of concern is the treatment of race and gender in the design of AI. On the one hand, technology has been considered neutral to race and gender because of disembodiment in the digital sphere. Especially on the Internet – so the story goes – everyone can be and do what they want (Daniels, 2015, p. 1381). Such attitudes follow a color-blind worldview, which allows White people to ignore structural inequalities between Blacks and Whites by reference to universalism and equality ("I don't see race because we are all equal", Bonilla-Silva, 2018). Yet, color-blindness ignores the genealogy of cyberspace (e.g. ,dominance of a White male perspective) as well as the visual culture online (Daniels, 2009, p. 19, 2015, p. 1378). As a result of color-blind approaches, designers have paid little or no attention to racial inequalities in or through technology (Benjamin, 2019b, p. 63).

On the other hand, designers increasingly aim at personalizing services to different users and therefore consider race and gender as distinguishing features of users. Although this is where color-blindness apparently ends, designers still consider the *use* of gender and racial

categories neutral because technology is continuously seen as neutral. Consequently, categories of race and gender become legitimate tools of classification in AI-based technologies without attention to the structural inequalities that they represent and reinforce (Benjamin, 2019b, p. 21). For instance, in the field of Machine Learning Fairness, race and gender are widely used as variables to ensure an equal outcome of algorithmic decisions for diverse groups. While the goal seems like a step in the right direction, the use of race and gender as *descriptive* demographic categories obscures their historical and normative baggage and social construction. It ignores the way these categories have been used in political projects of oppression, for instance in the early U.S. census, where African-Americans were not counted properly to privilege White people (Hanna et al., 2020, p. 5). In fact, gender and race categories continue to be used as tools (technologies even) to manage access to political influence and restrict some groups' human rights, as the example of racialized voter targeting by the Trump campaign 2016 shows (Sabbagh, 2020).

A third aspect worth mentioning here is the proliferation of hate speech and white supremacist content with little regulation on behalf of social media companies and the government. While companies claim to prioritize racism online and may hire a Black spokesperson, even obvious racial aggression remains visible in social media networks (Benjamin, 2019b, p. 29). There are immediate and more subtle consequences of this lack of regulation. One is the translation of online hate into physical attacks against Black women (Daniels, 2009, p. 7). Another consequence is the slow erosion of accepted rights and political processes because white supremacists provide "alternative" interpretations of historical events such as slavery. Shielded by anonymity on the Internet, they question established values and ideals such as equality, insinuating that one can have an opinion on racism (as opposed to racism is prohibited by law and goes against democratic values, Daniels, 2009, p. 8).

Cultural Oppression

This section traces the oppression of Black women through AI-mediated representations. Symbolism and cultural imagination of identities play a crucial role in upholding or dismantling social hierarchies. Throughout American history and in contemporary media, racist and sexist stereotypes of Black women have geared to their perception as inferior (Gammage, 2017). AI systems can reinforce such harmful representations, for example through the racialized representation of robots.

Cave and Dihal (2020) show that imaginations of AI in films, drawings, and stories are overwhelmingly racialized as White. Humanoid robots and other machines (such as chatbots)

carry physical features that resemble a white face, and voice assistants use standard middle-class American English instead of African-American Vernacular English. The robot "Sophia" from Hanson Robotics is an example that enjoys big popularity in society and popular culture (Cave & Dihal, 2020, figure 1). This racialization of AI as White has implications. It can unconsciously convey the notion that technology design and development is only for White people. It can further reinforce racial hierarchies that position machines over Black people as much as White people over Black people (Cave & Dihal, 2020). A study by Bartneck et al. (2018) indeed shows that consumers attribute race to differently colored robots and exhibit bias towards those robots associated with Black male identities.

Bias in the representation of AI may be unconscious or intentional. Designers who are White may consider a pale skin color as neutral and normal, thus unconsciously choosing this design for their robotic creation (Cave & Dihal, 2020, p. 10). However, Benjamin (2019b) questions whether the representation of robots as White is so unintentional. She discusses an advertisement from the 1960s which paints the picture of a robot as a dehumanized servant, whereas the ad reads: "Slavery will be back" and "Slavery will be here to stay" (p. 57, figure 1.2). In the past leading to the present, imaginations of robots by White people often involved a desire to dehumanize and command inferior 'beings' (Benjamin, 2019b, p. 56). This resembles White attitudes during slavery, when Black women were forced to work not only in the cotton fields but also as servants in the household (hooks, 2015a, 24ff). Going a step further, the Whiteness of robots may also be an intentional move towards an imagined White utopia, in which BIWOC as domestic workers are entirely removed from the White family's home (Cave & Dihal, 2020, p. 14; Rhee, 2018, p. 94).

Another case of stereotyping of Black women through AI concerns search engines like Google. Noble (2018) investigates "the ways in which search engine results perpetuate particular narratives that reflect historically uneven distributions of power in society" (p. 71). Analyzing Google searches of the keyword "Black girls," she shows how the racialized sexualization and commodification of Black girls resembles broader narratives about Black women in U.S. society. In an initial search of the term in 2011, Noble (2018) notices that the websites at the top of the search results advertise (free) porn or sex with Black women (cf. figure 2.1, p. 67). Even dating sites represent Black women (or "girls") as hypersexualized. Unsurprisingly, Google searches for girls of *any* ethnicity generate similar search results (Noble, 2018, p. 71). However, given the history of Black women's representation in American media, their continued and exacerbated misrepresentation through AI is particularly concerning (Harris-Perry, 2013).

Noble (2018) briefly hints at the historical construction of Black women as Jezebels, Sapphires, and Mammies (p. 98). These images have their origin during the enslavement of African people. On the one hand, Black women were perceived as exotic and desirable sex objects that were particularly wanton compared to White women. They were the subject of gruesome sexual abuse by White slave owners and even after the end of slavery, White men widely raped Black women (hooks, 2015a, 58ff). On the other hand, Black women were represented as Mammies or matriarchs. The "iconic" Aunt Jemima pancake brand popularized the image of the old and unattractive Black woman who happily serves the White family (hooks, 2015a, 84f). As a matriarch, the Black woman is considered head of the own household, stripping her husband of his masculinity. Hence, as hooks (2015a, 80f) puts it:

> Just as whites used the myth that all black women were sexually loose as a way to devalue black womanhood, they used the matriarchy myth to impress upon the consciousness of all Americans that black women were masculinized, castrating, ball-busters (80f).

Some of these stereotypes are still used today, especially in movies and popular culture (West, 2018). While stereotypes in classic media contribute to the cultural oppression of Black women, the case of online search engines is a particular one. Search engines are gatekeepers of information (Bozdag, 2013, p. 209). They have the power to define and signify sometimes ambiguous concepts. Furthermore, search engines like Google enjoy the standing of a legitimate and neutral source of information, expressed by the saying "just Google it" (Noble, 2018, p. 85). They are even used in high schools and higher education for assignments or research. If we encounter biased information around a subject that is new to us, we may accept the information as truth and reality (Noble, 2018, p. 65). With the enormous reach of Google and YouTube, these services have the power to control representations of concepts and identities. This is concerning not just because of the concentration of power in the hands of a few (predominantly White and male) designers situated in Silicon Valley, but due to a capitalist framework. The representation of knowledge and identities is thus bound by private commercial interests rather than public democratic values (Noble, 2018, 91f).

AI & Resistance by Black Women

"The imagination is a tool of resistance. Creating stories with people of color in the future defies the norm. With the power of technology and emerging freedoms, black artists have more control over their image than ever before" (Womack, 2013, p. 24).

Resistance is an integral part of Black feminism and Black women's lives. Oppression cannot be thought without resistance, because Black women's experiences of oppression are

intrinsically linked with the desire to transform the unjust situation (Collins, 2000a, p. 9). Resistance here does not necessarily refer to the national mobilization of protestors. Simply surviving in an anti-Black and patriarchal system is a form of resistance (Collins, 2000a, p. 9). Similarly, intellectual labor can be a form of resistance (Spivak, 2018). Historically, Black women were exploited as free or cheap labor and lived in poverty, which left little or no time and resources to educate oneself (Collins, 2000a, p. 4). Another form of resistance is imagination of the future, which allows the oppressed to break free from the limiting possibilities of the present reality (Womack, 2013, p. 14). And of course, the tradition of Black women's struggles incorporates political resistance through protest à la Rosa Parks or the formation of movements. Most prominent is the contemporary Black Lives Matter movement, which is explicitly intersectional and shaped by queer Black women (Clark et al., 2018; Khan-Cullors & bandele, 2017).

In the context of AI, resistance takes many forms. During protests for democratic reform in Hong Kong in 2020, protestors sprayed the lenses of surveillance cameras, which fed data to AI-based face recognition systems used by police (Kantayya, 2020). Black feminist activists use AI-based services such as Twitter to campaign via hashtags for their rights, e.g., #BlackLivesMatter (Lumsden & Harmer, 2019, p. 10). Other forms of resistance develop new frameworks for AI-based platforms, e.g., "platform cooperativism," where a community owns and control everything that happens in the platform (Poster, 2019, 163f). Some use story-telling as a form of resistance, e.g., in Afrofuturism, a creative field of science fiction that cultivates visions of Black people's future with technology (Womack, 2013). Art is indeed increasingly leveraged to question and reimagine AI.[22] Due to scope limitations, in the following, I highlight political resistance and approaches to AI that emerge out of the academic community.

Political Activism

The film "Coded Bias" traces the political activism of researcher Joy Buolamwini against facial recognition technology in the USA (Kantayya, 2020). After Buolamwini discovered biases against Black women in facial recognition algorithms (Buolamwini & Gebru, 2018), she founded the Algorithmic Justice League and campaigned against the use of flawed software. In June 2019, Buolamwini testified in front of the U.S. House Committee on Science, Space and Technology about the dangers of algorithmic bias in face recognition. She demanded several actions to reduce algorithmic bias, among them incorporating ethics education in computer science curricula, diversifying AI development teams, auditing benchmarks and implementing

[22] see for instance the ai.assembly by Neta Bomani, https://aiassembly.io

fair data collection processes, prioritizing an intersectional perspective in the assessment of algorithmic harms, and public funding for company-independent AI development. She also suggests an accountability tax that AI companies should pay, of which the revenue would be redistributed to independent developers that serve public interests (Buolamwini, 2019b).

The Algorithmic Justice League has also engaged in awareness-raising by hosting workshops. One example is the *Drag vs. AI* workshop on 7 October 2020. Three drag queens hosted the one-hour program, first highlighting how facial recognition technology classifies faces according to heteronormative gender expectations and by exhibiting a white male bias. The creative part of the workshop taught participants acts of resistance. They were instructed to use make-up in order to confuse facial recognition technology and prevent the system from detecting their faces as their own.

But Black feminists not only dismiss harmful AI-based technologies. They also *leverage* AI-based technologies to campaign for their political interests. In these cases, activists use AI to their benefit such as generating visibility of Black feminist issues. A prominent example is hashtag activism via the platform Twitter (Jackson et al., 2020). While Twitter algorithms are not without bias (Trillò, 2018), they have successfully been used by Black feminists to create counterpublics, which constitute narratives or accounts of Black female experiences that would otherwise not have been picked up by mainstream news (Jackson et al., 2020, p. xxvii). On the one hand, these are hashtag campaigns that raise awareness of the violation of Black women's rights, e.g. #MeToo, #BlackLivesMatter, #SayHerName (S. Williams, 2016). On the other hand, hashtags are used to show solidarity and counter hate speech against Black women and girls in social media platforms, e.g. #StandWithJada (S. Williams, 2015).

New Frameworks for the Study of AI-based Technologies

Black feminists have also proposed new frameworks for the study of AI-based technologies that reveal intersecting discrimination or algorithmic bias. Daniels (2015) stresses the need to move away from the prevailing narrative that the Internet is a neutral space, free of racism and sexism. Instead, researchers and observers must acknowledge that race has been and still remains a factor in the structure of the Internet: whiteness shapes the Web, from its genealogy to design choices to content. Therefore, a lens that reveals this "white racial frame" (Feagin, 2013) should be adopted to uncover racism and bias online (Daniels, 2015, p. 1387).

Noble; Noble (2016, 2018) introduces another framework for the research of artificial intelligence and emerging technologies: Intersectional Black Feminist Technology Studies. A central component of this new method and framework is to produce counternarratives that

change the discourses about African Americans and women as technologically illiterate. Rather, it is important to study how Black people consume and influence technology beyond the Black community and how they shape current and future designs of AI-based products (Noble, 2018, p. 172). Benjamin (2019b) proposes "Race Critical Code Studies" that considers who has access to technologies but also how technologies are produced in the first place. At the heart of this research agenda are investigations of anti-blackness in technologies and how race itself was established and continues to work as a technology to regulate social relations and manifest White supremacy (Benjamin, 2019b, p. 45).

Technology Designs for Social Justice

Finally, Black feminist approaches and the concept intersectionality have inspired new design solutions. Costanza-Chock (2020) offers a paradigm change that de-colonizes design. Their approach to design removes the White male heterosexual norm from the design process and instead centers marginalized users. The guiding question for the design process is the following: how are certain users erased in or through technology? Following this reflection, technology should then be designed so that it dismantles systems of oppression (such as race, gender, sexuality, ability). This process requires an explicitly intersectional perspective instead of a single-axis analysis of technology's impact.

Intersectionality has gained traction in the computer sciences and especially in human-computer interaction (Kumar & Karusala, 2019). Erete et al. (2018) argue that designers need increased attention to the diversity of privilege and oppression to design for marginalized communities. The authors propose to consider three design principles: understanding and attending to context, reflecting one's own socio-economic and cultural context, and disclosing dissent or diverging perspectives within one project. Other works draw on the concept intersectionality to reflect on the representation of users and affordances for them in the technology (Schlesinger et al., 2017; Wong-Villacres et al., 2018).

However, Y. Rankin and Thomas (2019) criticize that works on intersectional design or intersectional human-computers interaction (HCI) ignore the tradition of Black feminist scholarship on intersectionality. The authors further call out the continuing exclusion and structural marginalization of black women in the field of human-computer interaction (also see J. O. Thomas et al., 2018). They point to existing work by black feminists in HCI that is rarely cited, and demand that HCI scholars build coalitions with black feminists.

Conclusion

"If we are not intentional about designing AI systems with equity in mind, we will replicate existing structural inequalities" (Buolamwini, 2019b).

This chapter has traced Black feminist perspectives on the design and use of AI. Black feminism provides a unique lens that renders visible the intersecting forms of oppression that Black women face in US society. Such intersecting oppressions are not just historically grown and exist in the physical world but further upheld by AI-based computer systems. The chapter reconstructed AI-based oppression in the economic, political, and cultural realms from exploitation in the digital economy to racist and sexist aggression in social media, and the reinforcement of biased search results. A Black feminist analysis of AI particularly points to *structural* oppression and highlights the danger of reinforcing structures of inequality due to pattern recognition.

Based on an intersectional analysis of AI's implications, Black feminist contributions developed strategies of resistance and new frameworks for AI that align with social justice goals. Activists and scholars have protested harmful AI products, demanding a moratorium on flawed technology. Among others, advocate Joy Buolamwini spoke before US Congress to campaign for a halt of the deployment of facial recognition technology. These activities aim at dismantling and transforming structural inequalities. Black feminist perspectives on AI thus propose radical innovation, rethinking the White and male centered foundations of technology design and automated systems. Race Critical Code Studies (Benjamin, 2019b) and Intersectional Black Feminist Technology Studies (Noble, 2018) along with justice-oriented design approaches (Costanza-Chock, 2020) offer new avenues for emerging technologies.

Chapter 3: Schwarzfeministische Perspektiven auf Künstliche Intelligenz: Erkenntnisse und neue Fragen zu KI-gestützter Gesichtserkennung und Überwachung

[This chapter was originally published in: Schelenz, Laura. 2021. "Schwarzfeministische Perspektiven Auf Künstliche Intelligenz: Erkenntnisse Und Neue Fragen Zu KI-Gestützter Gesichtserkennung Und Überwachung." *Femina Politica – Zeitschrift für feministische Politikwissenschaft* 30 (2-2021): 73–93. https://doi.org/10.3224/feminapolitica.v30i2.07, reproduced with permission from the Verlag Barbara Budrich]

Ob Pflegeroboter, Übersetzungssoftware, Risikoanalysen oder die Verteilung staatlicher Sozialhilfe, Künstliche Intelligenz (KI) wird mittlerweile zur Optimierung und automatischen Entscheidungsfindung in nahezu allen Bereichen des Lebens eingesetzt. Im Zuge dessen vermehren sich kritische Stimmen, die auf ethische Bedenken von KI hinweisen, insbesondere unter dem Stichwort AI Ethics (AI = Artificial Intelligence). Schwarze Feminist*innen bieten hier eine eigene, strukturelle Kritik an den risikobehafteten Implikationen von KI für das Leben von Schwarzen, indigenen und Frauen* of Color. So argumentieren einige Schwarze Feminist*innen, dass KI zur Unterdrückung von Schwarzen Frauen* beitrage, indem bestehende Systeme von Ungleichheit durch Verzerrungen in den Datensätzen, voreingenommene Designs und ausbeuterische Rahmenbedingungen in der Technikentwicklung verstärkt werden (Benjamin, 2019a, 2019b; Noble, 2018). Angesichts dieser einschlägigen Beiträge zur Erforschung von KI-Systemen stellt sich die Frage: Inwiefern produzieren Schwarzfeministische Ansätze neue Erkenntnisse und werfen Fragen rund um die ethische Entwicklung und Anwendung von KI auf? Das heißt, inwiefern ergänzen und entwickeln Schwarzfeministische Ansätze den bestehenden Diskurs weiter?

Geleitet von diesem Interesse beleuchtet der Artikel das Potenzial Schwarzfeministischer Instrumente zur Interpretation und Diskussion KI-gestützter Innovationen. Dabei steht das Thema Gesichtserkennung und Überwachung im Vordergrund. Schwarze Feminist*innen haben in den letzten Jahren eindringlich vor den Risiken von Gesichtserkennungstechnik gewarnt (Buolamwini und Gebru (2018)), aber auch die Praxis der Überwachung historisch einbettet (S. Browne, 2015). Im Folgenden werden drei zentrale Schwarzfeministische Instrumente herangezogen, um ihr Potenzial zum Verständnis und der Kritik von Gesichtserkennung und Überwachung herauszuarbeiten.

Das erste Analyseraster ist Unterdrückung und ist im Schwarzen Feminismus eng mit einer intersektionalen Analyse von Herrschaftssystemen verbunden (Collins, 2000a, p. 4; Crenshaw, 1989). Zweitens beziehe ich mich auf die Dialektik aus Hypervisibility/Invisibility, welche

ebenso wie Unterdrückung die Lebensrealität Schwarzer Frauen* prägt. Drittens betrachte ich Formen des Widerstands, welcher integraler Bestandteil Schwarzfeministischer Theorie und Praxis ist (Collins, 2000a, p. 9). Der Artikel argumentiert, dass sich mithilfe dieser drei Analyseraster ergänzend zur Mainstream-Debatte um AI Ethics neue Erkenntnisse und Diskussionspunkte im Hinblick auf KI-gestützte Gesichtserkennung und Überwachung generieren lassen. Beispielsweise werfen sie provokante Fragen auf, die in der AI Ethics-Debatte bislang wenig Anklang fanden: Kann eine Technik wie die Gesichtserkennung ethisch entwickelt und eingesetzt werden? Müssen nicht die Logiken der Überwachung selbst überdacht werden?

Der Artikel nutzt eine US-amerikanischen Strömung des Schwarzen Feminismus. Die Beispiele der Gesichtserkennung und Überwachung stammen ebenfalls aus den Vereinigten Staaten und sollten in der deutschsprachigen Debatte stärker thematisiert werden, denn Überwachungsdynamiken mithilfe von KI finden auch in Deutschland Anwendung (Biselli, 2020). Hier bieten Schwarzfeministische Ansätze eine Vorlage für die Analyse und Kritik der in Deutschland verwendeten Techniken, welche ebenfalls zur Unterdrückung marginalisierter Gruppen beitragen (ebd.). Der Artikel analysiert Schwarzfeministische Konzepte aus einer Weißen, Europäischen Perspektive, weshalb meine Interpretation Schwarzfeministischer Konzepte womöglich Einschränkungen mit sich bringt.

Schwarzfeministische Instrumente: Theoretischer und Methodischer Rahmen

Drei Analyserahmen sollen exemplarisch herangezogen werden, um die Potenziale Schwarzfeministischer Theorie und Praxis im Hinblick auf die Debatte um KI-gestützte Überwachung zu eruieren.

Unterdrückung: Laut Patricia Hill Collins bezeichnet Unterdrückung eine zeitlich ausgedehnte ungerechtfertigte Unterordnung einer Gruppe unter eine andere sowie eingeschränkte Zugänge zu gesellschaftlichen Ressourcen:

> "Oppression describes any unjust situation where, systematically and over a long period of time, one group denies another group access to the resources of society. Race, class, gender, sexuality, nation, age, and ethnicity among others constitute major forms of oppression in the United States" (Collins, 2000a, p. 4).

Die tatsächliche Ausprägung einer Unterdrückung ist von Kontext zu Kontext verschieden, da sie an das jeweilige Gesellschaftssystem eines Landes bzw. einer Region angeschlossen ist (Collins, 2000a, 227f). Intersektionalität bietet das methodische Werkzeug, um Unterdrückung in ihrer konvergierenden Natur sichtbar zu machen und zu kritisieren. In ihrem einflussreichen Artikel postuliert Kimberlé Crenshaw (1989), dass Schwarze Frauen* aufgrund ihrer

strukturellen Position an der Schnittstelle verschiedener Unterdrückungssysteme (race, gender, class) besondere Formen der Diskriminierung erleben (142).

Hypervisibility/Invisbility: Dies bezeichnet eine Dialektik aus (unerwünschter) Sichtbarkeit Schwarzer Frauen* bei gleichzeitiger Marginalisierung. In einer rassistisch und sexistisch geprägten Gesellschaft sind Schwarze Frauen* häufig als ‚the Other' den prüfenden Blicken staatlicher Institutionen (Yuval-Davis, 2011, 48f), aber auch Weißer Menschen im Beruf oder im sozialen Kontext ausgesetzt (Sue et al., 2007). Schwarze Frauen* sind somit im negativen Sinne ‚hypervisible,' denn sie werden stetig überwacht und kontrolliert. Diese Sichtbarkeit bedeutet jedoch nicht, dass eine Schwarze Frau* tatsächlich im politischen Sinne gesehen wird. Wenn es um die politische Partizipation oder im beruflichen Kontext um Expertise geht, werden Schwarze Frauen* zumeist ignoriert und marginalisiert. Ein Beispiel ist die Einbindung Schwarzer Frauen* in Diversitäts- oder Gleichstellungsaufgaben. Hier wird meist die einzige Schwarze Frau* eines Unternehmens für unvergütete Diversitätsarbeit rekrutiert, aber verunglimpft, wenn sie strukturelle Machtverhältnisse oder Rassismus anspricht (S. Ahmed, 2009, 48f).

Widerstand: Im Schwarzen Feminismus kann Unterdrückung nicht ohne Widerstand gedacht werden. Die oft gewaltsamen und diskriminierenden Erfahrungen Schwarzer Frauen* sind eng mit ihrem Bestreben verbunden, ungerechte Verhältnisse zu transformieren (Collins, 2000a, p. 9). Dabei kann einfaches Überleben in einem anti-schwarzen, patriarchalen und klassenbasierten System bereits eine Form des Widerstands sein (ebd.). In ähnlicher Weise kann intellektuelle Arbeit eine Form des Widerstands sein (Spivak, 2018), denn Schwarzen Frauen* standen aufgrund von Versklavung und Ausbeutung wenig Zeit und Ressourcen für ihre Bildung zur Verfügung (Collins, 2000a, p. 4). Eine weitere Form des Widerstands ist die Visualisierung einer alternativen Zukunft, wie etwa im Afrofuturismus (Womack, 2013). Zuletzt gehört zur Tradition des Schwarzfeministischen Widerstands natürlich der politische Protest. Ein Beispiel ist die Black Lives Matter-Bewegung, die sich in vielen Teilen als explizit intersektional und queer-feministisch definiert (Clark, Dantzler und Nickels (2018)).

Die Einführung der theoretischen und praktischen Instrumente zeigt, dass es vielseitige Formen der Anwendung dieser Analyserahmen gibt. In Anwendung auf ein Themenfeld (z.B. KI) könnten also verschiedene Blickwinkel angelegt werden. Um den Beitrag weiter einzugrenzen, wird es im Folgenden bei der Untersuchung von KI vorrangig um politische Aspekte wie der Unterdrückung politischer Teilhabe und racial profiling gehen.

Ethische Bedenken zu Künstlicher Intelligenz

In den letzten Jahren haben sich zunehmend kritische Stimmen zum Design, zur Anwendung und zu den Rahmenbedingungen von KI-Entwicklungen geäußert. Teilweise liegen die Probleme der Technik in ihrer Beschaffenheit. KI umfasst eine Reihe von Methoden und Anwendungen in der Informatik, die darauf abzielen, menschliche kognitive Fähigkeiten zur Lösung komplexer Probleme zu simulieren (Misselhorn, 2018, p. 161). KI hat zwei zentrale Aufgaben: (1) die Analyse von großen Datenmengen (Big Data) und (2) die Modellierung von Zukunft, unter anderem Vorhersagen über das menschliche Verhalten (Malik, 2019). Für beide Aufgaben greift die Technik auf Mustererkennung zurück, was schließlich auch die Grundlage für automatisierte Entscheidungen durch KI liefert (Mittelstadt et al., 2016, 3). Bei der Mustererkennung ist KI zwar effizienter als der Mensch, bringt aber einige Risiken mit sich. Wird ein Computermodell etwa mit einer verzerrten Datenbank trainiert, dann reproduziert das Modell den eingespeisten Bias (Malik, 2019). Häufig arbeiten Unternehmen mit öffentlich zugänglichen Datenbanken (etwa im Internet verfügbare Text- oder Bildquellen), die in einer rassistisch und sexistisch geprägten kapitalistischen Gesellschaft zumeist a) die Informationen von Minderheiten und marginalisierten Gruppen exkludieren oder b) Stereotype reproduzieren. Hier kommt es dann zum Algorithmic Bias (Zou und Schiebinger (2018)), welcher zu unfairen Entscheidungen bei der Vergabe von Jobs, Krediten, medizinischen Behandlungen oder sogar Risikoeinschätzungen zur Entlassung von Straftäter*innen führt (O'Neil, 2016).

Ein klassisches Beispiel von Algorithmic Bias ist der Fall von Software zur Früherkennung von Hautkrebs. Der Einsatz der Technik kann Leben retten, wirkt in der Praxis allerdings benachteiligend für Menschen mit dunkler Hautfarbe (Lashbrook, 2018). Das liegt daran, dass der Algorithmus auf einer unausgeglichenen Datenlage, nämlich überwiegend mit Bildquellen von heller Haut, trainiert wurde und damit Anzeichen von Hautveränderungen auf dunkler Haut schlecht erkennt. Dies hat reale Konsequenzen für Schwarze Menschen, da aufgrund rassistisch geprägter medizinischer Ungleichbehandlung ohnehin weniger Schwarze als Weiße Patient*innen eine Hautkrebserkrankung überleben (Lashbrook, 2018). Die AI Ethics-Gemeinschaft weist hier auf die Notwendigkeit hin, Datensätze von Grund auf inklusiver zu bauen und detaillierter zu annotieren, z.B. die Bilddaten mit qualifizierenden Kategorien (race, gender, etc.) zu versehen (Zou und Schiebinger (2018)). Ein weiterer Vorschlag ist das sogenannte Auditing von Algorithmen, bei denen KI eingesetzt wird, um Bias in KI-gestützter Software zu identifizieren (ebd.).

Manche sehen die Problematik jedoch tieferliegend. Laut Ayanna Howard und Jason Borenstein (2018) reflektiert algorithmischer Bias bzw. Bias in Datensätzen die Vorurteile, die wir seit unserer Kindheit internalisiert haben (1530). Ruha Benjamin (2019b) zeigt, dass

Technik ähnlich menschlicher Assoziationen Daten mit teils stereotypen Profilen verbindet. Deutlich wird das am Beispiel von Vor- und Nachnamen: Ob ‚Alice Tang' oder ‚Tyrone Jackson', Namen suggerieren in diesem Fall, Informationen über die Geschichte und das soziale Umfeld einer Person zu liefern (ebd., 6). Nicht nur erkennt und reproduziert eine Maschine solche Assoziationen, sondern sie legitimiert diese Informationen auch noch als wahr (ebd., 21).

Nach wie vor wird häufig angenommen, dass KI neutraler sei als das menschliche Analyse- und Urteilsvermögen (Wachter-Boettcher, 2017, 14f). Diese Haltung ist im Bereich des kommerziellen Technikdesigns weit verbreitet und verbunden mit vorherrschenden Werten wie Meritokratie und Gleichheit. Hierzu gehört auch, dass Designer*innen bewusst Unterschiede aufgrund der Hautfarbe ignorieren, und zwar mit dem Anspruch ‚für alle' Nutzer*innen designen zu wollen. Jessie Daniels (2015) spricht von ‚Colorblind Design' und argumentiert, dass eine solche Herangehensweise historisch gewachsene Ungleichheiten verschleiert, wie unter anderem die Tatsache, dass digitale Innovationen in einem Weiß und männlich dominierten Kontext in den USA gewachsen sind. Über die Werthaltung hinaus trägt auch der neoliberale wirtschaftliche Rahmen dazu bei, dass Diskriminierung durch KI-Systeme unzureichend adressiert wird. Im Zuge des Surveillance Capitalism können Daten ungebremst fließen, und große IT-Unternehmen verfügen über enorme Gestaltungsmacht (Zuboff, 2019). Global gesehen führen die genannten Rahmenbedingungen zu einer Verfestigung bestehender Machtrelationen zugunsten westlicher Akteure, was auch unter dem Begriff Digitaler Kolonialismus diskutiert wird (Coleman, 2019; Tuzcu, 2020).

Unterdrückung durch KI-gestützte Gesichtserkennung und Überwachung

Um tiefer in die Diskussion einzusteigen, fokussiere ich im Folgenden auf KI-gestützte Gesichtserkennung. Diese Technik wird mittlerweile an vielen öffentlichen Plätzen in den USA, in Großbritannien, aber auch in Deutschland eingesetzt. Am Flughafen wird die biometrische Gesichtserkennung zum automatischen Abgleich von Pässen verwendet (MONITOR, 2020). An Bahnhöfen oder in Fußgängerzonen wird schon lange Kameraüberwachung (CCTV) eingesetzt – zunehmend werden hier auch Gesichtserkennungsfunktionen integriert. In der Londoner Innenstadt testet die Polizei KI-gestützte Gesichtserkennung zur Prävention von Kriminalität (Kantayya, 2020). Weit verbreitet ist der Einsatz von Gesichtserkennungssoftware bei Strafverfolgungsbehörden. Diese greifen

unter anderem auf private Firmen wie Clearview zurück, die öffentlich zugängliche Fotos von Gesichtern aus sozialen Medien herausgefiltert und zu einer Datenbank zusammengeführt haben. Strafverfolgungsbehörden gleichen dann mithilfe dieser Datenbank Fotos von Straftäter*innen mit den in der Datenbank hinterlegten Bildern ab. Zahlreiche Beschwerden laufen gegen Clearview, da das Unternehmen die Einwilligung der Rechtsinhaber*innen der Bilder nicht einholte (ACLU, 2020; Kantayya, 2020). Während eine solche Missachtung des Rechts auf Privatheit all diejenigen betrifft, die Fotos in sozialen Netzwerken hinterlegt haben, kann sie für marginalisierte Gruppen besonders gravierenden Konsequenzen haben.

Zum einen werden Schwarze Frauen* häufiger zu Unrecht aufgrund der verzerrten Software als Straftäter*innen identifiziert und verhört oder verhaftet. Als prominentes Beispiel wird im Diskurs häufig der Fall von Robert Julian-Borchak Williams angeführt. Es ist der erste bekannte Fall, bei dem ein Afro-Amerikaner aufgrund eines fehlerhaften Abgleichs mit einer Bilddatenbank verhaftet wurde. Bei seinem Verhör sah er das Foto des gesuchten Straftäters, hielt es neben sein Gesicht und prangerte den eindeutigen Fehler an: „You think all Black men look alike?“ (Hill, 2020). Während der prominente Fall einen Mann betrifft, wird im Diskurs wenig auf die Erkenntnisse der Studie „Gender Studies“ von Joy Buolamwini und Timnit Gebru (2018) verwiesen, die eine schlechtere Performanz von Gesichtserkennungsalgorithmen für Schwarze Frauen* als für Männer konstatierten. Die untersuchten Algorithmen von Microsoft, IBM und Face++ wiesen eine signifikant schlechtere Performanz bei Frauen* mit dunkler Hautfarbe auf und klassifizierten sie oft fälschlicherweise als männlich (Buolamwini und Gebru (2018, p. 10)). Essenziell für den Erkenntnisgewinn über die Fehlerraten der Algorithmen war die intersektionale Analyse. Buolamwini und Gebru bildeten vier Datengruppen aus den Bildquellen: „darker females, darker males, lighter females, and lighter males“ (ebd., 2). Diese intersektional gedachte Aufteilung in Subgruppen ermöglichte ein Verständnis der Performanz von Gesichtserkennungssoftware für Schwarze Frauen* und konnte so bislang vernachlässigte Defizite der Technik aufdecken. Während die Defizite zwar vor allem die Fähigkeit der Software betreffen, Frauen als ‚Frau‘ zu klassifizieren, lassen die Erkenntnisse dennoch Schlüsse über die Fähigkeit von Software zu, Bilder mit denen in einer Datenbank abzugleichen.

Zum anderen kann sich das schlichte Aufzeichnen der Bilddaten und die Kapazität des KI-gestützten Auswertens durch Behörden auf die Bereitschaft von Minderheiten auswirken, an öffentlichen Veranstaltungen wie einer Demonstration teilzunehmen. Denn selbst wer friedlich demonstriert, muss damit rechnen, dass die eigenen Bildquellen in große Datensätze gespeist werden und das Bild zum Abgleich von Fahndungsfotos verwendet wird (MONITOR, 2020).

Gepaart mit einem höheren Risiko der KI-bedingten Fehlidentifizierung für Minderheiten vermag die sicherheitsbehördliche Überwachung von Protesten noch größere Bedenken über die Teilnahme an politischen Aktivitäten auszulösen. In den USA spielt hier historisch gewachsene, strukturelle Unterdrückung eine Rolle, denn aufgrund weit verbreiteter Polizeibrutalität gegenüber Schwarzen Frauen* (Ritchie, 2017) kann Profiling über Gesichtserkennung zu einer Lebensbedrohung für Schwarze Frauen* werden.

Historische Einbettung: Hypervisibility/Invisibility von der Versklavung bis zur Bürgerrechtsbewegung

Eine Einbettung der genannten Defizite in die historischen und gegenwärtigen sozialen Kontexte zeigt die strukturelle Natur dieser (algorithmischen) Diskriminierung auf. Überwachung durch Techniken des visuellen ‚Erkennens' Schwarzer Menschen ist in die Zeit des transatlantischen Sklavenhandels zurückzuführen. Bereits im 18. Jahrhundert wurden in den USA Schwarze Menschen (zumeist Versklavte) mit damals vorhandenen Techniken wie Laternengesetze, Anzeigen über geflüchtete versklavte Menschen oder Passsysteme für Weiße Menschen sichtbar und identifizierbar gemacht (S. Browne, 2015, p. 66). Dabei handelt es sich jedoch nicht um eine Sichtbarkeit in der politischen Öffentlichkeit im Sinne der demokratischen Partizipation, sondern um restriktive Maßnahmen, die der systematischen Unterdrückung Schwarzer Menschen im Rahmen der Plantagenökonomie dienten. Überwachung durch Technik ist also historisch betrachtet inhärent mit der Kontrolle von Schwarzen Menschen verbunden (S. Browne, 2015, 32; 50).

Simone Browne beschreibt in Dark Matters: On the Surveillance of Blackness einige der zu Zeiten der Sklaverei verwendeten Techniken: Ein Beispiel sind die oben erwähnten Laternengesetze. Hierbei handelt es sich um die Anordnung, dass Schwarze Menschen während der Dunkelheit eine Kerze oder eine Laterne mit sich tragen müssten, um stets im Dunkeln sichtbar zu sein (ebd., 67). Somit sollte die Mobilität Schwarzer Menschen nur unter Beobachtung Weißer Menschen möglich sein. Andere Techniken zur Überwachung Schwarzer Menschen waren geschriebene Dokumente, wie etwa die Bekanntmachung der Flucht von Versklavten in Zeitungen, die Produktion von Pässen, ohne die sich Versklavte nicht einmal auf der Plantage frei bewegen durften, und offizielle Listen mit Namen von ehemaligen Versklavten (The Book of Negroes, ebd., 68) In diesen Dokumenten wurden Schwarze Menschen detailliert beschrieben, nicht nur ihre Gesichtszüge und Hautveränderungen, sondern auch Kleidung und Charakterzüge (ebd., 54).

Dieser kurze Rückblick in die US-amerikanische Geschichte zeigt, wie eng die Überwachung Schwarzer Menschen mit ihrer Unterdrückung durch Weiße verwoben ist. Die heutige Fortschreibung von Racial Profiling (ACLU, 2021) zeigt, dass rassifizierte Überwachung als Vermächtnis aus der Sklaverei nach wie vor Anwendung findet. Da also Schwarze Frauen* historisch betrachtet die Subjekte rassistischer Überwachungstechniken sind, scheint es überraschend, dass heutige Gesichtserkennungssoftware Schwarze Frauen* schlechter zu identifizieren vermag als Weiße. Wie lässt sich dieser scheinbare Widerspruch erklären? Hier kommen wir zu der Ambivalenz von Hypervisibilität bei gleichzeitiger Unsichtbarkeit. Während Schwarze Frauen* als die ‚Anderen' in einer Weiß dominierten Gesellschaft stets überaus sichtbar sind und ihr Verhalten von kritischen Blicken begleitet wird, fehlt ihnen die Anerkennung als politische und gesellschaftliche Subjekte (S. Browne, 2015, p. 57).

Die US-amerikanische Gesellschaft ist um die Überlegenheit von Weißsein, Männlichkeit und Heterosexualität herum strukturiert (Hull, Bell Scott und Smith (2010)). Dies führt dazu, dass Weißsein, Männlichkeit und Heterosexualität als ‚normal' wahrgenommen und damit für Angehörige dieser Gruppe unsichtbar wird (Eggers et al., 2017). ‚Anders' auszusehen oder sich ‚anders' zu verhalten, wird damit überaus sichtbar in öffentlichen Räumen, gleichzeitig jedoch als ‚nicht normal' abgewertet. Für Schwarze Frauen* bedeutet diese Dehumanisierung unter anderem eine Nichtbeachtung ihrer Bedürfnisse, Interessen und Rechte. Schwarze Frauen* wurden in der Geschichte durchgängig missachtet, zum Beispiel in der Weiß und bürgerlich geprägten Frauenrechtsbewegung zu Beginn des 20. Jahrhunderts und in der männlich dominierten Bürgerrechtsbewegung der 1960er-Jahre (hooks, 2015a, 3ff). Diese Unsichtbarkeit und Marginalisierung Schwarzer Frauen ist ein wichtiger Motivator für den schwarzen Feminismus (Collins, 2000a, p. 9).

Für den Fall der Gesichtserkennungssoftware bedeutet die Ambivalenz von Hypervisibility/Invisibility nun, dass Schwarze Frauen* aufgrund struktureller Unterdrückung zu den Zielpersonen einer KI-gestützten Überwachung zählen und die Technik in Kontexten eingesetzt wird, in denen Schwarze Menschen unter Generalverdacht stehen – in der Stadt, bei Protesten, am Flughafen. Während der Black Lives Matter Proteste in New York City im August 2020 setzte die Polizei beispielsweise Gesichtserkennungssoftware ein, um Demonstrant*innen im Nachhinein wegen Vergehen bei den Protesten aufzusuchen (Vincent, 2020). Gleichzeitig funktioniert die Technik jedoch nicht richtig aufgrund einer verzerrten Datenbasis, welche auf die Unsichtbarkeit Schwarzer Frauen* und damit einhergehend auf einen Mangel an Informationen über sie zurückzuführen ist (siehe Algorithmic Bias weiter oben, Lashbrook (2018)). Die hohe Fehlerrate bei der Gesichtserkennung Schwarzer Frauen*

in Kombination mit dem Einsatz der Technik zur Profilbildung und Strafverfolgung von Schwarzen Menschen deuten auf etablierte, strukturelle Unterdrückungsdynamiken hin, die eine Schwarzfeministische Perspektive hervorhebt.

Schwarzfeministischer Widerstand gegen Überwachung und Gesichtserkennung

Basierend auf der Erkenntnis, dass KI-gestützte Gesichtserkennung strukturelle Unterdrückung reproduziert, haben Schwarze Feminist*innen Widerstand gegen die Technik geleistet. Nach der Studie „Gender Shades" gründete Joy Buolamwini die Algorithmic Justice League (AJL) und stellte sich gegen den Einsatz von Gesichtserkennungssoftware in ihrer gegenwärtigen Form. Die Dokumentation „Coded Bias" verfolgt ihren politischen Aktivismus in den USA (Kantayya, 2020). Zu dem Zeitpunkt mobilisierten auch andere Kräfte gegen den Einsatz von KI-gestützter Gesichtserkennung (ACLU, 2020). Erste Erfolge von öffentlichkeitswirksamen Kampagnen gegen Gesichtserkennung wurden in San Francisco, Oakland und Somerville in den USA verzeichnet. Hier haben die Städte den Einsatz der Technik gebremst. Gegenwärtig wirbt die Kampagne #banthescan der Menschenrechtsorganisation Amnesty International für ein Verbot der Software in New York City. Im Juni 2019 sagte Buolamwini vor dem U.S. House Committee on Science, Space and Technology über die Gefahren von algorithmischer Verzerrung in Gesichtserkennungssystemen aus (Buolamwini, 2019b). Während diese Aktivitäten hochrangige politische Akteure adressieren, gibt es weitere bottom-up Aktivitäten Schwarzer Feminist*innen gegen Gesichtserkennungssoftware.

Ein Beispiel ist der Drag vs. AI-Workshop am 7. Oktober 2020, online veranstaltet von der AJL. Drei Drag Queens führten durch das einstündige Programm und zeigten zunächst auf, wie Gesichtserkennungstechnologie Gesichter nach heteronormativen Geschlechtererwartungen klassifiziert und dabei einen Weißen männlichen Bias aufweist. Im praktischen Teil des Workshops lernten die Teilnehmer*innen ‚drag technologies' zu nutzen, um die Gesichtserkennungstechnik zu verwirren. In anderen Worten lernten die Teilnehmenden, sich so zu schminken und zu verkleiden, dass die Technik ihre Gesichter entweder nicht erfassen kann (z.B. über gemalte Schatten und schwarzweiß Kontraste im Gesicht) oder Geschlecht und Alter fehlidentifiziert (z.B. Falten schminken, mit Geschlechternormen spielen).

Diese Praxis konstituiert das, was S. Browne (2015, 21f) als „dark sousveillance" bezeichnet. Personen, die Subjekte der Überwachung sind, beobachten ‚von unten' und versuchen, die Überwachungstechnik zu verstehen, um sie zu unterlaufen und sich Freiheiten zu verschaffen. Insbesondere das Make-Up Tutorial des Drag vs. AI-Workshops erinnert an Praktiken von Schwarzen während der Versklavung. Wegen der One Drop-Regel galten Menschen, die eine*n

Schwarze*n und Weiße*n Vorfahre*n hatten als Versklavte. Sogenannte ‚quadroons,' also Menschen, die nur zu einem Viertel afrikanischer Abstammung waren, wurden demnach als Versklavte betrachtet, konnten sich aber relativ frei bewegen, da sie als Weiß gelesen wurden. Make-up und Kleidungsstil halfen dabei, sich als Weiß zu ‚verkaufen' und somit dem Status als versklavter Person zu entkommen (ebd., 54).

Hier zeigt sich, dass ‚Unsichtbarkeit' in einem Gewaltstaat eine Überlebensstrategie sein kann. So werfen die Arbeiten von S. Browne (2015) und die Aktivitäten der AJL die Frage auf, ob es im Kontext von KI-gestützter Gesichtserkennung nicht sinnvoller ist, die Daten von Minderheiten nicht einzuspeisen. Zwar würde eine weiterhin unausgeglichene Datenlage zu Fehlidentifikationen führen, aber eine Optimierung der Technik zur Überwachung und Kontrolle von Schwarzen Frauen* würde verlangsamt. Gedanken der Subversion von destruktiven Techniken finden sich vor allem in der abolitionistischen Argumentation wieder, die im Folgenden noch näher beleuchtet wird.

KI-gestützte Gesichtserkennung zwischen Reform und Abolition

In den vorangegangenen Teilen habe ich gezeigt, wie Schwarzfeministische Analyserahmen einen produktiven Beitrag zur Diskussion um KI-gestützte Gesichtserkennung und Überwachung leisten. Zum einen zeigen sie neue Erkenntnisse auf. Dass Schwarze Frauen* nicht korrekt von Gesichtserkennungssoftware erkannt werden, lässt sich nur mithilfe einer intersektionalen Analyse aufdecken (vgl. Buolamwini und Gebru (2018)). Gleichzeitig bieten Schwarzfeministische Konzepte die Werkzeuge, um Sachverhalte historisch einzubetten und entsprechende Widerstände zu entwickeln. Als zentrale Erkenntnis lässt sich festhalten, dass „Algorithmic Bias" strukturell gedacht werden muss und dass einfache technische Lösungen keine Antwort auf Unterdrückung sein können (Noble, 2018, 6f). Diese Perspektiven ergänzen den Mainstream-Diskurs im Bereich AI Ethics und zeigen gleichzeitig dessen Grenzen auf. Wenn es bei AI Ethics um die Qualität von Datensätzen und die Schaffung von Transparenz, Erklärbarkeit und Fairness geht, dann werden dabei immer noch nicht die zugrundeliegenden Diskriminierungsstrukturen von Technik beleuchtet (vgl. die Diskussion von Kind et al., 2021). Im Falle von KI-gestützter Gesichtserkennung und Überwachung haben Schwarze Feminist*innen daher eine Reihe von Antworten entwickelt, die ein Kontinuum von Reform bis hin zu abolitionistischen Ansätzen darstellen. An einem Ende des Kontinuums steht die Arbeit von Informatiker*innen wie Joy Buolamwini und Timnit Gebru. Sie erarbeiten Möglichkeiten, Datensätze aufzubereiten und Computermodelle mit dem Ziel einer möglichst hohen Genauigkeit bei der Gesichtserkennung zu trainieren (vgl. den Datensatz „Gender

Shades“ von Buolamwini und Gebru (2018)). Dabei dienen zeitweilige Moratorien dazu, Fehlerquoten sichtbar und nachvollziehbar zu machen bzw. zu minimieren. Vor dem Hintergrund, dass die Technik trotz bekannter Gefahren Anwendung findet, sind solche Reform-orientierten Ansätze hilfreich, denn eine technische Verbesserung der Software könnte die Gefahr von Fehlidentifikationen für Schwarze Frauen reduzieren.

Am anderen Ende des Kontinuums stehen abolitionistische Ansätze, also solche, die eine grundsätzliche Abschaffung von Überwachungstechniken fordern. Hier argumentieren Schwarze Feminist*innen, dass die Logik von Überwachung und damit einhergehende Datensammlung selbst inhärent rassistisch sei. Auch fehlerfreie Technik würde die Entfaltung Schwarzer Menschen beeinträchtigen, da sie sie immer in einem westlichen, Weißen, heteronormativen und neoliberal ausgerichteten Normensystem gefangen gehalten würde (Benjamin, 2019a, 2f). Diesem Argument folgend ist es für Schwarze Frauen* nicht wünschenswert, in ein System integriert zu werden (zum Beispiel über das Einspeisen von Daten), welches der Unterdrückung von Schwarzen Frauen* dient. Anna Laureen Hoffmann (2020) bezeichnet den Diskurs um die Integration von Minderheiten in Datensätze als „diskursive Gewalt“: Positiv konnotierte Begriffe wie Diversität oder Inklusion würden benutzt, um Minderheiten in ein für sie gewaltsames System zu integrieren (ebd., 11).

Abolitionistische Ansätze verfolgen also entgegen der Integration von Minderheiten in Datensysteme ein radikales Umdenken des Verhältnisses von Technik und Gesellschaft. Hier bietet der Afrofuturismus einen Ort für den Ausdruck Schwarzfeministischer Zukunftsvisionen. Technosoziale Imaginationen der Zukunft bieten Schwarzen Autor*innen und Künstler*innen die Möglichkeit, neue Gesellschaftsmodelle und -verhältnisse zu denken. Die Mitbegründer*in des Afrofuturismus Ytasha Womack (2013, p. 24) bezeichnet dies als Widerstand:

> The imagination is a tool of resistance. Creating stories with people of color in the future defies the norm. With the power of technology and emerging freedoms, black artists have more control over their image than ever before.

Die technischen Möglichkeiten, von denen Womack hier spricht, nutzen beispielsweise die Künstler*innen und Aktivist*innen der Hyphen-Labs. Sie entwickeln unter dem Namen Neurospeculative Afrofeminism Methoden, um rassistische und sexistische Denkweisen, die Schwarze Frauen* über sich internalisiert haben, aufzubrechen. In einem Video suggerieren sie, dass über Therapie-ähnliche Sitzungen mit Virtual Reality Brillen Schwarzfeministisches Wissen oder futuristische Fiktionen von Freiheit vermittelt werden können. Auch über technisch innovative braids ließe sich in dieser Vision das Gehirn zum Empowerment von Frauen* stimulieren (Benjamin, 2019a, p. 11; Hyphen-Labs, 2019). Hier verbindet sich dann Schwarzfeministischer Widerstand mit der Nutzung von KI zur Befreiung Schwarzer Frauen*.

Ausblick: Lehren für die Diskussion um KI in Deutschland

Auch in Deutschland wird Gesichtserkennungstechnik hergestellt (siehe Cognitec in Dresden) und von Sicherheitsbehörden am Flughafen oder zur Strafverfolgung eingesetzt (MONITOR, 2020). Zwar gibt es in Deutschland andere Ausprägungen von Unterdrückung als in den USA, aber die zugrundeliegenden Logiken (wie etwa Weiße Vorherrschaft) sind vergleichbar und gehen ursprünglich vom europäischen Kontinent aus (siehe Kolonisierung und Versklavung auf dem afrikanischen Kontinent sowie Transatlantischer Sklavenhandel). So lässt sich vermuten, dass sich auch in Deutschland die Unterdrückung von marginalisierten Gruppen wie etwa Migrant*innen und Geflüchteten, Rom*nja, Sinti*zze und Jüd*innen durch KI-gestützte Überwachung reproduziert und intensiviert (vgl. Biselli und Meister (2019)). Ein Beispiel ist die Spracherkennungssoftware, die das Bundesamt für Migration und Flüchtlinge (BAMF) bei der Verifizierung von Dialekten und somit von Fluchtgeschichten einsetzt. Hier stellen sich nicht nur Fragen zur ethischen Anwendung angesichts fehlerhafter Technik, sondern Schwarzfeministische abolitionistische Argumente finden Anklang: Selbst der Einsatz einwandfreier Software bliebe doch ein „Automatisiertes Misstrauen“ (Biselli, 2020) und erschwert bis verunmöglicht das Wahrnehmen der Rechte auf Asyl und Freizügigkeit. Im Juni 2021 haben Datenschutz-Behörden der EU das Verbot von Gesichtserkennungssoftware in öffentlichen Räumen gefordert (European Data Protection Board [EDPB], 2021). Vor dem Hintergrund der anstehenden EU-Regulierung von KI-Systemen ist dies ein wichtiger Schritt. Es bleibt allerdings offen, ob damit auch ein grundsätzliches Umdenken der Beziehung von Gesellschaft und Technik sowie gesellschaftlicher Unterdrückungssysteme einhergeht.

PART II:

STUDYING DIVERSITY *IN* TECHNOLOGY FROM A BLACK FEMINIST PERSPECTIVE

Chapter 4: Diversity Concepts in Computer Science and Technology Development: A Critique

[This chapter was originally published in: Schelenz, Laura. 2022. "Diversity Concepts in Computer Science and Technology Development: A Critique." *Science, Technology, & Human Values*, 1-26. https://doi.org/10.1177/01622439221122549, reproduced with permission from SAGE publications]

Diversity has received heightened attention in computer science (CS) and technology development (Chi et al., 2021).[23] On the one hand, diversity is promoted in the workforce of the technology industry. Programs targeted at underrepresented groups encourage their inclusion in design teams. One prominent example is the Google-funded organization "Black Girls Code" (Noble, 2018, p. 64). On the other hand, computer scientists and designers increasingly leverage diversity as a concept in the designs of their products. Areas in CS that increasingly engage with diversity include personalized recommendations (Ricci et al., 2015), diversity-sensitive human-computer interaction (Himmelsbach et al., 2019), and design approaches for social justice (Costanza-Chock, 2020). Furthermore, works on structural discrimination through algorithmic mediation (Benjamin, 2019b) and data bias (D'Ignazio & Klein, 2020b) deal with the risks of a lack of representation and flawed interpretations of diversity.

This paper is *not* concerned with the former aspect: diversity in technology development teams. In this paper, I am interested in conceptual understandings of diversity that are implicitly or explicitly employed to design technologies. I understand diversity concepts as notions of "difference" that are filled with different meaning depending on the context and objective of their use. Diversity is then considered a *descriptive* concept, which nevertheless has to be acknowledged as value-laden since definitions of difference organize social relations and produce systems of inequality (Subramaniam, 2014; Yuval-Davis, 2011, 47f).

Despite increased interest in diversity in CS, actual diversity *concepts* remain understudied. What or whom do computer scientists and designers consider diverse? How do they define diversity? What are the consequences of leveraging these diversity concepts in terms of social justice? I assess diversity concepts in two subfields of CS, personalized recommendations and human-computer interaction (HCI), by conducting a literature review of academic works between July 2018 and July 2021. These two fields are major areas of CS with

[23] In this paper, I use "computer science" and "technology development" synonymously because both research and development go hand in hand.

significant output in publishing and product development. They are interesting because they deal with differences between users at scale, such as modelling users, to personalize content and improve the user experience. I reconstruct diversity concepts from the literature by mapping their meaning and clustering similar meanings into categories.

Drawing on Black Feminism and Critical Race Theory as well as literature that is inspired by these theoretical frameworks, I argue that diversity concepts in CS fail to account for *structural* inequalities between and within groups of users. Diversity attributes largely rely on flat, binary, and individual-level notions of difference. This poses the risk that "diversity-aware" technologies may stereotype and exclude marginalized users, thereby reinforcing existing inequalities. These findings may be attributed to the highly commercialized context of the fields under investigation. The discussion raises the possibility to (re)link diversity to notions of redistribution and access without economic calculation of diversity's benefits.

Investigating the use of diversity concepts in CS is important because it gives researchers of technology insights into the less obvious mechanisms of a) legitimizing harmful technologies and b) sorting and classifying users into oppressive hierarchies. On the one hand, diversity is largely regarded as "good" (Vertovec, 2012). Diversity is associated with inclusion and pluralism, and "diversity-aware" technologies may naively be perceived as fair and just. Their risks become obscured. On the other hand, diversity concepts are powerful tools. They are performative, which means that they "do" something. They transport norms and values, bring about social relationships (who belongs and who does not), and even establish our social order. They not only describe but *prescribe* difference (S. Ahmed, 2012; Haslanger, 2012, p. 241; Yuval-Davis, 2011, 47f). Diversity categories and concepts then have real effects on those subjected to them.

The paper's critique is relevant for scholars in the humanities, social sciences, as well as Science and Technology Studies, who are interested in diversity. It extends critiques about the structural nature of algorithmic discrimination (Benjamin, 2019b; e.g., Noble, 2018) to technology that heavily draws on diversity concepts. It thereby traces the origin of structural discrimination to conceptualizations of "difference" that are embedded in the technology. The paper can further offer insights for computer scientists and designers on how *not* to use diversity in their designs.

The paper is structured as follows. The second section offers background on diversity concepts: What are they and why are they relevant? Then, I introduce the theory (section 3) and methodology (section 4) used to conduct a critical analysis of contemporary diversity concepts in CS. The fifth section presents the findings of my analysis, where I cluster diversity concepts

in the fields of personalized recommendations and HCI. The critique (section 6) draws on works from Black Feminist and Critical Race scholars as well as literature inspired by the former to uncover the problems with the diversity concepts identified in section 5. Finally, the discussion and conclusion (section 7) reflect on ways towards designing technology that aligns diversity with social justice.

Diversity Concepts in Society and Technology

While diversity is popular and many people affirm its importance, diversity remains an ambiguous concept. In general, a diversity concept is an abstract idea that can be operationalized in various ways. According to Steel et al. (2018), the term "concept of diversity" refers to "an understanding of what constitutes diversity, abstracted from questions of which attributes are relevant to diversity in a specific context" (762). The authors look at diversity within and across groups of people, and differentiate between egalitarian, representative, and normic concepts of diversity (762, 767). While egalitarian diversity concepts measure the variation of people within a group, representative and normic concepts of diversity measure the divergence of a group from group proportions in society or a given non-diverse norm (a hegemonic group). This mapping of diversity concepts helps analyze diversity discourses, but the authors remain at an abstract level and do not expand on the concrete population of diversity concepts in a given context. In contrast to their work, I am interested in how diversity is *defined* in a particular setting, not in the logic of its measurement. Hence, I look at *conceptual diversity as some notion of difference*.

Diversity as "difference" is constructed in numerous fields of study. In biology and medicine, diversity refers to different species or bodies (Roughgarden, 2013). In international relations, diversity refers to the pluralism of values, norms, and interests of nation-states (J. Williams, 2016). In business management, diversity is seen as different skills of employees (Jack, 2016). In sociology, behavior and routines can differ between individuals and groups (Gildemeister, 2008). Although widely used, these conceptual definitions of diversity are contested. In the examples above, the authors cited challenge the pronounced perception of diversity. A growing critical movement evaluates the "relationship between the operations of power and the construction of difference" (Steyn, 2018, p. 6). Critical Diversity Studies investigates how difference is leveraged in projects of oppression like colonization but also how diversity discourses are transformed through social movement and identity politics (ibid).

A critical perspective on diversity asserts that definitions of difference are not neutral but carry values and must be analyzed against the historical contexts of their conception. One

example is race. In Subramaniam (2014), the author traces the origins of definitions of "variation" and highlights that they have power and politics: "Questions of genetic variation in human and nonhuman organisms [in the natural sciences] are deeply linked to questions of diversity and difference in human populations steeped in tortured histories of slavery, colonialism, and genocide" (7). Understanding the normative underpinnings of diversity concepts and how they operate in real life is crucial because categories of difference are widely used to make sense of the world. When we encounter different people, we usually engage in "social sorting" (Lyon, 2003). We sort the encountered into one of many categories: female or male, Black or White, heterosexual or homosexual, abled or disabled, etc. This sorting is the result of an internalization of discriminatory images, including stereotypes, of different groups. Through media, politics, and historical story-telling, society produces imaginations of "normal" and "abnormal" identities. In the USA, the norm is White and male (Hull et al., 2010). This society-mediated categorization into norm and "Other" leads to stereotyping and the production of stigma (Goffman, 1990; Yuval-Davis, 2011, p. 47).

Technologies similarly rely on the mechanism of social sorting (Lyon, 2003). They draw on diversity categories to classify people in order to make decisions. At the same time, widespread techno-optimism or techno-determinism suggests that these processes of categorization are neutral (Benjamin, 2019b, 8f). In other words, while we may be aware of social mechanisms of (discriminatory) sorting in the form of bias, we often trust technology to do a better job because of its reputation as rational and progress-oriented. Techno-optimism paired with the idea that diversity is inherently "good" (cf. Vertovec, 2012) may therefore prevent researchers from scrutinizing diversity concepts embedded in technologies. To minimize this gap, the present paper contributes a critical diversity perspective on concepts of difference in CS.

Black Feminism, Critical Race Theory, and Diversity

Critical perspectives on diversity and its coneptualization in society and technology are advanced particularly in Black Feminism and Critical Race Theory. Black Feminism is a critical social theory that combines scholarship and activism around topics concerning Black women.[24] One product of Black Feminism is intersectionality, an analytical tool that helps understand

[24] While there are multiple Black Feminisms, this paper focuses on US-centric thought, which has emerged from Black Women's resistance to their enslavement in early US history (Guy-Sheftall 1996, 2f). The focus on American Black Feminism stems from the author's background in American Studies.

multiple, interlocking systems of oppression (Crenshaw, 1989). Following an intersectional analysis, persons inhabit different positions at the intersection of various systems of oppression (race, gender, class, sexuality, ability etc.) which determine their daily lived realities (Collins, 2000a, 24ff; Collins & Bilge, 2016, p. 2). These systems are formed through discourse, policies, law, attitudes, and social practices. They produce an ”unjust situation where, systematically and over a long period of time, one group denies another group access to the resources of society“ (Collins, 2000a, p. 4). As a result of these systems, different social groups are subjected to different privileges and oppressions (Collins & Bilge, 2016, p. 2).

Critical Race Theory deals with “race” and racism in America. It differs from liberal or conservative scholarship in the way it criticizes civil rights discourses (Crenshaw et al., 1995, pp. 13–14). A distinguishing feature of the theory is its structural analysis of oppressive systems. This means that scholars look beyond the individual level in explaining racism. They see racism as a structural, institutionalized form of anti-Blackness (Crenshaw et al., 1995, p. 14). Another distinguishing feature is the critique of color-blindness. While race is socially constructed, it has real consequences for Black people and must therefore be recognized and addressed (Bonilla-Silva, 2018, p. 8; Crenshaw et al., 1995, p. 26).

Black Feminism and Critical Race Theory inform critiques of diversity in the humanities, sociology, and political science. For instance, S. Ahmed (2012) charges that diversity has become a management practice that incorporates marginalized groups into White, neoliberal institutions that continue to derogate Black women. Diversity is presented as the “happy” work of inclusion (14), while the challenging work such as grappling with racism is dismissed (ibid, 64). Social inequalities are ignored by liberal diversity discourses (Dhamoon, 2010). Originally, the Civil Rights Movement adcovated for inclusion in terms of redistribution and remedy for historical injustice so that oppressed groups had a fair chance of participating in economic and political opportunities. A backlash against affirmative action resulted in legal cases that established a looser commitment to diversity (Nash, 2019, 23f). Diversity was framed as beneficial for everyone instead of focusing on the needs of historically marginalized groups. As a result, representation in organizations compared to the inclusion of a few affluent Black people rather than transforming power relations (Berrey, 2015, p. 9). Hence, mainstream diversity discourses invoke a sense of tolerance while at the same time upholding a racial order that benefits White people (ibid, 269).

Methodology

This paper conducts a qualitative analysis of diversity concepts in CS. I first conduct a literature review in two areas of CS: personalized recommendations and HCI. The former field is interesting because personalization demands that we consider how we can meet the different preferences of users (Ricci et al., 2015, pp. 5–7). HCI is interesting because, with increased globalization, the field has started to adapt technologies to (culturally) diverse users (Allard & Blanchard, 2011). Both fields are thus motivated to explore differences between users. This invites us to critically examine how difference is conceptualized. The search included terms such as "diversity," "difference," "technology," AND "recommender system," "recommendation", "personalization," OR "human-computer interaction." It was executed in the ACM Digital Library, IEEE Explore, Google Scholar, and the University of Tübingen library catalogue with a filter limiting publications to dates between July 2018 and July 2021. From there, I included further works mentioned in the references. The 120 papers I reviewed are annotated in a publicly available Google Document.[25]

Two challenges were encountered during the review. First, a search for "diversity" may include literature on diversity in design teams. While this research is crucial, my goal in this paper is to investigate the diversity *concepts that are inscribed* in technology. My guiding research questions helped limit the material: How do computer scientists define diversity concepts? What or who is considered diverse? The focus on the conceptual notion of diversity is inspired by works that point to the value-ladenness of technology (Friedman & Nissenbaum, 1996). Values that are inscribed in technology have an impact on the performance of the technology in the real world. Benjamin (2019b) notes that with certain values, racial orders that disadvantage Black people are also introduced into technologies (8-10). It is therefore crucial to critically examine what kind of values are embedded in technologies, both in terms of how they are defined (input data) and what they "do" (effects). Second, diversity concepts can be implicit, and literature presenting them does not necessarily appear in a search for "diversity." They must first be reconstructed to be included in a critical analysis.

After compiling and annotating the list of material, I mapped the meaning of different diversity concepts and categorized them. The categories were built inductively from the material instead of using pre-defined categories. Finally, I assessed the different diversity concepts by drawing on existing critiques of diversity that are inspired by Black Feminism and Critical Race Theory. This step was guided by the following research question: What are the

[25] The document is available at https://docs.google.com/document/d/1Ry9o6rzMqtbbhrmOD-MCRLQXZkymYDtIyUXjlS3dHzI/edit?usp=sharing

consequences of leveraging said diversity concepts in terms of social justice? The critique is presented in section 6, and possibilities to address the critique are deliberated in section 7.

Diversity Concepts in CS and Technology Development

This section presents the results of my analysis of diversity concepts in two subfields of CS: personalized recommendations and HCI. Personalized recommendations aim at optimizing the recommendation of items (a book, a video, a tourist activity) to a user. Recommendations are the cornerstone of the most profitable contemporary services like YouTube, Amazon, Netflix or Spotify (Rocca, 2019), and the field can be characterized as heavily commercialized (cf. publications in the conference RecSys). At the same time, recommendations become increasingly vital in areas such as information and democracy (Helberger, 2019). HCI deals with the experience of users in their interaction with a system. This includes the usability of audio and visual features in an interface, mouse and keyboard use, but also how users immerse themselves in augmented and virtual realities and deal with 24/7 screen time and social connectivity (Norman & Kirakowski, 2018, p. 3). The field is particularly interested in how we use technologies and conscious of the fact that technology today affects "all sectors of humanity: rich and poor, gifted and disadvantaged, those who relish technical sophistication to those who prefer to see technology as transparent" (ibid. 4).

Diversity Concepts in the Field of Recommendations

In the field of personalized recommendations, diversity concepts are either representations of items that should be recommended to a user ("item diversity") or representations of users that are targeted with recommendations ("user diversity," cf. figure 1).[26] Item diversity in recommendations is considered the variation of items in a pool of items (e.g. movies of different genres in a film database, Kaminskas & Bridge, 2017). According to Castells et al. (2015), "diversity generally applies to a set of items and 'pieces,' and has to do with how different the items or pieces are with respect to each other" (884). The focus on item diversity (identified in 18 papers) can be explained with the way that personalized recommendations work. The idea is to adjust recommendations to the preferences of users (personalization). Since designers have little or no information about a *new* user, they derive information about the user from their

[26] Note that not every paper in the corpus was attributed to a category in figure 1, as some papers did not mention diversity in a way that was relevant to answering my research question. At the same time, some papers included several understandings of diversity (e.g. item diversity AND diversity as a strategy to achieve user satisfaction) and were thus included in more than one category.

interaction with a recommended item. Based on the user's past interaction with items, the system recommends items that it determines could also be interesting to the user (Rocca, 2019).

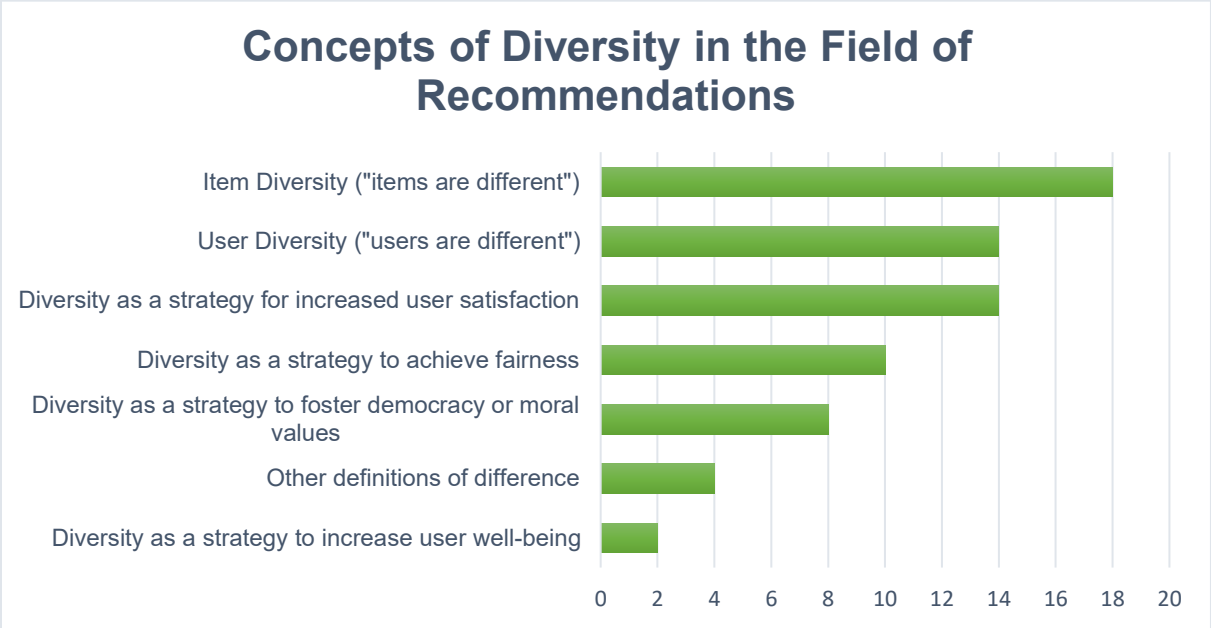

Figure 1: This figure shows the different concepts of diversity identified in the literature review of the field of recommendations and how often they appeared.

In the material studied for this analysis, item diversity refers to different content that is recommended to users. Holtz et al. (2020) consider item diversity as different podcasts in the music streaming service Spotify. But item diversity does not necessarily refer to *material* items. A recommended item can also be a person like someone to follow on Twitter (Skenderi et al., 2019). Item diversity can also be attached to user diversity. Jin et al. (2018) consider the recommendation of musical pieces and seek to enhance the diversity of recommended items (in simply terms, not recommend the same old song). They state that "diversity refers to the diversity of a recommended list measured by an intra-list similarity metric" which speaks to their understanding of diversity as item diversity (293). However, they acknowledge that users have different individual traits which can influence their perception of recommendations. They therefore combine item and user diversity (292).

User diversity indeed becomes increasingly popular in recommender systems. For instance, in content-based recommendations, instead of leveraging the item-user interaction profile, the system considers additional diversity qualities of an item or a user, e.g. the age of the user (Rocca, 2019). User diversity is then taken into account as a context factor, and demographic categories or personality traits are factored into the computer model (Wang et al., 2020). Abascal et al. (2019) personalize user interfaces for people with disabilities and thus

consider users' different physical, mental, and cognitive abilities. Kravčík et al. (2019) consider (cognitive) skills and affective responses of students in e-learning systems, while also laying out the importance of adapting intelligent mentoring systems to the user "based on ethnic diversity, cultural self-perception, preferences" (21). Users can also be diverse in their "situational contexts" such as "location, time and role" (Lakehal et al., 2018), their relationships to other users in a social network (Sanz-Cruzado & Castells, 2018), and – coming full circle – their preferences for diversity in recommendations (Loecherbach et al., 2021).

Apart from item diversity and user diversity, researchers and designers regard diversity as a strategy to achieve something else. Particularly item diversity is considered of strategic value. One idea is that, since computer systems could only approximate the "real" preferences of users, diversifying the item base can "optimize the chances that at least some item pleases the user" (Castells et al., 2015, p. 883). This falls into the category of user satisfaction, which is the most prominent strategy related to diversity (observed in 14 of the reviewed papers). Diversity is also seen as a strategy to increase fairness in recommendations. The idea is that diversifying the item base (e.g. music titles) gives authors of items (musicians) better chances of being recommended to users (C. Bauer, 2019). Some designers explicitly focus on niche items that usually fall through the cracks due to the common popularity bias in recommendations (Silva & Durão, 2020). Here, diversity as fairness is seen as equal opportunity for items (or authors of items) to be recommended. Recently, researchers additionally paid attention to fairness on the user side by considering the equal opportunity of users. So-called multi-sided fairness is especially important when users receive recommendations for economic opportunitities such as job advertisements (Burke et al., 2018). Finally, diversity is mentioned in conjunction with news recommendations and the need to foster democracy through the diversification of information, e.g. different sources, political views (Helberger, 2019). Sørensen (2019) points to the responsibility of public broadcasting services to recommend diverse content as they have an educational mandate.

Diversity Concepts in the Field of HCI

In their quantitative analysis of diversity in the CHI conference (how often was diversity mentioned in papers published in the conference), Himmelsbach et al. (2019) show that diversity has become increasingly relevant in HCI. Prominent diversity concepts include demographic criteria of users such as gender, age, sex, and education. Similar to Himmelsbach et al. (2019), my analysis finds that user diversity was the most prominent diversity concept and appeared in 23 papers (cf. figure 2). Although user diversity is not equivalent to *human*

diversity (for instance, Hirskyj-Douglas and Lucero (2019) explore dog-dog communication mediated via technology), the overwhelming amount of user diversity concepts refers to human diversity. HCI as a field is more human-centered than personalized recommendations in the sense that diversity usually refers to some human quality rather than material items. Goethe et al. (2019) define diversity modelling as user modelling: “To recognize diversity, the designer considers the kind of user frequenting the system, ranging from different skill levels, to users from diverse backgrounds and different skill sets and abilities” (96). Lazar and Barbosa (2019) consider user diversity as “expertise, age, gender, educational level, language, and culture, as well as disabilities and situational impairments” (1203). Other user diversity concepts include relationship to self (Kou et al., 2019), behavior (Banovic et al., 2019), and “gender-related cognitive styles” (A. Anderson et al., 2021). A variation of user diversity is the idea that bodily movement is diverse. This becomes relevant in cases where physical aspects such as hand movement matter for the usability of a system (Sun et al., 2020).

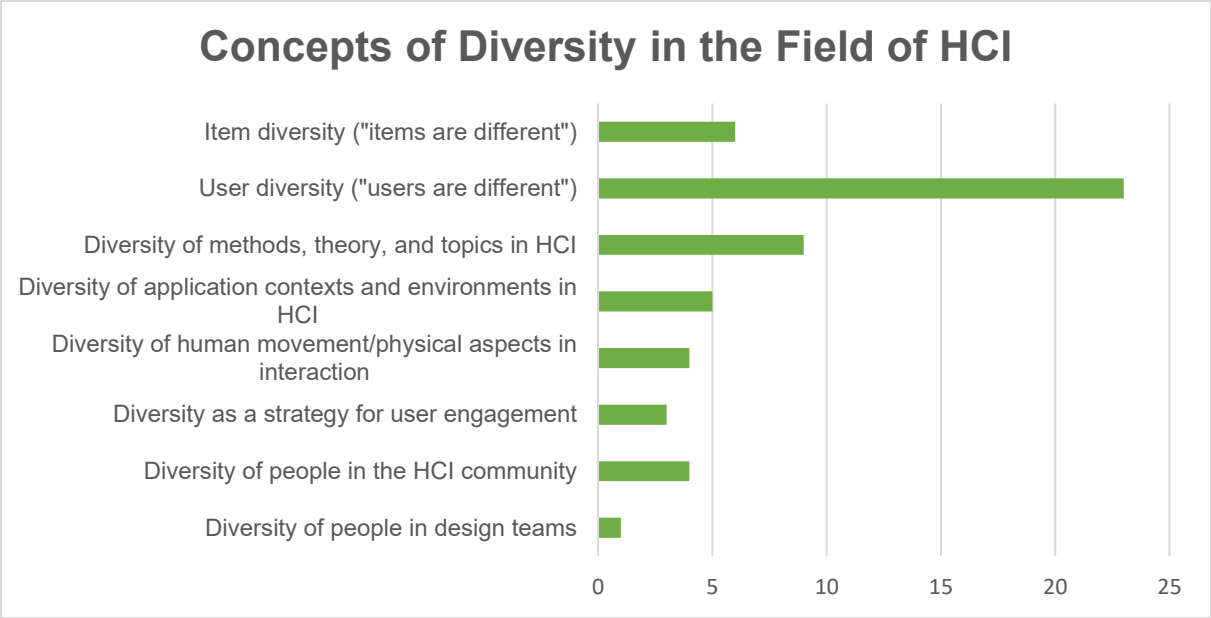

Figure 2: This figure shows the different concepts of diversity identified in the literature review of the field of HCI and how often they appeared.

Diversity also describes the methods, theories, and topics in HCI. The field is highly interdisciplinary. Researchers combine approaches from different fields such as CS and biology, and use microbes to adjust wearable devices or design interfaces with color and smell (Pataranutaporn et al., 2020). Hsu and Nourbakhsh (2020) use community citizen science methods and Monaco (2020) introduce covert impairment methods to study how users react to interruptions in a system. Similarly, HCI deals with a diversity of application contexts.

Nwokoye (2018) investigate the use of technology in Nigerian prisons and Pataranutaporn et al. (2021) explore how HCI can support human life and work in space. Less prominent but still relevant is the idea that diversity can be used as a strategy. Goethe et al. (2019) highlight two roles of diversity: to personalize interfaces for increased user satisfaction and to reach a wider audience by increasing accessibility for a range of people. Finally, item diversity has also appeared in HCI but only with reference to recommendations (see the previous section).

Critique of Diversity Concepts in CS

After reconstructing concepts of diversity in personalized recommendations and HCI, this section sheds some critical light on the diversity notions uncovered. From a Black Feminist perspective, it is striking that diversity not necessarily refers to human or societal notions of diversity. In the highly commercialized field of recommendations, system designers foreground the diversity of *items* (a book, a video) rather than users. Thereby, they think of users primarily as *consumers* of items rather than complex human beings. This said, user diversity plays a role and even more so in HCI. Yet, even when user characteristics are centered, the diversity dimensions utilized to describe users are presented as self-evident, natural, and value-neutral. Binary demographic criteria are named as standard measures of difference . Such definitions of diversity are problematic, because they a) filter people with multifaceted experiences into allegedly homogenous groups, and b) obscure the fact that these categories are not natural but socially constructed.

The use of demographics as diversity markers renders significant inner-group differences invisible. The practice suggests that for instance women are so similar that their experiences can be expressed in summary. This denies the reality of social inequalities. Depending on a woman's positionality in society, she is affected by different and intersecting systems of oppression (Collins, 2000a, 4; 227). Black women are not only affected by gender oppressions but also race oppressions (and class, disability, sexuality, etc., Crenshaw, 1989). Hence, experiences among women differ. Additionally, Black feminists criticize that the category of "women" is commonly understood to mean "White women," which renders Black female experiences invisible. Similarly, the group of African Americans is commonly understood to mean "Black men." The particular lived reality of Black women is thus erased (Hull et al., 2010). Since gender is usually divided into "male" and "female," this further marginalizes people who do not subscribe to one of the two labels. Queer, intersex, and transgender users are harmed in this process of forced categorization (Keyes, 2019).

Apart from erasing structural identities and obscuring social inequalities, researchers and designers in CS fail to account for the *social origins* of categories. Demographic categorization is not a product of nature but formed through practices and discourses shaped by power relations in society (Haslanger, 2012). Relatedly, demographic categories are not neutral. They invoke expectations, roles, and internalized images of the "diverse" groups, e.g. gender roles or racial stereotypes (Haslanger, 2012, p. 239). Diversity categories then create and manifest difference. This is how diversity categories become instrumental in projects of oppression.

One example is "race" and the use of data on race to create "fair" technologies. Hanna et al. (2020) criticize that researchers in Machine Learning Fairness lack awareness that "race" has a history of legitimizing Black people's oppression. The early US Census did not consider enslaved Black people as full persons and Census data was used to control immigration (ibid, 5). Collecting "race" data is then inherently interwoven with determining who belongs in a country and who is *not* protected. Using measures that go back to such oppressive practices can not only reaffirm harmful categories but further have unintended consequences when using such data to create "fair" technology. Hanna et al. (2020) suggest studying the *effects* of racism instead of race but also to consider alternative metrics to improve fairness (6, 8f).

User diversity is further conceptualized as different *cultures* of users (Goethe et al., 2019; e.g., Lazar & Barbosa, 2019). The idea is to be mindful of diverging values and practices across contexts. Yet, concepts of culture often rest on the notion that there is "a" culture within a given context, with a unique shared story that allegedly materializes in common cultural practices. But culture is complex, involves sometimes contradictory practices, and can hardly be captured in categories (Yuval-Davis, 2011, pp. 44–45). Furthermore, critics warn of the obsession with culture as a marker of difference, which can be attributed to mainstream diversity discourses that build on theories of liberal multiculturalism (Dhamoon, 2010). These theories tend to obscure social inequalities by reference to cultural difference. Global, regional or local inequalities become equated with cultural specificity. For instance, in the USA, Black communities may be perceived as suffering disproportionate lack of access to resources due to their distinct values rather than structural discrimination (Ford, 2009). Such accounts of "racial culture" obscure the fact that difference in social status is an outcome of long-standing oppression instead of group-specific traditions (Collins & Bilge, 2016, pp. 13–18; Ford, 2009).

To illustrate the problem with culture, let's imagine a system that accommodates different users by allowing for culturally specific communication practices. Such a system may exhibit a blind spot when it comes to marginalized users. State-sponsored discrimination can lead to illiteracy and poverty which affects the ability to communicate and even prevents access

altogehter. If the system further ascribes specific communication patterns (e.g. of semi-literate migrants) to culture, this may naturalize *social* inequalities as *cultural* specificity. Hence, a naïve priorization of cultural difference in computer systems may prevent access, accommodation, and justice.

Diversity in CS is also understood as a strategy to increase user satisfaction. In personalized recommendations, diverfying the range of items that are recommendated to a user can increase the user's satisfaction (Eskandanian & Mobasher, 2020). This likely fosters further engagement with the system which generates more profit for the system owner. Also concepts of user diversity in HCI are linked to satisfaction. Goethe et al. (2019) state: "Addressing issues regarding the accessibility and meaning of engagement for users from different cultures, ethnicities, socio-economic status can help designers to reach their product to the wider audience, while it can also offer more power to users from relatively deprived communities" (96). Here, serving marginalized users is subordinated to the goal of expanding the consumer group. The commercial interest dominates. Ideas of diversity as a means to make money are far removed from Black Feminist ideas of social justice such as those formulated by the Combahee River Collective (Guy-Sheftall, 1996, p. 231). Matters of representation, access, and visibility should not be weighted against their economic implications.

Diversity is further seen as a strategy to increase fairness. Two concerns should be highlighted here. First, as stated above, designers often define protected groups by reference to demographics – "gender", "race" – without awareness that these categories constitute systems of oppression (Hanna et al., 2020). Second, fairness is largely conceptualized as equal opportunity. While we can appreciate the importance of equal treatment as a means to fairness, fairness does not equal justice (ibid., 9). Fairness as equal opportunity does not remedy historical injustice. It rather obscures long-standing social inequalities that stem from centuries of structural discrimination against historically oppressed groups (Young, 1990, p. 12; Yuval-Davis, 2011, p. 77). S. Ahmed (2012) warns that "an equality regime *can* be an inequality regime given new form, a set of processes that maintain what is supposedly being redressed" (8). For example, a system that recommends quality jobs to Black and White users equally, does not take into account that White people benefit from privileges which put them ahead in a competition for quality jobs. It further neglects that White users have broader access to quality job postings via their networks. Although the recommender system in question is fair, it may not qualify as just, because it does not account for the structural disadvantages that particularly Black women face during a job search.

Finally, concerns arise from the overwhelmingly positive attitude towards diversity concepts. Considering diversity in CS *in some way* is associated with progress. Himmelsbach et al. (2019) note "an encouraging trend" with diversity increasingly present in HCI (8, 12). However, it does not matter how *much* researchers leverage diversity but *how* they understand diversity. If more and more scholars rely on binary and individual-level notions of "race" and "gender," nothing is won from a critical perspective. In Intersectional Black Feminist Technology Studies (Noble, 2016), diversity is not merely a matter of user representation. Rather, it determines how social relations of domination and subordination are (re)produced through the social construction of gender, race, and class, and mediated by technology. Himmelsbach et al. (2019) acknowledge the "need for progressive critical diversity research" (9). Yet, it remains a concerning trend that diversity is increasingly detached from its original motivation, such as equal rights, redistribution, and transformation of power relations. In their review of diversity initiatives in HCI conferences, Strohmayer et al. (2018) warn that "through a 'combination of overuse, imprecion, inertia and self-serving intentions' [13], the term 'diversity' has been seemingly removed and re-appropriated from its original idealistic intentions."

Discussion and Conclusion: Towards Principles for Diversity-aware Technology

The shortcomings of diversity concepts examined in this paper encourage us to explore other ways of engaging diversity. Black Feminism and Critical Race Theory provide the frameworks for critically studying diversity concepts. While both schools of thought are skeptical of mainstream diversity discourses, they also offer insights on "difference" that prioritize *structural* and intersectional-type analyses, and render visible social inequalities (Collins & Bilge, 2016; Dhamoon, 2010). Black feminists have further proposed design solutions for social justice. Erete et al. (2018) apply an intersectional lens to technology development, promoting three principles: understanding and attending to context, reflecting one's own socio-economic and cultural context, and disclosing dissent within a project. Costanza-Chock (2020) proposes to center marginalized groups of users in design. Noble (2018) and Benjamin (2019b) offer critiques of the construction of identities online and structural discrimination. D'Ignazio and Klein (2020b) have proposed design principles for the collection and use of data that build on a Black feminism.

From my analysis and considering the above-mentioned calls for social justice designs, some considerations can be made for technology design that leverages concepts of diversity. First, to avoid the pitfalls of working with individual-level diversity categories, designers are

encouraged to focus on *social* aspects of diversity. This means that questions of diversity cannot be framed as economic opportunity. Diversity has to be (re)linked to social justice and aligned with anti-capitalist values that prioritize representation, access, and redistribution. This would center the well-being of marginalized communities and decenter exploitation and profit.

Second, designers should not only make explicit who and what they consider diverse but further contemplate what their diversity understanding *does* to users and in society. Are notions of difference that manifest hierarchies of power reaffirmed? Are people classified into homogenous, static boxes that harm their self-perception and obscure *real* difference? When considereing differences between users, instead of drawing on demographic or cultural user representations, it may be beneficial to consider users' actual lived realities, particularly their experiences of privilege and oppression (Collins & Bilge, 2016, p. 2). Considering the experience of users does not mean collecting data on users' experiences and subjecting users to harmful data collection processes (Hoffmann, 2020). Rather, it means qualitative engagement with stakeholders, understanding their struggles, and how technology can support them.

As a third suggestion, and in alignment with Erete et al. (2018), it may thus be prudent to *contextualize* diversity categories by considering structural experiences. Attending to context allows us to link common diversity categories to their discriminatory effects. If we want to consider gender, we have to understand users' experiences of gender discrimination as well as how gender interacts with other systems of oppression. Elsewhere, I have shown how diversity concepts can be contextualized for the design of recommendation systems (Schelenz, 2021). Nevertheless, further research needs to explore whether and how considerations of social difference and inequalities can be married with design approaches in highly commercial contexts.

Chapter 5: Diversity-aware Recommendations for Social Justice? Exploring User Diversity and Fairness in Recommender Systems

[This chapter was originally published in: Schelenz, Laura. 2021. "Diversity-Aware Recommendations for Social Justice? Exploring User Diversity and Fairness in Recommender Systems." In *Adjunct Proceedings of the 29th ACM Conference on User Modeling, Adaptation and Personalization*, edited by Judith Masthoff, Eelco Herder, Nava Tintarev, and Marko Tkalčič, 404–10. New York, NY, USA: ACM, https://doi.org/10.1145/3450614.3463293, reproduced with permission from ACM]

Diversity and fairness have received increased attention in the field of personalization and recommendation (especially in the FairUMAP workshops). The two concepts are related. For instance, a diversification of recommended items is considered key to fairness. One example is the recommendation of search results. Presenting a diverse range of sources and representations of race and gender will increase justice towards populations that have previously been rendered invisible (Noble, 2018).

Initially, diversity in recommendations was a question of user satisfaction (we don't want to recommend the "same old" items (McNee et al., 2006)). Plus, diversity was used as a strategy to "optimize the chances that at least some item pleases the user" (Castells et al., 2015, p. 883) given the uncertainty about users' actual preferences. In this vein, methods such as re-ranking were developed "to achieve a balance between diversity and accuracy." Yet today, diversity is increasingly considered in fairness-related efforts (Sonboli, Burke, et al., 2020) and re-ranking is a tool to right wrongs in a recommender model.

This paper deals with the relationship of diversity and fairness from a social justice point of view (not a user satisfaction point of view). Reviewing literature in the field, I investigate how diversity is leveraged in recommender systems and evaluate the implications for fairness from a Black feminist perspective. The paper finds that, while item diversity may be an effective tool to increase fairness, the way researchers in the field currently leverage 'user diversity' compromises these efforts. In particular, I argue that a naive employment of user diversity models may compound previous injustices. Concern arises especially from the common understanding of diversity categories (gender, age, education, skills, practices, personality) as neutral, individual-level characteristics.

Black feminist theory helps verbalize these concerns. The theory is well suited to this analysis because it provides the vocabulary and methods to reveal *structural* inequalities, which are obscured by mainstream diversity concepts and rhetoric. In the methodology section, I

highlight the theoretical underpinnings of Black feminist theory. In the section on diversity and fairness, I tentatively relate the two concepts. In the section on item diversity vs. user diversity, I provide background information on the differences between item and user diversity. To elaborate the main argument, the section "user diversity in recsys [i.e. recommender systems]" provides a critique of 'user diversity.' Then, a hypothetical use case better illustrates the mentioned concerns about fairness. Following my critique, I offer tentative solutions for the handling of user diversity concepts. The section "contextualizing user diversity" proposes to contextualize user diversity by mapping users' experiences of privilege and oppression in a given diversity dimension, and the paper finally offers tangible recommendations to system designers.

Methodology and Theory: A Black Feminist Approach to Fairness

For the critical analysis of the relationship between diversity and fairness in recommendations, I first conducted a literature review of how researchers and designers in the field understand "diversity." The papers were selected by a keyword search in the ACM Digital Library and Google Scholar, combining the keywords "diversity," "personalization," and "recommendation." From there, I searched the references of the papers for further material. To ensure that the analysis is up to date, I further included 73 full papers from the ACM conference on User Modelling and Adaptive Technologies (UMAP) 2020 and the ACM conference on Recommender Systems (RecSys) 2019.

Then I reviewed the concept "diversity" from a (US-centric) Black feminist perspective. Black feminism is a critical social theory (Collins, 2000a) that has long dealt with questions of diversity from a social justice perspective. Contrary to other theories of social justice, which attend to fairness from a single axis lens, Black feminism advances an intersectional perspective. Intersectionality renders visible the existence of double or multiple converging forms of oppression that shape the lived reality of Black women (Crenshaw, 1995). These intersecting forms of oppression exist not just in the physical world. Algorithmic bias and discrimination disproportionally affects Black women (Benjamin, 2019b; Buolamwini & Gebru, 2018; Noble, 2018).

According to a Black feminist vision of "fairness," societal and technological systems must render visible and dismantle oppression and structural inequalities between social groups (Benjamin, 2019b; Costanza-Chock, 2020). Contrary to common perception, this does not necessarily mean equality of opportunity or equal treatment (Crenshaw, 1988, p. 1346). Due to historically grown discrimination, society does not represent a level playing field. Algorithms

may compound prior injustice "by carrying it forward into another domain" (Hellman, 2020, p. 828). Fairness thus means that a system dismantles oppression and prevents the compounding of prior injustice.

Diversity and Fairness

Diversity has both a conceptual and a normative component. Conceptually, diversity refers to the difference of many "things." Normatively, diversity is linked to discussions of multiculturalism and pluralism (Yuval-Davis, 2011, p. 39) but also calls for inclusion and justice: here, diversity is debated in terms of belonging, equal access to resources or matters of recognition (Fraser, 1996).

Fairness refers to the equal treatment of human beings. Rawls (1971) stresses that people should have equal opportunities, while inequalities resulting from different levels of talent and capabilities must be mitigated. However, this account of fairness may not hold in real life because it neglects (historical) discrimination that leads to an asymmetric playing field. New accounts of fairness should thus correct for historical injustice (Crenshaw, 1988; Hellman, 2020).

In the field of recommender systems, diversity and fairness are linked. On the one hand, researchers invoke diversity's normative quality by referring to diversity as a moral imperative. Ekstrand et al. (2019) stress that the legacies of historical discrimination may influence recommendations, which has implications for different groups of users. Definitions of fairness in recommendations thus build on notions of moral diversity, such as inclusion, non-discrimination, and justice. According to Sacharidis (2020, p. 313), "fairness means that the system exhibits certain desirable ethical traits, such as being non-discriminatory, diversity-aware, and bias-free."

On the other hand, conceptual diversity is leveraged to develop fairness-aware methods. The diversification of items (through re-ranking) is a prominent approach to increase fairness in recommender systems (Sonboli, Eskandanian, et al., 2020). Diversity from a fairness perspective wants to increase the number of different items that are recommended to a user, especially in a way that benefits protected classes as producers of items (Ekstrand et al., 2019).

Finally, some works consider diversity among users in fairness endeavors (Burke, 2017; Fu et al., 2020; Leonhardt et al., 2018). The focus here is on the performance of recommendations for different groups of users. This is an important new avenue. Especially Burke's 2017 "multi-sided" fairness hints to a tension between item-side fairness and user-side fairness, which is further explored in this paper (Burke, 2017; Korhonen et al., 2019).

Item Diversity vs. User Diversity in Recommender Systems

In order to further clarify the tension between fairness on the item side and fairness on the user side, let us first consider the diversity of items and users in detail. Research that explicitly deals with diversity in the field of recommender systems mostly refers to a) item diversity and b) personalizing the level of diversity in item recommendations. “Diversity” is considered the dissimilarity between items in an item pool (Kaminskas & Bridge, 2017). Castells et al. (2015, p. 884) provide a formal definition: “Diversity generally applies to a set of items and ‘pieces,’ and has to do with how different the items or pieces are with respect to each other.”

Eskandanian and Mobasher (2020) differentiate between individual diversity and aggregate diversity. Whereas aggregate diversity promotes a wide coverage of different items from an item pool (and thus mitigating a possible popularity bias that favors the recommendation of popular items), individual diversity describes the variability of items recommended to a user. On the individual level, diversity is also adapted to user's individual diversity tolerance (maximum variety vs. feeling overwhelmed, Eskandanian et al., 2017).

User diversity is considered rather implicitly in recommender systems and refers to information about the user. Since information about users is mostly gathered through implicit feedback (user-item interaction), the interactions of users with items can be diverse. Diversity in user-item interaction can be leveraged to build sub profiles or style profiles of users (Iqbal et al., 2019). However, Burke (2017) as well as Sacharidis et al. (2020) raised awareness about the difficulty to predict diverse user preferences based solely on implicit feedback. This difficulty is amplified in cold start scenarios.

Research increasingly deals with the diversity of users beyond implicit feedback. Frolov and Oseledets (2019) propose combining user-item interaction and additional user attributes to improve the quality of recommendations. Costa and Dolog (2019) propose a context-aware recommendation model that focuses on temporal aspects to contextualize user preferences. Dudzik et al. (2020) also stress the importance of contextualizing the user, e.g. observing their individual emotional reactions to videos, which may be triggered by personal experiences and activated memories.

Hence, we see a shift in attention from item diversity or user-item diversity to user diversity. While item diversity as a tool to build recommender systems and increase their fairness still dominates the field, emerging perspectives are interested in “What's in a user?” (Millecamp et al., 2020). This is a welcome shift as recommender systems may become more human-centric. However, some risks emerge for the fairness of a system.

How "Fair" Recommender Systems Risk Compounding Previous Discrimination

By now, it has been established that technology can reinforce existing structures of inequality (Benjamin, 2019b; Buolamwini & Gebru, 2018; see works on data and algorithmic bias, Noble, 2018). This section highlights a similar bias, which results from the way that user diversity categories are used in the design of recommender systems.

User Diversity in Recommender Systems - Challenges

The field of personalized recommendation draws on a series of diversity dimensions that describe types or groups of users. A preliminary mapping of diversity in two ACM conferences (RecSys 2019 and UMAP 2020) reveals 7 dimensions of diversity (see Table 1). Users are considered diverse in their demographics (e.g. age, gender, occupation), psychology (including personality and affect), physiology, culture, skills, social practices, and relationships. An additional area of diversity relates to specific functional preferences or preferences in the realm of user-computer interaction (e.g. how much transparency or control users wish to have over a system, Millecamp et al., 2020). This area, however, was not documented because it appeared less relevant to the research question.

A bias may emerge from the way that these categories are understood and leveraged in the development of recommender systems. In particular, three problems have implications for fairness. First, the diversity categories are treated as self-evident, natural, and value-neutral. Second, diversity categories are employed with little theoretical foundation and reflection of the social construction of diversity categories. In other words, what is missing in respective research is a written section that not only specifies the diversity concept but briefly lays out its origins and premises, and reflects possible implications of the diversity concept when applied within a recommender model. Third, the diversity of users is considered on an individual rather than a structural level. These three shortcomings usually occur simultaneously and are connected to each other. A number of examples exist:

Constantinides et al. (2020) study the effects of culture on users' interaction with picture passwords. The authors make transparent why they believe that "culture" is relevant to the subject but the authors take for granted that culture can be divided into "Eastern vs. Western"; in this case, the readers (and authors) may benefit from further theorizing the meaning of culture and how culture can be operationalized in a system (Constantinides et al., 2020, p. 44). Emerson et al. (2020) explore the "diversity" of students and speak of "student-level characteristics" and "differences in prior knowledge" [p. 66] yet it remains unclear what

"student characteristics" and "knowledge" entail; how are these concepts operationalized to produce (fair) recommendations? Millecamp et al. (2020) explore the relationship of "personal characteristics" and preferences for explanations in music recommender systems. Although the authors define personal characteristics in page 174, it remains unclear why this particular combination of diversity features is more relevant than others. Banskota and Ng (2020) and Mauro et al. (2020) optimize services for users on the autism spectrum by drawing on users' (dis)abilities as a diversity dimension; while these efforts are important to increase access and fairness for impaired users, the authors remain vague as to their conceptualization of disability and thus, in an unfortunate scenario, may risk reinforcing static ideas of "disability"; the cause could benefit from further reflection on disability, e.g. a specification whether authors adopt a social or biological/medical model of disability (Haegele & Hodge, 2016) and the implications of such a decision for the recommender model.

Another example is the common use of the Big Five personality model without reflecting its origins, premises, and implications for the particular use case (Alves et al., 2020; Sertkan et al., 2020; Wang et al., 2020); while the personality model may be popular and widely respected, it may nevertheless require contextualization in a particular use case. In one example, Wang et al. (2020, p. 270) nicely lay out the demographic and psychological user characteristics (age, gender, personality) in their study on the perception of serendipity. While this extensive reflection of diversity dimensions can be considered a best practice case, concerns arise from a lack of awareness of users' differences due to social status. How do societal structural dynamics (such as gender inequalities) shape the experiences and preferences of users? Might financially strained elderly female users have different preferences for serendipity than female users who are comfortably situated? Here, social context matters and even within diversity categories, intersectional experiences may result in divergent preferences.

The above examples were chosen to illustrate the difficulty of picking, operationalizing, and justifying diversity concepts in the development of computer models. Although attending to the diversity of users in the first place is commendable, risks emerge from a superficial treatment of diversity. Diversity concepts do not originate in a vacuum. They are socially constructed through discourse and practices (Haslanger, 2012). They are embedded in a given political and social context that determines what these diversity categories "do." For instance, diversity concepts (e.g. race, gender) can be used to establish hierarchies and a particular social order, where some groups are privileged and some are oppressed (Collins, 2000a). Hanna et al. (2020) have argued that the widespread understanding of race as an individual descriptor rather than a system impedes fairness efforts in machine learning. I pick up this line of argumentation

and argue that the way we currently leverage user diversity in recommender systems may impede fairness. The following hypothetical use case better illustrates this argument.

Diversity dimensions	Diversity of users	Sources
Demographics	Geographical location, age, gender, occupation	Frolov and Oseledets (2019), Zhao et al. (2019), Pei et al. (2019), Wang et al. (2020), Tian et al. (2020), Neve and Palomares (2019)
Psychology	Personality, psychology, cognition, emotions, affect	Alves et al. (2020), Landoni et al. (2020), Sertkan et al. (2020), Wang et al. (2020), Winecoff et al. (2019)
Physiology	Physiology, heart rate, stress levels, physiological reactions	Kundinger and Riener (2020), Zepf et al. (2020)
Culture	Sociocultural background, culture, language	Constantinides et al. (2020), Palmero Aprosio et al. (2020)
Skills/Experience	Skills, abilities, cognitive abilities, performance of different tasks, mental health, mental capabilities, knowledge, experience	Banskota and Ng (2020), Barz et al. (2020), Emerson et al. (2020), Herbig et al. (2020), Mauro et al. (2020), Papamitsiou et al. (2020), Rhim et al. (2020), Wu et al. (2019), Berndsen et al. (2019)
Social practices	Interests, activities, practices	Herzog and Wörndl (2019), Tian et al. (2020)
Relationships	Social relations, social interactions	Costa and Dolog (2019), Fan et al. (2019)

Table 1: Concepts of user diversity in the field of personalized recommendations

Hypothetical Use Case: Educational Diversity and Historical Discrimination

A public broadcasting website suggests documentaries to users. The pool of documentaries ranges from sports, nature, and travel to history, science, and politics. It includes diverse producers and featured groups within its pool. Rather than relying solely on user-item interaction, the system takes into account the “real” diversity of users. Diversity is understood

as educational level (high school diploma, associate's degree, bachelor's degree, master's degree and doctorate - which is asked when users create an account). The designers of the system assume that education is an indicator of preferences for documentaries: highly educated users are considered more interested in scientific documentaries rather than sports.

In this example, user diversity is explicitly considered in the process of designing a recommender system, contrary to classic collaborative filtering. Whereas the system "only" recommends documentaries, it adds to extra-institutional education for viewers and can inspire them to adopt interests or role models from the documentaries. In the absence of formal educational opportunities for marginalized groups (e.g. due to costs), recommendations for scientific documentaries can become an issue of social justice.

On the item side, diversity/fairness is achieved by including documentaries produced by or featuring diverse groups. On the user side, however, previous race and gender disparities in educational opportunities (Delmont, 2016) may result in recommendations of scientific documentaries to mostly privileged groups. The bias has emerged because designers considered user diversity as individual-level categories that are detached from larger social relations. They ignored the fact that members of different groups have different starting points that determine their chances to access the same resources. Because of historical discrimination (e.g. Jim Crow in the United States), marginalized groups (especially Black women) have not gained the same level of education as Whites and males. The legacies of such discrimination materialize today in a smaller number of Black women in science and higher education (Cheney & Shattuck).

Although the scenario itself is set up to incorporate biases (e.g. designers are said to treat education as indicator of preferences), the potential for unfairness emerges primarily from the treatment of diversity dimensions as individual-level characteristics rather than structural ones. The point of the scenario is then to illustrate a particular form of bias, namely the compounding of previous discrimination. The hypothetical use case reveals that fairness goals require closer attention to the way we conceptualize and utilize diversity categories. While diversity as a tool to achieve fairness on the item side may be effective, unfairness on the user side compromises item fairness.

Contextualizing User Diversity: Looking at Privilege and Oppression

In order to avoid compounding previous injustices in society via recommender systems, user diversity must be leveraged in a way that considers and dismantles structural inequalities. The first step to achieve this goal is to contextualize user diversity dimensions. Any diversity dimension (gender, race, education, disability, social practices etc.) involves a structural context that mainstream diversity discourses often neglect. This context relates to political and social

circumstances that determine how different members of society fare in this society. Black feminist and critical race scholars have pointed particularly to structural differences (inequalities) in society, which materialize in people's different experiences of oppression and privilege (Anderson & Middleton, 2018; Collins, 2000a; Crenshaw, 1988, 1995). Contextualizing user diversity dimensions then means understanding and attending to users' different levels of privilege and oppression in a given political and social context.

Oppression refers to a situation, where "systematically and over a long period of time, one group denies another group access to the resources of society" (Collins, 2000a, p. 4). The Black Feminist concept intersectionality highlights the multiple, converging forms of oppression that for instance Black women face (Cooper, 2016; Crenshaw, 1995): they are affected by gender configurations and race hierarchies, as well as heteronormative and classist value systems.

Privileges are the often unnoticed or taken for granted advantages that a group holds. McIntosh (1988, p. 4) offers an account of "white privilege" which renders visible the unearned advantages that white people experience, such as being sure that neighbors will be neutral or pleasant to them or that media productions feature people who look like them. Oppression and privilege are dynamic. Whether a social group is oppressed depends on the so-called "matrix of domination" (Collins, 2000a, 227f), which describes the particular organization of oppression in a society. This can be different from one location to another.

In order to reveal the structural context of a diversity dimension, I propose to map users' experiences of privilege and oppression within a diversity category. Wong-Villacres et al. (2018) have conducted such an exercise in a development project using ICT in the Global South. They unpacked the degrees and forms of "penalties" and "privileges" that their beneficiaries experience. First, they reviewed their (ethnographic) data and found common themes or subjects that shape users' realities. Then, they produced narrations of the life circumstances and socio-economic dynamics experienced by the users. Finally, they identified points of intervention, where the technology can alleviate penalties.

In the hypothetical use case, our diversity dimension was education. Fairness efforts were compromised because the designers picked educational "degree" as a diversity dimension. They did not take into account that the educational system in the USA does not afford everyone the same opportunity to gain a degree of higher education. To mitigate the bias, designers could map the privileges and oppression of groups of users in primary, secondary, and tertiary education by studying critical race and feminist works on (higher) education (for example Crenshaw et al., 2019; Morgan, 2019). This contextualization can help reiterate the choice of diversity dimensions and better reflect implications for the recommender model. Based on their

new insights, designers may decide to refrain from using “degree” to operationalize educational diversity. They may also become inspired to question the premise that education indicates preferences for scientific documentaries. Finally, they may decide to deliberately leverage recommendations to mitigate the oppression they mapped.

Following Wong-Villacres et al. (2018), there is potential in ethnographic studies of user groups to understand their needs and preferences. Similarly, participatory approaches are appropriate design methods to understand the specific experiences of privilege and oppression of affected user groups. Designers may seek access to so-called identity groups (e.g. refugees, Muslims, transgender women, Latinx) and conduct focus groups with members of these groups or expert interviews with NGO workers representing these groups to understand group members' experiences with the educational system.

Having said that, thinking of users in terms of identity groups has its limits. After all, a Black feminist intersectional lens highlights differences *within* identity groups depending on a person's position at the intersection of multiple oppressive systems. Furthermore, thinking of users as members of identity groups can have an essentializing effect (Young, 1990). Nevertheless, from a social justice point of view, identity groups are bound together by certain structural experiences. Precisely these experiences motivate the groups' political struggles for social justice (Combahee River Collective, 1996). It thus seems legitimate to start an inquiry with different identity groups and, from there, map diverse experiences of oppression and privilege which may transcend different groups. Yet, the process of contextualization of diversity requires further elaboration in future work.

Recommendations: Towards Diversity-aware Recommendations

The following recommendations and guiding questions can help designers of recommender systems avoid the compounding of previous injustice when they leverage concepts of user diversity:

1. Making a Concept of User Diversity Explicit

This may seem trivial, but it is essential to spell out what diversity dimensions we take into account, when we consider user context factors in computer models. This means not only stating a diversity category but also explaining which theoretical model we build on, and why the diversity category is relevant.

Guiding questions: What kind of diversity are we talking about? Who and what do we consider diverse? Why do we prefer this diversity concept over others?

2. Reflecting a Concept of User Diversity

Reflecting the diversity concept that lays the ground for categories of user diversity is important to understand what assumptions about human difference we buy into. Diversity concepts have been formulated in a given historical and societal context. Diversity concepts are socially constructed, often with a specific goal in mind, and thus carry baggage.

Guiding questions: What are the origins of the diversity concept? Who has developed it in which social and political context? What are the limitations of the diversity concept?

3. Contextualizing a Concept of User Diversity

Using existing diversity categories can lead to the unintentional reinforcement of injustices because the way these diversity categories are commonly understood obscures *structural* inequalities. It is therefore crucial to consider the different experiences of privilege and oppression of users in a given diversity dimension.

Guiding questions: What are privileges and oppression that users face in a certain diversity dimension? What are the cultural and historical contexts that shape a users' current experience in a diversity dimension?

Conclusion

This paper discussed challenges for user-side fairness and specified a bias emerging from the conceptualization and application of user diversity categories. It highlighted concerns that categories of user diversity are often taken for granted without questioning their origins or meaning in the larger social structure. A hypothetical example involving educational diversity illustrates how recommender systems may accidentally compound historical injustice if they disregard the structural context of diversity dimensions. In order to avoid the identified pitfall, designers can contextualize user diversity by mapping structural differences (inequalities) between users, which materialize as different experiences of privilege and oppression. Preliminary suggestions for the contextualization of user diversity are given in the paper. These are only initial steps to increase fairness with regard to user-side diversity, and further research should provide solutions along a justice-oriented design process.

Chapter 6: Diversity and Social Justice in Technology Design: Reflections on Diversity-aware Technology

[This chapter was originally published in: Schelenz, Laura. 2023. "Diversity and Social Justice in Technology Design: Reflections on Diversity-Aware Technology." *International Journal of Critical Diversity Studies* 5 (2). https://doi.org/10.13169/intecritdivestud.5.2.0033, open access, CC-BY, and reproduced with permission by Pluto Journals][27]

Diversity initiatives as a means to include representatives from marginalised communities in design teams have been reviewed and challenged in the field of Critical Diversity Studies, also in this journal (Chi et al., 2021; Twine, 2018). Yet, there is little critical engagement with the actual diversity concepts leveraged in the design of a technology. Diversity as a concept is increasingly used to design so-called diversity-aware technology. In such designs, a notion of diversity (e.g. the cultural diversity of users) is operationalised and employed as a metric to optimise the technology. Diversity concepts are thus embedded in the technology. Researchers and developers thereby seek to address ethical challenges of systems that use artificial intelligence (AI) such as algorithmic bias and discrimination (Benjamin, 2019b; Noble, 2018; Zou & Schiebinger, 2018).

One example of diversity-aware technology from the US-American context is diversity-aware datasets. Here, researchers actively consider diversity criteria (usually gender and skin colour) to ensure the representation of marginalised groups (Buolamwini & Gebru, 2018). Since data is used to train an AI-based system, the equal representation of the technology's stakeholders in the dataset is key to optimising the technology for them. Other examples of diversity-aware technology include methods in machine learning fairness (Oneto & Chiappa, 2020) and diversity-aware social media (Schelenz et al., 2021). However, there is an overwhelming focus on individual-level diversity categories (such as demographics, personality, culture) to describe technology stakeholders. This obscures structural differences such as social inequalities among the stakeholders and reinforces the bias researchers seek to address (Schelenz, 2022b).[28] Additionally, since diversity is widely perceived as "good" (Vertovec, 2012), technologies that appear to be diversity-aware may prematurely be seen as positive interventions and escape critical analysis.

[27] Please note that, due to the instructions of the journal where this chapter was first published as an article, this chapter uses British English instead of American English.

[28] By technology stakeholders, I refer to users but also those who are targeted or affected by the technology without actively making use of it.

This article scrutinises the idea of diversity-aware technology. I differentiate between individual-level diversity-aware technology and structural-level diversity-aware technology to shed light on the shortcomings of the former and the potential of the latter.[29] These families of diversity-aware technology can also be distinguished by reference to mainstream and critical diversity-aware technology. Both types draw on a) descriptive diversity concepts that offer a definition of difference (e.g. gender, race, age, values and norms, personality, capability, experiences, etc.) and b) normative diversity concepts that are value statements used by societal actors to create a sense of inclusion, tolerance, and justice. Table 1 illustrates the difference between individual and structural-level diversity-aware technology.

As a descriptive concept, diversity must be scrutinised because definitions of difference have been used to establish human hierarchies in the politics of nation-state formation, "ethnic hygiene", eugenics and colonisation (Rusert, 2017; Subramaniam, 2014). As a normative concept, diversity carries value. When I refer to a normative concept, I mean that it holds an idea of what is right and wrong. Since values are contested, there can be competing normative diversity concepts. For instance, multiculturalism has undermined other normative concepts like social justice and the redistribution of power that dominated the American Civil Rights Movement (Berrey, 2015; Dhamoon, 2010).

In order to highlight the potential of critical or structural-level diversity-aware technology, I define and develop criteria for a structural-level diversity-aware design practice. Thereby, I draw on Black feminism (BF) and critical race theory (CRT). Both theories are rich in material on diversity as they criticise contemporary diversity narratives and practices but also highlight structural differences in societal power relations.[30] This paper engages in a US-centric discussion of diversity. Most examples of diversity-aware technology and the theories used to analyse them are situated in a US-American context. This bias stems from my background in American Studies. The article discusses this limitation and acknowledges that Global South perspectives on design might have a vastly different take on diversity-aware technology.

As shown by drawing on BF and CRT, the potential of structural-level diversity-aware technology lies with the (re-)mergence of diversity concepts with considerations of social justice. Structural-level diversity-aware technology is defined as a tool that considers the

[29] Individual-level diversity-aware technology is synonymous with mainstream diversity-aware technology because I have observed it to be the common or default design practice Schelenz (2022b), see chapter 4 in this dissertation.

[30] While this article engages Black feminist thought, it is written from a White perspective. This means that the application of the theory, which itself is grounded in real-life experiences of discrimination, may be limited due to the privileged position from which it is engaged. The concepts of Black and White as well as Women and Men are capitalised on to illustrate their social construction.

experiences of privilege and oppression of its stakeholders and enables just societal relations while dismantling oppression. The five design criteria are: self-reflection of designers, mapping stakeholders' experiences, contextualising individual-level diversity dimensions, enabling just societal structures and dismantling oppression. The definition and criteria are helpful for designers as well as critics to use in their own designs or to compare existing diversity-related technologies to a desired design practice.

Furthermore, offering a conceptualisation of diversity-aware technology in terms of structural notions of diversity seeks to emancipate the term and design practice from widespread individual-level notions of diversity. Diversity has been socially constructed to be associated with individual-level notions such as demographic and cultural difference. This means that diversity can be redefined and mean something else (Haslanger, 2012, p. 242). In fact, diversity originally pertained to the redistribution of power and resources during the American Civil Rights Movement (Berrey, 2015). The benefit of defining structural-level diversity-aware technology stems from making visible the gap between mainstream practices of diversity-aware technology and social justice-oriented designs. Ultimately, the goal is to associate the term with design practices that seek transformation towards social justice and offer inspiration to designers who work with "diversity".

<table>
<tr><td colspan="2">Diversity-Aware Technology</td></tr>
<tr><td colspan="2">A technology that leverages a) a definition of diversity to account for differences among stakeholders of the technology, and/or b) a normative notion of diversity, usually to make a
value statement.</td></tr>
<tr><td>Individual-level (or mainstream) diversity-aware technology</td><td>Structural-level (or critical) diversity-aware technology</td></tr>
<tr><td>Definition of difference:
demographic differences
cultural differences
differences in personality and psychological aspects
(dis)ability
different social practices</td><td>Definition of difference:
different experiences of the stakeholders
different levels of privilege and oppression
structural inequalities
power relations</td></tr>
</table>

Normative notion/value of diversity:	Normative notion/value of diversity:
• tolerance • pluralism • multiculturalism • inclusion • fairness • equality	• social justice • representation • redistribution of resources and power

Table 1: Juxtaposition of individual-level and structural-level diversity-aware technology

Black Feminist and Critical Race Perspectives on Diversity

BF and CRT have produced and inspired critical perspectives on diversity and technology development. BF in the context of the US is the tradition of Black feminist thought in written and oral testimony, music, art and academia.[31] It theorises the experiences of Black women as a group, including their intersecting oppression and acts of resistance (Collins, 2000a, 1ff). CRT engages critically with the concept of race, and how imaginations of race have material effects and shape the lived reality of Black people in America. The field is related to Critical Legal Studies and traces how legal institutions produce and uphold oppressive race relations (Crenshaw et al., 1995). The literature brought into discussion in this section was chosen because it represents core critiques of contemporary diversity discourses and practices, the meaning of diversity and difference, and diversity in technology development. It is relevant for the conceptualisation of structural-level diversity-aware technology discussed below.

Critiques of Contemporary Diversity Narratives and Practices

BF and CRT have produced critiques of contemporary diversity narratives and practices. American scholars in these fields charge that framings of diversity established after the Civil Rights Movement of the 1960s obscure social inequalities and undermine structural change. Berrey (2015) and others (Holmes, 2015; Peller, 1995) argue that mainstream understandings of diversity represent watered-down accounts of social justice-oriented ideas like equal rights and the redistribution of power. After the Civil Rights Movement, there was a backlash against such ideas as well as policies like affirmative action. In court cases debating affirmative action, the use of racial categories in decisions concerning the admission of students to a university was defended on the grounds that this policy would increase diversity in the student body *(Grutter v. Bollinger, Grutter v. Bollinger 539 US 306, 2003, 23 June 2003)*. Diversity was

[31] The focus on an American-centric version of Black feminism stems from my expertise in this area, especially my background in Gender Studies and American Studies.

determined as a legitimate reason to factor race into admission decisions, protecting affirmative action but at the same time eroding diversity's transformative core. Following the afore-mentioned court decisions, universities established a looser commitment to "diversity work" (Nash, 2019, 23f). S. Ahmed (2012, p. 78) criticises that diversity became a practice of the privileged, which allowed institutional leaders to include minorities but simultaneously uphold their own power. Corporations adopted the diversity rhetoric established in the above-mentioned court cases to advocate pluralism in the workforce while undermining structural changes to organisational hierarchies (Jack, 2016; Vertovec, 2012, p. 293).

With regard to diversity narratives at the government level, Berrey (2015) shows that diversity was framed as "integration" after the Civil Rights Movement. In this narrative, diversity is the practice or policy of including previously marginalised racial and ethnic groups into societal and cultural institutions. However, as Berrey (2015) demonstrates, racial hierarchies remained intact and racism was played down. Black people or immigrants were expected to integrate into White institutions in a one-sided manner (p. 3).

In yet another narrative dominating public discourse, diversity is associated with the value of cultural pluralism. According to Dhamoon (2010), understandings of diversity are primarily informed by theories of liberal multiculturalism. Here, difference between social groups is seen as cultural difference. Tolerance of different cultures is considered key to societal cohesion. Dhamoon (2010) criticises this focus on culture rather than social status in public discourses in Canada. She argues that the focus on culture neglects differences between social groups that are based on social inequalities and an imbalance of power in society (ibid, p. 2). Furthermore, critics charge that cultural pluralism is celebrated only as long as the "other" cultures do not violate liberal values or threaten the essence of the dominant culture (Dhamoon, 2010, p. 7; Yuval-Davis, 2011, 55f).

These critiques highlight how the recognition of structural inequalities in society was replaced in public discourse by giving priority to individual-level notions of diversity like skin colour, ethnic looks, and cultural background. The replacement of structural notions (recognition of social inequalities, redistribution of power) with individual-level diversity detaches diversity from a commitment to social change and thus becomes non-threatening to the existing social order. Individual-level diversity concepts then represent what Nash (2019, 23ff) calls an "apolitical" or even "antipolitical" language and practice of diversity.

What Makes People "Diverse": Oppression and Structural Difference

BF and CRT have further produced studies on the creation, meaning, and effects of structural differences in society. They render visible structural differences between and within social groups and track the origins of these differences to the systemic oppression of social identities. In this way, they show what really makes people different, namely the way people are subjected to oppression and privilege.

One concept that helps understand different experiences of oppression is intersectionality. The Black feminist tool of analysis helps understand how systems of oppression interact and bring about a specific lived reality for those at the intersection of multiple systems of oppression (Collins & Bilge, 2016; Crenshaw, 1989). Systems of oppression are social constructs like gender, race, class, ability and heterosexuality that are used to order societies around a universal norm. Those who depart from the norm are considered "Others" and are actively "othered" by the dominant societal actors and, as a result, treated as less (Desmond & Emirbayer, 2009, p. 345; Goffman, 1990; Yuval-Davis, 2011). Here, definitions of human difference come in because they determine who represents the norm. In the example of race, definitions of human difference go back to "race science", a movement that emerged together with the transatlantic slave trade and European colonialism in an attempt to legitimise these violent projects (Desmond & Emirbayer, 2009, p. 338; Rusert, 2017, p. 8). Through historical constructions of race, the norm is considered White and the "Other" is considered Black (Hull et al., 2010). This means that the experience of Black women as a group is different from the experience of White women because their subjection to gender norms is compounded by constructions of race.

Intersectionality then generates the following insights: People or groups are different not because of their race, gender, class, abilities or sexual orientation but because socially constructed ideas of race, gender, class, ability and sexual orientation are used to subordinate some and privilege other groups. According to Collins (2000a), oppression is organised differently in different societies. This is theorised by the so-called "matrix of domination", which refers to the broad organisation of power within which intersecting oppression grows. In the US, oppression is organised via schools, housing, employment, government and other social institutions (Collins, 2000a, p. 227).

A point of clarification is important here because diversity and intersectionality are often conflated (Luft & Ward, 2009; Nash, 2019) and intersectionality tends to be absorbed by diversity discourses (S. Ahmed, 2012). Intersectionality is neither a theory nor a practice of diversifying teams or even redistributing power to different social groups. Intersectionality is also not a theory of identity that describes difference along categories of gender, class and race

(Cooper, 2016, 386f). The way that intersectionality speaks to diversity (in terms of a conceptual notion of difference) is by highlighting the different experiences of oppression of Black and White women. In other words, intersectionality provides the tools of analysis that help reveal how Women are structurally diverse in their experiences of oppression.

Beyond experiences of oppression, an intersectional analysis may also reveal experiences of privilege. Privilege represents the flip side of intersectionally constituted oppression. It refers to the unearned advantage of a group, again originating from the way power is structured in society (McIntosh, 1988). In Black feminist scholarship, it remains contested whether intersectionality should focus only on oppression or simultaneously theorise privilege (Allen, 2016, 10f). Nash (2008, 9f) suggests that intersectionality could be expanded to study experiences of privilege and how multiple forms of privilege come together or the combination of privilege and oppression come together in a person's lived reality. Dhamoon (2011) addresses the complexity of lived reality by referring back to Collins (2000a). She argues, also drawing on Fellows and Razack (1998), that individuals are usually shaped by both experiences of oppression and privilege at the same time (or privilege and penalty, as Dhamoon says, 2010, 2011). Oppression and privilege thus are not binary sources of structural difference. Rather, people are subjected to "differing degrees and forms of penalty and privilege" (Dhamoon, 2011, p. 235).

This is also true over time. Systems of oppression are not fixed or static but continuously reproduced, adjusted and reconfigured. Manifestation of inequality (new forms of racism like colour-blindness, Bonilla-Silva, 2018) and social change via movements like Black Lives Matter is possible (Clark et al., 2018). In this way, structural difference can be regarded as a fluid concept of diversity, and understanding the way people are really different requires the continuous examination of power.

Motivated by insights into the structural oppression of Black women, Black feminists in the USA have advocated for social justice. Black feminism relies on the combination of thought and practice, and explicitly seeks social justice not only for Black women but for all oppressed groups (Collins, 2000a, 22ff). The goal of social justice is best expressed by the statement of the Combahee River Collective, a group of Black feminists founded in 1973, which addresses the experience of Black women in the US (Guy-Sheftall, 1996, 231ff). Here, social justice is achieved through the destruction of systems of oppression that keep marginalised groups at the bottom and privileged groups at the top. Social justice speaks to diversity (in a structural sense) as it is a strategy to minimise structural inequalities. Social justice is also a

normative concept (of diversity) that appeals to the consciousness of society, conveying the moral notion that injustice is not acceptable.

Diversity and Social Justice in Technology Development

Recent years saw an increase in frameworks and critiques of AI-based technologies that draw on American BF and criticise the Western, liberal, capitalist, White, male norm underlying mainstream design practices. These frameworks and critiques share commonalities such as exploring how different worldviews, ways of life, and groups of people are marginalised through technology; demonstrating how social relations become inscribed in technology and how technology mediates social relations to create hierarchies of power; and advocating for a transformation of design towards social justice-oriented norms and practices. The following works offer important insights that can be synthesised and used in the next section to conceptualise diversity-aware technology from a structural perspective.

Noble (2018, p. 177) develops "Black Feminist Technology Studies", a framework of analysis that focuses on Black people's interaction with technology. With this framework, Noble gives priority to studying the experiences of marginalised groups and how marginalised groups' interaction with technology influences the dominant culture. She thereby acknowledges and attempts to mitigate structural differences between groups of technology users in how they gain researchers' attention and are represented. With another framework of analysis, "Race Critical Code Studies", Benjamin (2019b) proposes to examine how technology upholds and reinforces a structural racial order that disadvantages Black people. Her methodological framework focuses on the way technology is designed, keeping in mind that race is a technology itself (ibid, p. 45). Benjamin thus highlights structural racism in technology and shows how technology works differently for different people.

Hankerson et al. (2016) stress the psychological effects of discriminatory technology. They argue that technologies are primarily designed for White people, often unintentionally so due to internalised societal biases. The authors state that "technology like any aspect of society is just a bit easier to use if one happens to be Caucasian" (p. 475). They are concerned about the effects of exclusionary design on the well-being of excluded groups including social isolation, frustration, and additional hardship (p. 481).

While the previous works are critical of racist and sexist designs, the following works have proposed design frameworks for social justice-oriented technology. Erete et al. (2018) are interested in how underserved communities can benefit from designs. The authors provide a design framework in which particular attention is paid to the diversity of privilege and

oppression of users. Three design principles are identified: considering context, self-reflection and documentation of conflict (p. 68). Costanza-Chock (2020, p. 19) similarly centres marginalised communities as the primary beneficiaries of design, and dismisses "universalist design principles and practices". Erete et al. (2018) amd Costanza-Chock (2020) offer design practices that serve disenfranchised stakeholders. Diversity awareness is then tied to the empowerment of marginalised stakeholders.

Wong-Villacres et al. (2018) offer design practices that they tie to intersectionality. Whereas intersectionality plays a role in all of the above design frameworks and critiques, Wong-Villacres et al. (2018) seek to operationalise the experiences of stakeholders to account for their social context in the design of technologies. The authors propose ethnographic methods to map the privileges and penalties of groups of target users. In their own use case, they look at the privileges and penalties of girls and boys in India when designing an educational tool.

Finally, D'Ignazio and Klein (2020b) provide a framework for the collection, analysis, and distribution of data. Their framework builds on intersectionality and borrows insights from Black feminist traditions. While the authors are mindful of the history of data use from the transatlantic slave trade to eugenics and contemporary surveillance, they argue that data can be used in a feminist way to transform unequal power relations (p. 17f).

Shortcomings of Individual-Level Diversity-Aware Technology

From a critical perspective on diversity and technology development, there are concerns about what I term individual-level or mainstream diversity-aware technology. Such technology is designed with a diversity concept in mind that fails to account for structural inequalities in society. I wish to highlight three examples where diversity is defined differently but, in all cases, diversity is conceptualised as an individual feature of a technology stakeholder and thereby obscures power relations among groups of stakeholders.

One example is diversity-aware datasets. They are motivated by research that suggests discriminatory technology emerges from unbalanced, non-representative data (Criado Perez, 2019; Malik, 2018). In other words, because the dataset used to develop and train an AI-based system does not represent the diversity of those affected by the system, it does not work for them. A prominent case is the balanced dataset for facial recognition technology compiled by Buolamwini and Gebru (2018).[32] In their study "Gender shades", Buolamwini and Gebru (2018) find that popular facial recognition software fails to recognise and classify Black women

[32] Another example is the FairFace dataset by Kärkkäinen and Joo (2019).

accurately. The authors established much higher error rates for darker-skinned people than for Whites. To remedy this bias, Buolamwini and Gebru (2018) propose a dataset with an equal representation of dark-skin and lighter-skin as well as male and female presenting faces.

While Buolamwini and Gebru (2018, p. 2) make explicit the structural nature of discrimination through facial recognition technology, the operationalisation of skin tone and gender in the proposed dataset follows binary individual-level demographic definitions of diversity. The authors acknowledge that facial recognition technology is frequently employed to profile and oppress minorities (ibid., p. 2). However, the recognition of such structural inequalities in society is not translated into structural-level considerations of diversity in the development of the new dataset. Instead, the use of individual-level demographic features (skin tone, gender) in the dataset suggests that those subjected to facial recognition technology are equally affected by the technology. As a result, the individual-level treatment of diversity obscures the risk that facial recognition technology operating with a balanced dataset is optimised to increase surveillance and the disproportionate profiling of marginalised stakeholders.

Another example relates to machine learning fairness. This field of study deals with the definitions and methods required to produce fair predictions (Oneto & Chiappa, 2020). Beyond fair datasets, the field develops methods to put constraints on computer models to ensure a fair outcome or adjust unfair outcomes *ex-post*. Concepts of diversity and fairness closely relate to each other in this field, as protected attributes are reference points to provide fairness. However, designers tend to conceptualise protected attributes like race as individual-level demographic diversity. They fail to acknowledge the fact that protected attributes represent a history of inequality (Hanna et al., 2020). Race, for example, is not simply a demographic feature, but a system of oppression that produces privileges for White people and barriers for Black and Brown people (S. Browne, 2015; Feagin, 2006).

Two risks should be highlighted in the way diversity is conceptualised in fairness methods. On the one hand, the use of individual-level demographic categories of race obscures the negative effects of race itself. Race has historically been used to subordinate Black communities which has resulted in a structural lack of resources as well as poor education and health (Desmond & Emirbayer, 2009). Whereas the educational and health disadvantages of Black people are then a result of the creation of the concept of race itself, an individual-level demographic definition of race may naturalise these dynamics and represent Black people as

less educated and more prone to illness (Hanna et al., 2020, pp. 6–8).[33] On the other hand, treating race as an individual-level demographic diversity concept suggests that we should treat racial categories as equal in fairness models. But equal treatment fails to respond to historical injustice and may thus continue it (Hanna et al., 2020, p. 9).

The third example is diversity-aware social media. Existing social media platforms have been criticised for a lack of diversity, e.g. because they tend to promote filter bubbles and echo chambers (Sacharidis et al., 2020). Diversity-aware social media such as theWeNet platform seeks to address these challenges by leveraging the diversity of users and connecting them according to the users' needs for complementarity or similarity (Schelenz et al., 2021). The design team of the platform, which includes myself, has defined diversity as different social practices, which are routine activities like cooking, dancing, playing the guitar. The idea is that the computer model connects users based on their need for diversity. If a user wishes to fulfil a complex task like repairing a bike, they are matched with users who own the necessary tools and skills. If they seek someone to join their reading club, they are connected with like-minded bookworms (Schelenz et al., 2021).

With regard to the operationalisation of diversity, the approach taken draws on social practices as individual-level notions of cultural difference. Upon reflection, this operationalisation of diversity may introduce bias due to the coded nature of social practices. Social practices are gendered and racialised experiences (Haslanger, 2004). They are tied to certain associations, media representations and roles. For instance, ballet is associated with White female images and rapping is associated with Black male representations. A practice is then not simply enacted by an individual but tied to societal expectations of behaviour. Expectations can amount to oppression if they materialise in barriers for people who wish to break with established norms (Haslanger, 2004). Roles attached to different structural identities must then be made visible through a critical, structural-level perspective on diversity to avoid the algorithmically mediated reinforcement of stereotypes.

From Individual to Structural Level: Reframing Diversity-Aware Technology

While reflections on diversity-aware technology could remain focused on the shortcomings of existing practices, it is important to sketch out an alternative understanding of diversity-aware technology. This shows that diversity-aware technology can be reinterpreted to align with social

[33] The page numbers in the citation of Hanna et al. 2020 (as included in my original paper published in the IJCDS) do not correspond with the page numbers in the publication by Hanna et al. 2020 posted in the ACM Digital Library. This is because I was working with a PDF version that did not have page numbers at the time.

justice-oriented notions of diversity, and that designers can change their practices to design for social justice.

A Definition of Structural-Level Diversity-Aware Technology

In the introduction, I distinguished between a descriptive and a normative concept of diversity. The former is a definition of difference while the latter constitutes a value statement. In order to reconceptualise diversity-aware technology, I define both these diversity concepts by drawing on BF and CRT.

Concerning the descriptive concept of diversity, Black feminist accounts of difference have stressed the importance of structural oppression in Black women's lives. Black women's experiences are shaped by interlocking systems of oppression such as gender, race, sexuality, classism, ableism, nationality and more (Collins, 2000a; Crenshaw, 1989). On the descriptive side of diversity, difference then refers to different lived experiences shaped by societal power relations. Anderson and Middleton (2018, p. 1) nicely capture this notion of difference when they state that "we live in a society that includes a population of people who are all similar yet 'different' in how they move through this world, experience this world, and are perceived by others in this world". According to Nash (2008), Dhamoon (2010), and Anderson and Middleton (2018), these experiences include not only experiences of oppression but also experiences of privilege. As a descriptive concept, diversity then refers to different experiences of privilege and oppression of society members.

This descriptive definition of diversity has two implications. First, the focus is on social difference rather than individual or cultural difference. Diversity has to do with social experiences that are made due to (self-identified or externally ascribed) membership in social groups (Crenshaw, 1989; Young, 1990). Second, the definition foregrounds socially constructed and human-made differences. In other words, a descriptive definition of diversity acknowledges that inequalities are produced by societal processes of power formation. Defining diversity for the design of diversity-aware technology thus involves an active analysis and critique of power (Dhamoon, 2011).

On the normative side of diversity, a Black feminist approach allows us to tie diversity to social justice. Drawing on the statement of the Black feminist Combahee River Collective (1996), diversity from a normative perspective is then the value of just societal structures and the moral imperative to dismantle all systems of oppression. Diversity work, in turn, would be the actual labour of countering oppression.

Bringing our descriptive and normative diversity concepts together, the following definition of structural-level diversity-aware technology can be generated: Diversity-aware technology is a tool that takes into account the different experiences of privilege and oppression of its stakeholders. The technology enables just societal structures and helps dismantle systems of oppression.

I use the term "stakeholders" as a translation of "society members" (in the descriptive definition above) to account for those involved in creating the technology and those affected by the technology. This includes particularly designer and user. Noble (2018) stresses that technology is not neutral but reflects the values, worldviews, and thus biases of their creators.

Benjamin (2019b, p. 69) warns that the naive belief in value-neutrality of technology encourages blind acceptance and "assuming in the process that our hands are clean". Designers thus have a responsibility to consider not just the experiences of those affected by their design but also their own experiences. This means reflecting on their own privilege and oppression and how these find expression in the design.

Regarding the promotion of just structures and the dismantling of oppressive systems, it means that structural-level diversity-aware technology actively contributes to this end. Here, it is helpful to invoke the idea of "captivating technologies": According to Benjamin (2019a), contemporary large-scale technological systems hold Black people captive by extending and optimising practices of surveillance and incarceration. In such cases, the technology itself reproduces inequalities by its very logic and purpose. In such cases, Benjamin (2019a, 3f) suggests that we need to abolish the technology and imagine alternative (technological) futures that are structurally different from contemporary ones.

Criteria for the Design of Structural-Level Diversity-Aware Technology

Based on the definition of diversity-aware technology above, five criteria for its design can be derived. Criteria 1–3 are geared towards the descriptive aspects of diversity whereas Criteria 4–5 target the normative understanding of diversity-aware technology. They are addressed to designers and critics of technology. Although the underlying societal inequalities that a structural view of diversity seeks to render visible cannot be solved by designers alone, the way that designers shape technology influences how society interacts. It thus makes a difference on a societal level how designers approach concepts of diversity in their products. At the same time, a defined best practice for diversity-aware technology can help critics of technology to compare existing designs to a desired practice, and thus help identify potential bias and shortcomings of mainstream technology.

Criterion 1: Designers of the Technology Reflect Their Own Experiences of Privilege and Oppression

A technology that takes into account the privilege and oppression of its creators requires the creators to self-reflect on their positionalities in society. Self-reflection has been identified as a core element in intersectional-type design processes (Erete et al., 2018, p. 68; Erete et al., 2023b, p. 37). In their critique of racist technology, Hankerson et al. (2016, p. 474) stress that "this is not to say that members of the [human–computer interaction] community are intentionally creating racist technology, as pre-existing social bias can be unintentional, one possible source is 'white privilege'". In order to prevent unconscious bias, designers should reflect on their own experiences of privilege and oppression (ibid., p. 481).

Criterion 2: Designers Analyse and Critically Review the Experiences of Privilege and Oppression of Their Targeted Beneficiaries

Designers usually have some idea of "the user" or other stakeholders for whom they create the technology. A popular method to visualise the user is to build so-called personae. However, personae mostly constitute stereotypes, as in "a young female college attendant" (Wachter-Boettcher, 2017, 32ff). Instead of attempting to represent users, it may be more fruitful to consider the different experiences of users and the needs that derive from said experiences. Erete et al. (2018, p. 68) suggest that designers engage with the environment of the user by mapping the community of the user and taking into account the historical oppression that community members experience. Apart from historical oppression, how privilege and oppression affect stakeholders today may also be mapped.

Wong-Villacres et al. (2018) provide an example of such a mapping exercise. Through ethnographic studies in India, they contextualise the educational environment of possible users of their technology with a study on the users' privileges and penalties. The authors analyse gender roles that affect girls and boys, the influence of parental values, and the modalities of access to education in rural vs. urban parts of India (p. 49).

Criterion 3: When Designers Work with Individual-Level Diversity Concepts (e.g. Gender, Age, Disability, Personality, Social Practices), They Still Consider the Different Experiences of Privilege and Oppression of Stakeholders

Although the article provides a descriptive definition of structural diversity ("different experiences of privilege and oppression"), this definition is not intended for operationalisation.

Data on privilege and oppression may be extremely sensitive data. The collection of such data would violate the normative orientation of diversity-aware technology ("dismantle existing power relations") because it would broaden the power of technology developers at the expense of the data owners (Hoffmann, 2020). Where an operationalisation of diversity is required (e.g. in the area of personalisation technology), designers can contextualise existing diversity categories to account for their social effects. This means that designers who work with individual-level demographic and cultural concepts of diversity consider the stakeholders' experiences of privilege and oppression in these dimensions. For instance, personality at first sight seems independent of structural discrimination. Yet the way people are perceived and classified into personality categories may incorporate biases. Historically, Black women have been perceived as angry, loud and aggressive in comparison to White women (West, 2018). Such stereotypes must be considered to ensure that they do not creep into the dataset.

Criterion 4: The Technology Enables Just Societal Structures

Criteria 4 and 5 aim at the practice-oriented aspects of structural-level diversity-aware technology. In BF, theory and practice cannot be divorced. They are both inspired by Black women's consciousness of their oppression and their desire to transform unjust relations (Collins, 2000a, p. 31). Enabling just societal structures means that the technology actively supports social change. In the context of structural racism in the US, technology that supports social change may improve Black people's voting experiences, provide free education, or support access to affordable housing and healthcare for Black communities.

Criterion 5: The Technology Contributes to Dismantling Systems of Oppression.

Core to seeking justice is dismantling unjust social relations that are upheld via systems of oppression. This means that structural-level diversity-aware technology counters binary thinking and the promotion of a norm. It also means that designers envision alternative, just systems that can replace present systems—"system" both in the technological sense but also as a social structure. Afrofuturism is an artistic movement that sees imagination as a form of resistance (Womack, 2013, 42f). Imagination, for example in the form of Black science fiction, allows for the visualisation of a peaceful and prosperous future for Black people against the limiting, oppressive realities of the present. Inspired by Afrofuturism, computer scientists at the Hyphen-Labs developed Neurospeculative Afrofeminism, a science-based art form that breaks with racist and sexist narratives about Black women and technology (Benjamin, 2019a, p. 11;

Hyphen-Labs, 2019). Their artistic products can be considered a best practice of structural-level diversity-aware technology.

The Benefits of a Structural Perspective on Diversity-Aware Technology

Having reframed diversity-aware technology to align with a structural understanding of diversity, we can now revisit our three examples stated above and illustrate how a design practice for structural-level diversity-aware technology is better able to respond to the challenges of AI-based technologies.

In the case of diversity-aware datasets for facial recognition technology (Buolamwini & Gebru, 2018), designers would have to map the experiences of privilege and oppression of the stakeholders they seek to include in the dataset in accordance with Criterion 2. A mapping of the stakeholders' experiences would render visible that minorities are disproportionately affected by surveillance and profiling through facial recognition technology (ACLU, 2021; S. Browne, 2015). Designers would then acknowledge that the inclusion of data from minorities might increase minorities' risk of surveillance. As a result, designers would question the technology itself and, considering Criteria 4 and 5, abandon its design. A structural-level diversity-aware approach thus provides the methods to a) understand in what ways the technology may harm stakeholders, and b) make an informed decision about the discontinuation of harmful designs. The latter point is important because, given neoliberal pressures to deliver products, designers may not have the tools to imagine the possibility to dismiss risky technology.

In fact, Criteria 4 and 5 for the design of social justice-oriented technology question the possibility that a dataset used for the design of facial recognition technology can be diversity-aware. This is because much of the facial recognition industry is entangled with abusive practices of surveillance and profiling (ACLU, 2021). Additionally, Hoffmann (2020) warns that subjecting minorities to abusive data collection practices that focus on extraction and exploitation of information may further harm already marginalised groups. Hence, as a matter of ethical professional practice, designers of datasets may want to be careful in referencing diversity as a quality of the dataset.

An exception could be a facial recognition system that predicts "white collar crime" as presented by Clifton et al. (n.d.) and picked up by Benjamin (2019b, p. 5). Such a system balances societal power relations by turning attention away from over-scrutinised poor Black neighbourhoods and to under-scrutinised rich White neighbourhoods, especially financial districts. Furthermore, if a dataset was generated with data from underrepresented groups for a

different type of AI-based system, one that uses data to empower minorities in the style of data feminism (D'Ignazio & Klein, 2020b), the designers could follow Criteria 1–5 to design a diversity-aware dataset. In such a case, the contextualisation of diversity categories (Criterion 3) is key to mitigating the risks of obscuring structural inequalities and violently reducing complex identities to mere data points (Keyes, 2019).

In the case of fair machine learning, a structural view of fairness methods would inspire designers to reconsider their focus. Instead of looking at individual sources of discrimination or bias in the dataset or algorithm, designers would consider the broader social environment as well as the power relations in which fairness methods are expected to produce fair outcomes (Hoffmann, 2019, p. 904). Mapping the stakeholders' experiences of privilege and oppression (Criterion 2) helps account not just for inequalities that are fed into the technology via historical data but also for structures of power that continue to produce inequalities despite equal treatment (e.g. due to the effects of historical discrimination). Designers could then consider additional measures to mitigate discrimination like affirmative action (Criteria 4 and 5).

Consider a fairness metric for college admissions that ensures equal likelihood of different social groups to be admitted to college independent of the size of their group in society or the number of applicants in their group. While equal opportunity is realised, a structural-level diversity-aware technology would additionally implement measures to compensate for the unequal opportunities of students from marginalised groups to submit an application in the first place. These could be measures such as free training classes for marginalised groups on how to prepare an application. A structural-level diversity-aware approach thus encourages designers to go beyond technical fixes.

In the case of the WeNet platform, an example of diversity-aware social media, designers rely primarily on social practices to classify their stakeholders (Schelenz et al., 2021). Here, social practices need to be contextualised in accordance with Criterion 3. Let's consider the practice of "working". A mapping of stakeholders' experiences of privilege and oppression with regard to "working" reveals a variety of practices from formal employment to part-time work, informal labour, and unpaid reproductive and care work. It also reveals that some stakeholders like Women and immigrants, and especially female immigrants, are more likely to engage in informal and care work due to a history of unequal gender roles (Haslanger, 2004). If "working" was operationalised only in terms of a work contract or formal employment, stakeholders who engage in informal and care work may be disadvantaged. Following a structural diversity approach, designers could instead consider a more granular operationalisation of "working" that grasps different aspects of the practice. The structural

perspective thus helps designers understand where modifications to the technology are necessary to mitigate algorithmic bias.[34]

Limitations

The conceptualisation of structural-level diversity-aware technology provokes a series of questions about its limited applicability. They are addressed one by one in this section.

Mapping the Experiences of Stakeholders

The mapping of stakeholders' experiences of privilege and oppression is time-consuming and requires interdisciplinary cooperation with ethicists, Black feminists, and critical race theorists. At the same time, one could ask: who has the authority to map stakeholders' experiences? Shouldn't stakeholders map their own experiences? This would require participatory design methods (Muller & Kuhn, 1993). Given the extensive resources required for a proper mapping exercise, structural-level diversity-aware technology may not be realised at scale but in smaller, local settings.

Another concern is that experiences of privilege and oppression may become naturalised through their mapping. Here, it is important to keep in mind that the mapping of stakeholders' experiences is a practice of analysing power. Since configurations of power change over time, a regular analysis of stakeholders' experiences is warranted.

Finally, mapping experiences of privilege and oppression may neglect other types of experiences. By focusing on experiences that result from social dynamics of power, we ignore experiences that result from individual choice, personal character, and cultural influence. Designers then show bias towards certain experiences instead of treating *all* experiences as relevant. However, diversity-aware technology is about designs that seek emancipation from unjust structures. In order to transform unjust structures, social dynamics of power have to be seen, recognised, and centred (Bonilla-Silva, 2018).

Practicality of Design Criteria

The criteria for the design of structural-level diversity-aware technology outlined in this paper remain vague. Especially the normative component captured by Criteria 4 and 5 raises concerns about measurement. If a technology only qualifies as diversity-aware when it dismantles systems of oppression, how much transformation of the status quo is necessary? To what extent

[34] Of course, in order to qualify as diversity-aware technology, the WeNet platform would also have to fulfil Criteria 1–2 and 4–5. Here, taking into account the own experiences of designers, local conditions, and global power dynamics is key because the WeNet platform operates in an international context.

does a single tool even have the capacity to change social structures? These questions are left to the designer to consider. It should be stressed, though, that developing diversity-aware technology requires extensive effort on the part of designers. It is thus unlikely that structural-level diversity-aware technology will become the new mainstream design method. Rather, the point here is to raise awareness that simply labelling a technology "diversity-aware" does not make it inclusive, fair, and just. Rather, an active effort in working towards social justice is required on the part of designers.

Socio-Cultural Context and Perspectives from the Global South

My reflections on diversity-aware technology are guided by BF and CRT, which are both theories originating in the US-American context.[35] The examples of diversity-aware technology reviewed in this article are also situated in a US-American or European context. This means that the article's conceptualisation of diversity-aware technology is limited to a specific cultural and geographic context in the Global North. In the US, a binary framing of oppressed vs. oppressor dominates academic discourses around experiences of privilege and oppression. Less attention is paid to the oppression of a marginalised group by another marginalised group (Dhamoon, 2015, 29f). The paper thus neglects how marginalised groups can be complicit in colonial practices (Dhamoon, 2015; Fellows & Razack, 1998). This said, experiences of privilege and oppression are not fixed but dynamic and change according to the specific organisation of power in a socio-cultural context (Collins, 2000a, 227f). Privilege and oppression can thus occur in various constellations.

Another limitation is the lack of attention to the colonial context of technology development in the West (Coleman, 2019). Scholars from the Global South have criticised the exclusion of perspectives from the Global South in debating societal and technological futures (Bon et al., 2022). Raju (2020) calls for decolonising information (museums, libraries, archives) in "a world dominated by Western power and privilege" (p. 1). Since libraries and archives are the precursors of commercial search engines and databases, it is vital to look at the epistemologies shaping such information spaces (Noble, 2018, p. 12). Dos Santos Tavares et al. (2022) offer a postcolonial perspective on inclusion in the digital society. They suggest that closing the digital divide can have unintended consequences. Research must consider whether digital inclusion fosters critical consciousness and empowerment rather than merely connecting

[35] There are other strands of Black feminism that are situated outside of the US. The Black feminist orientation leveraged in this paper, however, originates in early American history and was a response by Black women to the enslavement of Black women. This article works with a US-centric version of Black feminist thought due to my expertise in this area.

users or—worse—revealing the "dark side of digital technologies" (p. 6). Bon et al. (2022) argue that technologies shaped by Western "corporate coloniality" exploit data from the Global South, reproducing historical systems of oppression (p. 63). Another problem is that Western high-end computing clashes with limited local infrastructure. Bon et al. (2022) thus call for co-creation with stakeholders on the ground and for "small-scale solutions, decentralised systems, and green, energy-efficient technologies" (p. 66).

Critical voices from the Global South question Western-dominated development practices, including technology development. Jimenez and Roberts (2019) propose that epistemologies from the Global South such as the Latin American "Buen Vivir" inform design as a way of decolonising the same. Escobar's 1995 critique of Western development trajectories and his 2017 reflections on design similarly offer alternatives to modernity-inspired futures (Escobar, 2012, 2018). Drawing on Latin American movements, Escobar advocates for cultural autonomy and stresses that "every community practices the design of itself" (p. 5). These views can be valuable for considering diversity in technology but also offer interventions regarding the idea of diversity-aware technology outlined in this article.

The Neoliberal Context of Technology Design and Development

This article has reviewed three examples of diversity-aware technology from the technology industry and research. These examples are situated in a neoliberal context, where pressure to create profitable innovation dominates the design process. One may argue that the capitalist approach to mainstream diversity-aware technology is fundamentally opposed to social justice. This raises the question of whether social justice-oriented diversity-aware technology can be designed within the technology industry. I do not have an answer to this question but suspect that social justice-oriented diversity-aware technology requires non-traditional funding and alternative spaces of creation.

Sustainable Development

So far, the article has not considered the environment. The anthropocentric focus on the human stakeholders of the technology undermines a vision of diversity-aware technology that prioritises nature and other species. However, since environmental protection is a form of social justice, a call for a responsible and reciprocal relationship with the planet may be implied in Criteria 4 and 5 for the design of structural-level diversity-aware technology. Future research may seek to combine considerations for human-centred diversity-aware technology with approaches to design laid out by Escobar (2018).

Conclusion

This article has highlighted the shortcomings of mainstream technology designs that leverage diversity concepts to build so-called diversity-aware technology. In a review of three cases (balanced datasets, fairness methods, and diversity-aware social media), the article stresses that such innovations fail to respond to algorithmic bias and discrimination because their designers leverage individual-level notions of diversity (demographics, personality, culture). Drawing on Black feminism and critical race theory, I have conceptualised critical or structural-level diversity-aware technology to offer a social justice-oriented alternative. A structural perspective allows designers to link diversity to social justice and build solutions that consider structural inequalities among their stakeholders. A structural diversity approach helps designers to discontinue harmful designs, go beyond technical fixes, and modify the technology to mitigate algorithmic bias. Designers can benefit from a definition and design criteria for structural-level diversity-aware technology by adopting elements of the practice into their own processes. Critics may use the definition and criteria to call out designers and criticise technologies that claim the label “diversity-aware” without an active analysis of power.

PART III:

BLACK FEMINISM IN A EUROPEAN CONTEXT?

DESIGNING WITH AND FOR THE MARGINALIZED

Chapter 7: Technology, Power, and Social Inclusion: Afghan Refugee Women's Interaction with ICT in Germany

[This chapter was originally published in: Schlenz, Laura (2023): Technology, Power, and Social Inclusion: Afghan Refugee Women's Interaction with ICT in Germany. In: *The International Journal of Information, Diversity, & Inclusion, 7*(3/4), 2023. DOI: 10.33137/ijidi.v7i3/4.40292]
[A change was made in this chapter compared to the original paper publication. The citation Galletta et al. 2019 was corrected because the original citation was faulty. The relevant citation is AbuJarour et al. 2019.]

At the beginning of 2022, Germany hosted around 240.000 asylum seekers from Afghanistan (Statistisches Bundesamt, 2022). Asylum seekers or refugees are people who leave their home country because they have a "well-founded fear" of being persecuted for their race, religion, nationality, or politics (Convention Relating to the Status of Refugees, 1951).[36] Afghan women face oppression in Afghanistan due to political, economic, and cultural conflict (Wörmann, 2003). Once arrived in Germany, Afghan asylum seekers face language barriers, discrimination, and demanding bureaucratic requirements such as responding to German-language letters from various from German administrations (Abdelhady et al., 2020; Die Beauftragte der Bundesregierung für Migration, 2023). To support refugees in overcoming these challenges, administrations and non-profit organizations (NGO) have teamed up with researchers to explore the potential of information and communication technologies (ICT) to foster the social inclusion of refugees.

Social inclusion can be understood as the possibility for "individuals, families, and communities" to "fully participate in society and control their own destinies" (Warschauer, 2004, p. 8). One example of an ICT project in Germany is the App "Integreat," developed to provide information to refugees (Schreieck, Zitzelsberger, et al., 2017). While such solutions carry potential benefits for newcomers, they can also inhibit use due to the influence of cultural factors and social inequalities. For instance, refugees may lack the knowledge and skills of navigating ICT in the first place (Alam & Imran, 2015; Sabie & Ahmed, 2019; Warschauer,

[36] For the purpose of this paper, there is no distinction between refugees and asylum seekers. The German legal system makes a distinction, though. Refugees have received a case review and were determined as being under threat of persecution in their home country – thus 'real' refugees. Asylum seekers have claimed to be refugees but have not yet received a review of their case by the German authorities. This paper is interested in the lived reality of Afghan technology users who have fled their country rather than in their legal status under the German system. Therefore, the terms refugee, asylum seeker and forced migrant are used interchangeably.

2004). It is therefore important to understand how social factors and power relations affect refugees' interaction with ICT and how solutions can be designed to their advantage.

This paper presents findings from an empirical study with 14 Afghan refugee women about their attitudes towards existing technologies and visions for future innovations that can support their social inclusion in Germany. How do Afghan refugee women in Germany experience ICT? What structural factors influence their interaction with technology? What are design features in an application that can support their settlement in Germany? Qualitative data from four focus group workshops held between March and June 2022 will be analyzed, including field notes, transcripts of audio recordings, drawings, and photographs.

For a critical analysis of the findings, the paper draws on Black feminist theory, which theorizes social inequalities and societal power relations (Collins, 2000a; Guy-Sheftall, 1996).[37] The theory is increasingly applied to the study of technology, and offers insights into the experiences of marginalized groups with mainstream technology designs (Benjamin, 2019b; Noble, 2016, 2018). The paper applies Black feminism as a way to extend its contemporary engagement with technology design to new cultural and social contexts, namely the case of Afghan refugee women in Germany. The application of the theory in this paper showcases its strengths in understanding complex configurations of power in different contexts.

While the paper does not claim that Afghan women are Black women or that their experiences are the same, the analytical methods of Black feminism are well suited to understand power dynamics at play in Afghan refugee women's interaction with technology. This is because Black feminism looks at structural experiences, i.e., experiences that go beyond an individual's lived reality and capture broader societal dynamics. In the case of Afghan women, although they are situated in a different cultural, religious, and social context than Black women, they too share experiences of marginalization that transcend individual experiences (Shahalimi, 2022a, 2022b). My application of Black feminism and my interpretation of Afghan refugee women's experiences is limited, though, by my own background as a White German citizen. Interpretations are made from an outsider perspective, and participants' lived reality cannot be fully grasped. Ethical implications of my positionality are considered in the methodology section.

Previous research on the use of ICT for social inclusion is scattered across various fields, including ICT4D (ICT for Development), Human-computer Interaction (HCI), Migration

[37] Black feminisms have been developed in numerous contexts, including Great Britain and African countries. This paper uses a US-centric version of Black feminism as Black feminists in the United States have engaged heavily with technology-related discourses in recent years, e.g., Benjamin (2019b); Noble (2016, 2018). The paper builds on these discussions and extends the application of Black feminist theory to other cultural and social contexts to show its strength as a framework for structural analyses beyond a specific historical and socio-cultural context.

Studies, Media and Communication Studies, and Information Systems. Previous works relevant to this paper study the use of ICT by refugees and the development of apps and services for social inclusion (Berg, 2022; Z. Chen et al., 2020; Fisher, 2018; Patil, 2019; Schreieck, Zitzelsberger, et al., 2017). In these works, ICT is often presented as an enabler of refugees' social inclusion in a techno-determinist way (see e.g., AbuJarour et al., 2019; Schreieck, Wiesche, & Krcmar, 2017). The present paper challenges this notion and argues that ICT can be a barrier to the social inclusion of Afghan refugee women due to their experiences of marginalization and biased design. Digital illiteracy, concerns over safety, and distrust of online communities limit Afghan women's participation in German digital society. The design of applications can be exclusionary if it consciously or unconsciously reproduces the needs and preferences of the majority society rather than considering the experiences of the marginalized.

The paper makes two major contributions to the academic and practitioner community engaged with ICT for social inclusion. First, it presents design features proposed by participants of the study for an ICT solution that could offer orientation to Afghan refugee women in Germany. The paper underlines the importance of centering marginalized people in design discussions to understand their lived reality, as proposed previously (Costanza-Chock, 2020; Erete et al., 2018). Second, the paper reflects on the broader infrastructural conditions for migrants' and refugees' integration in Germany. The German migration management system is designed to increasingly rely on digital tools, which creates barriers for low-literate groups of refugees. Expanding in-person services instead of developing another app may give Afghan refugee women a real opportunity to become members of German society. The paper then calls to rethink the taken-for-granted benefits of ICT-based social inclusion as presented in AbuJarour et al. (2019), among others.

Contextualization: A Black Feminist Perspective on Technology, Power, and Social Inclusion

The analytical perspective applied in this paper is inspired by Black feminism, a critical social theory developed by Black women which foregrounds the analysis of power, structural inequalities, and social relations in society (Collins, 2000a; Guy-Sheftall, 1996). It was developed as a form of resistance against the oppression of Black women in the USA during and after slavery. Black feminism has developed analytical instruments such as "intersectionality" to understand and counter systems of oppression (Collins & Bilge, 2016; Crenshaw, 1989, 1991).

Intersectionality, while contested as to its scope and meaning for a Black feminist agenda (Dhamoon, 2011, 2015; Nash, 2008, 2019), is about the interplay of structures that discriminate. Systems of oppression are socially constructed concepts like race, gender, and class that promote certain ideas about 'the norm' and 'the abnormal' and thus structure social relationships between groups of different skin color, genders, and economic power (Davis, 2011; Hull et al., 2010; Yuval-Davis, 2011). Intersectionality and the study of systems of oppression have been picked up and developed further in numerous contexts, including in Great Britain (Mirza, 1997), Germany (Kallenberg et al., 2013; Salem, 2018), and in the realm of technology studies (Benjamin, 2019a, 2019b; Noble, 2018; Noble & Tynes, 2016) and technology design (Erete, 2021; Erete et al., 2023b). The present paper adds to this application across context by focusing on the experiences of Afghan refugee women with technology in Germany.

A concept relevant to this paper is "diversity," as it relates to the notion of social inclusion. Black feminists have criticized mainstream diversity discourses based on their own experiences of oppression and as part of their critical analysis of power. According to this critique, common diversity narratives promote the social inclusion of Black women without fundamental, structural changes in society (S. Ahmed, 2009, 2012; Crawley, 2006). Here, diversity and social inclusion relate because diversity efforts are expected to bring about social inclusion. Nash (2019, p. 24) criticizes the apolitical nature of diversity narratives in mainstream liberal discourses that were focusing on diversifying college campuses. Diversity is associated with positive feelings, tolerance, and multiculturalism rather than a political call for changes in racial hierarchies (S. Ahmed, 2012, p. 53; Berrey, 2015).

Nash also criticizes that the concepts diversity and intersectionality are often mixed together, conflated or appropriated: "Where diversity is a project of including bodies, intersectionality is an antisubordination project, one committed to foregrounding exclusion and its effects" (Nash, 2019, p. 24). The inclusion of bodies as a diversity practice is then considered insufficient to counter existing hierarchies of power as those 'bodies' would be included in an existing unjust system. Intersectionality renders power inequalities visible, which is a first step towards transformation. This perspective is relevant to the paper as it raises questions about ICT-based social inclusion (if practiced as the inclusion of 'bodies' in digital services without regard to power) as a potentially non-transformative practice that keeps the marginalization of Afghan refugee women in place.

Black Feminist Technology Studies expands the Black feminist theoretical framework to investigate questions of power in the design, development, distribution, and use of

technology (Benjamin, 2019b; Noble, 2016, 2018). In Human-Computer Interaction (HCI), Black feminists have criticized the Western-centric and White design of technology . This argument relates to the question, "who designs" the digital services that diverse populations are using? With a predominantly White and male perspective in Silicon Valley (boyd, 2019; Wachter-Boettcher, 2017, p. 20), technology may incorporate the implicit biases of this group and marginalize the needs and preferences of users of Color (Benjamin, 2019a; Hankerson et al., 2016). Noble (2018, p. 9) states: "The human and machine errors [coming to light in recent years] are not without consequence, and there are several cases that demonstrate how racism and sexism are part of an architecture and language of technology." This has to be understood against the history of the United States, where Black people were enslaved and continue to feel the legacies of discrimination imposed by White people (Alexander & West, 2012; ACLU, 2021; Hinton, 2021). This perspective is relevant to the current paper as Afghan refugee women are situated in a position where they experience structural discrimination (Shahalimi, 2022b; Wörmann, 2003), and the effects of these experiences should be taken into account when designing (digital) services for social inclusion, cf. (Erete et al., 2018).

Furthermore, a critical perspective on diversity efforts such as developing ICT for refugees' social inclusion sheds light on the politics involved. According to Keshavarz (2020), when designing ICT for refugees, we have to keep in mind that European border regimes with high-tech 'shields' against refugees (cameras, tracking devices, biometric surveillance) are designed as well. They are designed to keep 'the other' out and away from European wealth and comfort (Keshavarz, 2020). From a Black feminist perspective on structural discrimination, the techno-socially mediated relations between Europe and refugees must be taken into account when discussing one-off ICT solutions that are promoted as key to social inclusion.

Background: Structural Experiences of Afghan Women

This section offers an overview of the structural experiences of Afghan women a) in Afghanistan, b) in Germany, and c) in interaction with ICT. Structural experiences refer to a lived reality that can be observed across individual context. Structural experiences are big picture trends of how a person belonging to a particular social group is received and treated in society. The section thus focuses on widely shared experiences of Afghan women in their home country, in the host country Germany, and when they engage with technology. Three insights from Black feminism are helpful in presenting the experiences of Afghan refugee women. Guidance from Black feminism does not imply that Afghan women are Black women or that

their experiences are the same. Rather, insights about Black women as a marginalized group in the USA offer inspiration for the analysis of marginalized groups in other contexts.

The first insight is that there are different configurations of oppression and privilege that may affect group members. Oppression refers to the systematic deprivation of resources of one group by another group, including material items, knowledge, and power (Collins, 2000a, p. 4). This oppression is based on perceived 'features' of group members (which are in fact social constructs marking difference), such as their race, ethnicity, religion, class, gender, and sexuality. Privilege means that a person enjoys unearned advantages in society, e.g. having a lighter skin color and therefore not being subject to everyday racism (McIntosh, 1988).

The second insight is about structural experiences of Black women. According to Collins (2000a, 25f), not every individual Black woman has to experience oppression for it to be a structural experience of most Black women. This means that a structural experience remains structural even if an individual has a different experience. The paper does not claim that experiences of Afghan refugee women are homogenous. There is a diversity in the group of Afghan refugee women with some being highly educated, running their own business, and being financially independent while others are socio-economically marginalized (Putnam, 2021; Shahalimi, 2022a). Yet the paper conveys the bigger picture (again, a structural understanding) of political, economic, and cultural influences on the lives of Afghan women.

Third, there is a dialectic of oppression and resistance that shapes the experiences of Black women (Collins, 2000a, p. 3). Generalizing this insight, it means that a marginalized group may be oppressed but simultaneously resist this oppression and therefore becomes an agentic group. In terms of ICT, an example is the appropriation by marginalized groups of technology that has been designed by and for members of the majority society (Martinez Demarco, 2023). Such appropriation can be part of an empowering strategy that uplifts the voices of previously underrepresented groups.

The third insight about the dialectic of oppression and resistance is particularly relevant for the presentation of Afghan women's experiences. Given persisting global narratives about the victimhood of Afghan women (Abdelkarim, 2021; Ellis et al., 2007), it is vital to avoid falling into the trap of reinforcing such narratives. For example, Afghan women should not be presented as helpless in terms of using ICT. As the study in this paper will show, Afghan women have their own suggestions for technology design. At the same time, it is important to recognize structural disadvantages, especially in the context of research with Afghan women. For example, Afghan refugee women's situation of deprivation in Germany may influence their

ability to give informed consent to participate in the study presented in this paper (Ellis et al., 2007, p. 469).

I will begin with the situation of Afghan women in Afghanistan. On August 15, 2021, the militant Islamist fundamentalist group Taliban took over the government in Afghanistan. Subsequently, the situation for Afghan women and girls has deteriorated to the extent that women are beaten and imprisoned for leaving the house without a male guardian (PBS Newshour, 2022). Also prior to the Taliban takeover, Afghan women experienced oppression from Taliban fighters and warlords. Given threats from these groups, Afghan women have barely received education, let alone political power (Joya & O'Keefe, 2010). Although educational opportunities increased significantly for girls between 2001-2020, the rate of women's literacy stands at around 30 percent (UNESCO, 2020). Traditional customs and strict interpretations of Islam further make it difficult for Afghan women to participate in public life, the economy, and politics (Nordland et al., 2019; Wörmann, 2003).

Insights into contemporary experiences of Afghan women are offered in Nahid Shahalimi's book *We Are Still Here: Afghan Women on Courage, Freedom, and the Fight to Be Heard*, first published in German in 2021 (Shahalimi, 2022a, 2022b). The author emphasizes the patriarchal nature of Afghan society with a divide between conservatives and liberals. Male members of the family have a strong influence on a woman's life, representing her in public and managing her wealth (Shahalimi, 2022b, p. 15). Progressive Afghans support women's education while conservatives prescribe women's role in the household (Shahalimi, 2022b, p. 16). Conservative forces like the council of Mullahs and the Taliban enforce strict rules that prohibit women to dance, sing, or follow a career (Shahalimi, 2022b, p. 30). Women in politics are threatened with death (Shahalimi, 2022b, p. 83).

These power configurations like gender hierarchies influence Afghan women's relationships with technology. The patriarchal Afghan society pressures women to be reserved, which means not sharing content, photos, and real names on social media (N. Ahmed et al., 2022, 3f; Hussain & Amin, 2018; Nader, 2020). Women in Afghanistan are also conscious of the fact that they could be (sexually) harassed online without legal remedy or support from the family (N. Ahmed et al., 2022, 3f). At the same time, ICT can be tools to circumvent rigid gender regimes. Putnam (2021) finds that Afghan business women rely on ICT to run their businesses anonymously, thus not visibly violating social norms (Putnam, 2021, 10f).

When fleeing from Afghanistan to Germany, Afghan women face hardship similar to other refugees (Berg, 2022, p. 128). In the host country, refugees face insecurities about the outcome of an asylum application and deportation (Suerbaum, 2023), as well as poor living

conditions and economic marginalization (H. Lewis et al., 2015). Germany has a special configuration of bureaucracy and power. There is enormous pressure from authorities on all refugees to fulfill bureaucratic obligations and follow the correct legal processes, which no less affects Afghan women (Abdelhady et al., 2020). Furthermore, Afghan refugee women experience Islamophobia and racism. Especially after 9/11, Afghan women have been portrayed as terrorists in Western media (Osman, 2019, p. 290). At the same time, Afghan women are considered victims in need of saving, which obscures their activism for women's rights, education, and political involvement (Joya & O'Keefe, 2010).

Indeed, Afghan women have been active in protesting their oppression in political campaigns, as local leaders, and online (Joya & Mallett, 2009; Shahalimi, 2022b, p. 25). ICT are used to support these political efforts to campaign for the rights of Afghan women because they allow women to voice their opinion online, receive remote education, and work from home (Code to Inspire, 2015; Shahalimi, 2022b, p. 49). A concrete example is the launch of social media campaigns such as #Whereismyname to support Afghan women's rights (Nader, 2020, p. 2).

State of the Art: ICT Use by Refugees and Designing for Social Inclusion

Previous research about the use of ICT for the social inclusion of refugees is scattered across scientific communities. ICT4D studies the potential of ICT to assist the self-help and development of refugees' livelihoods in host countries (Bock et al., 2020; Schreieck, Wiesche, & Krcmar, 2017). Human-computer Interaction (HCI) also has a research stream on ICT and development that deals with refugee and migrant experiences (Sabie et al., 2021). Information Systems investigates the use of ICT by refugees. Examples include the workshops "Empowering Refugees with Technology: Best Practices and Research Agenda" at the European Conference on Information Systems 2017, and "Leveraging Technology for Refugee Integration: How Can We Help?" at the International Conference on Information Systems 2016 (AbuJarour et al., 2019). Communication and Media Studies is interested in the digital, multimedia, and information literacy of refugees (Alam & Imran, 2015; Lloyd et al., 2013). New fields have also formed around ICT and (forced) migration such as "Digital Migration Studies" (Leurs & Smets, 2018).

This section brings together previous research from different disciplinary backgrounds, covering the experiences of refugees with ICT in host countries and the design of ICT for refugees' social inclusion. My own research extends prior work in the following ways: It fills a gap since Afghan refugee women's use of ICT in Germany has not received attention so far.

Although previous research has looked at refugees' experiences in Germany, these refugees have been groups of various nationalities. Conducting research specific to Afghan women sheds light on a group that is situated uniquely in German society due to social, cultural, and political factors. This paper also challenges previous research that exclusively promotes the benefits of ICT without critical reflection. While I do not deny the potential advantages of using ICT for isolated goals related to social inclusion. Yet, as this paper will show, there are limitations to using ICT for social inclusion and they should be considered in a German program for integration and migration management.

In previous works studying refugees' use of ICT, refugees were portrayed as heavy users of technology (Patil, 2019) and ICT were characterized as refugees' "lifelines to the past, present, and future" (Fisher, 2018, p. 100). In the German context, Kreß and Kutscher (2016, p. 89) reported that the first thing refugees do is to buy a SIM card and create a Facebook, WhatsApp, and email account. In a study with refugee women in Germany, Berg (2022) showed that ICT are crucial to the women's well-being because they offer access to language training, emotional support from family members back home, and distraction through entertainment. ICT also provide access to spirituality as online videos of religious leaders reciting the Quran can be consumed without being present in a mosque (Akca, 2020, p. 72). A significant barrier to ICT use is the lack of Internet access in refugee accommodations in Germany (Berg, 2022, p. 130). This was also shown in a study with unaccompanied refugee youth who use public wifi in fast food chains (Kreß & Kutscher, 2016, 90f).

While ICT are essential to communicate with families back home, there are also downsides. Unaccompanied refugee girls experience surveillance by families who are concerned about their behavior in Germany (S. Thomas et al., 2018, 61f). Stories of war at home reach refugee children via social media which may cause psychological disstress (S. Thomas et al., 2018, p. 62). Witteborn (2015) found that communication with families exposes refugees to unrealistic expectations (find good employment and send money to Afghanistan), because families are unaware of the hardship refugees experience in Germany. Social media can offer a remedy by allowing refugees to create their own identity and speak out "against restrictive asylum laws and social marginalization" (Witteborn, 2015, p. 363).

Some research focuses specifically on the use of ICT by Afghan refugees. In Canada, Afghan refugee women use ICT to learn English and find accommodation (Quirke, 2012, p. 535). In a study with Afghan refugees in Iran, Jauhiainen et al. (2022, 3;10) noted that Afghan refugees who grew up in Iran, a neighboring country of Afghanistan, had access to ICT in their youth while Afghan refugees born in Afghanistan lacked such access. Afghans living abroad

may thus have more experience with ICT than Afghans in Afghanistan. In a study with Afghan Sikh refugees in India, Pandey and Ilavarasan (2019, p. 335) found that social media increased their social and economic capital. Facebook groups were utilized to share cultural events in the host country that reminded Afghans of their own traditions.

Previous works have emphasized the potential of ICT to support the social inclusion of refugees in host societies. AbuJarour et al. (2019) presented a research agenda for "ICT-enabled integration of refugees." They argued that ICT empower marginalized groups through access to digital services (e.g., mobile health apps), by collaborating with others (e.g., sharing experiences), and connecting them to the public (AbuJarour et al., 2019, p. 876).

In Germany, several apps have been developed with the motivation to ease the settlement of refugees: the App "Ankommen" (English translation: Arrival) by the German Federal Ministry of Migration and Refugees; the App "Welcome" by Heinrich & Reuter Solutions GmbH; the App "Integreat" by Tür an Tür – Digitalfabrik gGmbH (Schreieck, Zitzelsberger, et al., 2017, p. 3). The app Integreat is a frequent best practice example for refugees' ICT-based social inclusion (for instance in AbuJarour et al., 2019). It allows local administrations to provide accurate and timely information in various languages and in a centralized manner to refugees (Schreieck, Wiesche, & Krcmar, 2017; Schreieck, Zitzelsberger, et al., 2017).

Information retrieval was also at the heart of a design probe for a chatbot that answers questions of refugees (J. Chen et al., 2019). The chatbot was developed together with economic migrants in Finland and breaks down information into easily consumable pieces, making information seeking more pleasant. In Australia, Almohamed et al. (2018) developed a prototype website that connects refugees to NGOs. Through the matching via the website, volunteers could assist refugees in questions of citizenship and visa applications (Almohamed et al., 2018, p. 93).

Brown and Grinter (2016) observed that language is a key challenge for refugees arriving in the USA. To assist refugees' interaction with volunteers, Brown and Grinter (2016) developed "Rivrtran" (Refugee IVR Translation), which provides translations between refugees and host families. Translations are made by external volunteers with a bit of delay, suitable for less urgent interactions (Brown & Grinter, 2016, p. 327). Weibert et al. (2019) designed "language wizard," an app which helps refugees in Germany find (free) language classes by easily navigating different types of classes, costs, and eligibility to enroll.

Getting around in their new environment is another challenge for refugees. To assist refugees in understanding accommodation and transportation in the USA, Baranoff et al. (2015)

developed “Lantern.” Based on near field communication, refugees can scan a sticker on a door or wall and receive immediate assistance. Such works focus on the benefits of ICT in addressing refugees’ challenges including information gaps, language, and navigation of new environments. Literature acknowledging the limitations of ICT use for social inclusion is rare.

Among the few works offering a critical perspective on ICT development for refugees’ social inclusion is Alam and Imran (2015), who investigated the relationship between access to ICT and social inclusion. They find that the availability of ICT for integration purposes is not enough. Refugees also need digital literacy skills and the financial means to afford broadband. Keshavarz (2020) criticized the ‘blind’ designing of ICT solutions for refugees while the same refugees are systematically prevented from entering the European Union or deported. He argued that designers lack awareness of the politics behind refugees’ ‘social inclusion’ and called for a justice movement. Critical perspectives on the design of ICT for refugees thus qualify the affirmative discourse that focuses only on the benefits of ICT.

Figure 1: Screenshot of the Dari-language App "Ask for Help"

Methodology

This section presents the methods used in the empirical study with 14 Afghan refugee women between the ages of 20 and 50 in a large city in Germany. The study was conducted in four workshops between March and June 2022 in cooperation with an NGO for the rights of Afghan women. Participants were recruited via the internal communication channels of the NGO including a Telegram chat group, phone, and personal contact. Participants had different levels of education; three participants had a high-school diploma or higher education and three participants did not go to school whatsoever; the other participants had between 3 and 7 years of school experience.

The study used focus group discussions to gather data. Focus groups are considered appropriate methods to generate insights into the attitudes of a group but also to identify disagreement (Bloor et al., 2001, 4ff). Conducting focus groups with ethnic minorities requires adjusting elements of the method to the culture at hand (Colucci, 2008). My colleague at the NGO and I thus made culturally sensitive considerations. Afghan society has a strict segregation of gender in public spaces. Interacting with men outside the family can be seen as immoral (Shahalimi, 2022b, p. 113). The instructor, translator, and participants were all female. Given my German nationality and lack of Dari language skills (Dari is one of the major languagues spoken in Afghanistan), the NGO worker served as a translator and cultural mediator. For refugees, German authoritative figures can be associated with government administrations and the struggle for legal recognition. The presence of the NGO worker, whom participants knew beforehand, provided reassurance. No meetings were scheduled during the month of Ramadan (an important Muslim religious practice). Due to some participants' illiteracy, focus group exercises used visual and oral elements (see table 1).

Exercise #	Name	Activity	Source
1	Pile sorting and ranking	Participants reflected on how ICT services created positive and negative experiences for them.	Colucci (2007, p. 1425)
2	Choosing among alternatives	Participants reflected on their preferences for communicating with an online community or family and friends.	Colucci (2007, p. 1425)
3	Testing the chatbot *Ask for*	The participants downloaded and tested the chatbot *Ask for Help*.	Z. Chen et al. (2020)

	Help		
4	Roleplaying	Two participants enacted a scene where a designer and a tester talk about the chatbot *Ask for Help.*	Colucci (2007, p. 1427)
5	Storytelling	The participants each told one sentence of a story about the development of an app.	Colucci (2007, p. 1426)
6	Drawing a picture	Participants drew a technology that can help Afghan refugee women adjust to life in Germany.	Colucci (2007, p. 1428); Fisher et al. (2016); Almohamed et al. (2020) (Almohamed et al., 2020)

Table 1: List of exercises conducted during focus groups

The focus groups involved the use of the chatbot "Ask for Help," which was developed in an EU-funded research project, where I worked as an ethics researcher. The text-based chatbot was built on top of Telegram and allowed users to ask any question into a community (anonymized). The chatbot's algorithms pick a user from the community who is deemed best suited to answering the question. The chatbot then allows communication and the sharing of expertise in a pool of individuals with similar or different social practices and skills. It can facilitate mutual help with specific tasks or questions. Although the chatbot was originally developed with students, its potential to support refugees should be explored. Z. Chen et al. (2020) showed that chatbots can be helpful tools for migrants and refugees during the orientation phase in a new country. For this reason, the chatbot was translated into Dari and tested with the group of Afghan women.

The study used elements from participatory design. Participatory design is a method that engages the stakeholders (mostly anticipated users) of a technology in its design (Bustamante Duarte et al., 2018). The primary objective of the study was to survey the attitudes of participants towards ICT, not necessarily develop an ICT solution. The prototype chatbot served as a starting point to solicit feedback on ICT design from participants. If the chatbot had been subject to development together with the participants, the study would have required the inclusion of participants in the first stages of participatory design (developing a scenario with participants, surveying needs, and collecting design suggestions, Bustamante Duarte et al.,

2018, p. 4). Participatory design was applied particularly when participants imagined future technologies for Afghan refugee women's social inclusion in Germany.

The analysis was conducted on my own field notes (in English) as well as transcripts of recorded focus groups (which were translated from Dari into English). I used the method of Qualitative Content Analysis (QCA, Mayring, 2000) to identify reoccurring themes that shape the discourse among participants. QCA is a method where empirical phenomena, experiences, and values are identified through qualitative text analysis and connected to theory. I circled words that stood out in the transcripts, color-coded statements according to emerging themes, and connected them by drawing lines to other text parts. The themes that emerged from the material were a) privacy and safety concerns, b) distrust, c) (digital) illiteracy, d) uncertainty and insecurity, and e) preferences for ICT solutions that constitute an alternative to existing designs. These themes were then considered in relation to Black feminist insights about structural inequalities and dynamics of power. The results of this theoretical contextualization of the empirical material is presented theme by theme in the results (section 6).

Ethical Considerations

The study was approved by the ethics committee of the Department of Philosophy at the University of [anonymized]. In preparation of the study, a data protection impact assessment was conducted in line with the European Union General Data Protection Regulation (European Parliament, 2016). The information sheet and consent form were translated into Dari and verbally explained to the participants. All participants signed the consent form, agreeing to the audio recording of focus groups and the collection of data via the chatbot. Despite this, questions remain about the quality of consent. The study took place at the NGO headquarters and an NGO worker was present during all workshops. Since the NGO provides important services to the women (language classes, counselling), consent may have been given out of respect or for fear of losing support from the NGO (Ellis et al., 2007, p. 466).

Beyond informed consent, research with refugees requires further ethical considerations. Power relations between the researcher and refugees are not equal. This inequality puts a moral obligation on the researcher to contribute to the improvement of participants' situation (Mackenzie et al., 2007, p. 301). The study considered this moral obligation in the following ways. The first workshop was dedicated to building trust and a positive working atmosphere. Breakfast was provided in all four workshops. All participants received a gift and a certificate at the end of the study. The nature of the gift (a voucher for a drug store chain) was disclosed at the end of the study to prevent that participants'

socioeconomic precarity coerces them to participate. Finally, the research presented here will be broken down into a one-pager in simple German and Dari and handed to participants.

Biases of the study may emerge from inclusion and exclusion factors. Facilitators at the NGO targeted particularly Dari-speaking Afghan women. The workshop was held in Dari/Farsi (the Iranian NGO worker spoke Farsi which is very similar to Dari except for specific vocabulary), and the chatbot "Ask for Help" was translated into Dari. Yet, Afghanistan is a multi-lingual country with Dari and Pashto as main languages and multiple other spoken languages. Some potential participants may have been excluded because Dari is not their first language. During recruitment, we asked potential participants to bring a smartphone. This may have been an exclusion factor because not all Afghan refugee women have access to a smartphone.

Positionality of the Author and Implications

My positionality as a German citizen has implications for the research and its results. I am an outsider in the group of Afghan refugee women. I was also perceived as an authority figure during the study, which means that interactions between participants may have been adapted to my presence. Interpretation of the data was done without consultation of the participants due to resource constraints. This poses the risk of misinterpretation. Moreover, due to my own privilege, I am unable to grasp the lived reality of Afghan refugee women. These constraints may lead to complicity in colonial practices of knowledge production about Afghan refugee women.

Following Lazem et al. (2022, p. 167), this paper falls into the category of a Western postcolonial approach rather than a decolonial practice. While the research sheds light on an underserved community, it speaks *about* Afghan refugee women and to a Western audience. As such, it does not actively counter Western colonial ways of knowledge production. I decided to pursue the research anyway to drive attention to the experiences of marginalized users of technology (Ellis et al., 2007, p. 466).

Left: Figure 2: Results of focus group exercise "pile sorting and ranking"

Right: Figure 3: Results of focus group exercise "choosing among alternatives"

Figure 4: Drawing: A car invisibly crosses borders

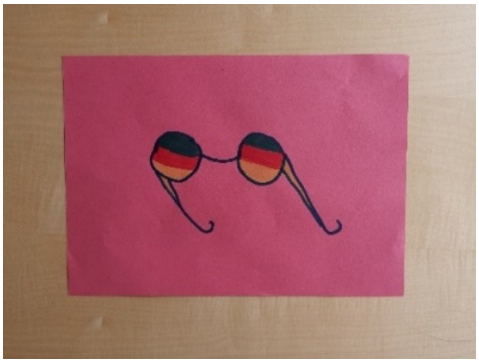

Figure 5: Drawing: Magic glasses

Results

Safety and Privacy Concerns

One theme emerging from the qualitative data was Afghan refugee women's concern for safety and privacy, especially safe communication with relatives in Afghanistan. In the exercise "pile sorting and ranking," participants were asked to sort pieces of paper with the logo of an ICT service to their emotions towards those services. This exercised helped them express their attitudes towards different apps (see figure 2). Many "loved" WhatsApp and imo (the latter was explicitly mentioned although it was not included in the given examples). One participant said that WhatsApp allowed her to communicate with her family and friends in Afghanistan. Several participants agreed but they also expressed the value of security. One participant stated: "We

like WhatsApp because it has good security. A WhatsApp account cannot be hacked easily." The group asked me about the likelihood that governments or individuals hacked their phones. They wanted to know how they can protect their phones and which mobile brands are the 'safest'. Participants' concerns about safety may be attributed to their experiences with persecution by militant groups. They may be afraid of the Taliban or other groups surveilling their families. This extends previous findings that point to refugees' awareness of government surveillance and abuse by malignant groups through ICT (Mancini et al., 2019, 8ff). In their interaction with ICT, Afghan refugee women may then want to rely on privacy-aware services but lack expertise in navigating the landscape of privacy affordances of different apps. Here, digital literacy training could help empower Afghan refugee women to protect their own and their family's privacy.

Another reason for valuing safety and privacy in online communication may be gender norms, as they prescribe `appropriate´ behavior for women in the physical world but also online. In Afghan society, it is considered inappropriate for women to interact with men outside of the family. Safety and privacy may prevent unsolicited contact from strangers and avoid conflicts in the family. One participant of the study complained that she was contacted by a stranger via Instagram. She was concerned that strangers could see her private photos. She asked how she could prevent strangers from contacting her on Instagram. Another participant shared her scepticism about the true identity of other users: "I'm saying, it is possible that someone writes to you and says they are this person. But they are lying in this moment and they are not this person." This reflects the fear of interacting with a fake profile.

Participants were also aware of the real-life consequences of privacy violations. One participant shared a negative experience with Facebook. Back in Afghanistan, she wanted to share a photo of her new haircut with a friend. She didn't understand that posts are seen publicly unless settings are changed, and she accidentally sent the photo to the public. Given the norm of wearing a headscarf in public, the accidental posting caused severe conflict in her family.

These findings confirm previous research. N. Ahmed et al. (2022, p. 3) highlighted the difficulty of Afghan women to juggle online visibility with family expectations and their own protection against sexual harassment. Especially girls are affected by the family's scrutiny (S. Thomas et al., 2018, 61f). In this context, N. Ahmed et al. (2022) discussed the challenge of designing ICT solutions for Afghan women. If conservative gender regimes motivate Afghan refugee women's interest in privacy, then culturally sensitive design in this case reaffirms patriarchal systems (N. Ahmed et al., 2022, p. 5).

Distrust

Another theme identified in the qualitative data was Afghan refugee women's distrust of online communities. This is reflected by the quote above on the threat of fake profiles. In the exercise "choosing among alternatives," participants were confronted with a topic and had to decide whether they would ask a question about this topic to an online group or to their friends and family in person (see figure 3). The topics ranged from finding employment and finding an Afghan restaurant to advice about a child's illness or getting married. The group disagreed in this exercise. Depending on the participants' prior experiences, they had higher or lower standards for trust. In general, though, participants were more on the sceptical side towards online communities. One participant expressed value in having a social media group consisting of her many friends and relatives. Given the large size of Afghan families, her question would reach numerous and trusted individuals. Another participant argued that a big group with unknown people always meant a lot of answers, which can be confusing. Asking just one (highly trusted) person guaranteed a straighforward and reliable answer.

For sensitive topics like marriage, one participant questioned other users' intentions in a bigger chat group: "There are some people that you don't know, and they don't want good things for you and will betray you behind your back." Another participant questioned the efficiency of asking about marriage in a big group: "Everybody has their opinion, and it will not have a good result." Given participants' scepticism towards asking questions into a chat group, especially for personal issues, chat groups may not be appropriate for seeking advice.[38] In-person interaction in smaller groups or one-to-one may be more valuable. This confirms previous research by Quirke (2012, p. 536) who reported in her study on Afghan refugee women in Canada, that "the most trusted settlement information sources are people [instead of technology], including family, friends, teachers and settlement workers." Hence, in-person contact when asking for help may be more attractive to Afghan refugee women than ICT solutions.

(Digital) Illiteracy

A relevant issue to the group of Afghan women was (digital) illiteracy. This became clear during the testing of the Telegram-based chatbot "Ask for Help." The onboarding process that allows users to register in the chatbot was too complicated and caused frustration. The onboarding consisted of multiple steps: a) locating the chatbot in Telegram, b) following

[38] These findings may not necessarily be specific to Afghan women but apply to various individuals who value their privacy. Nevertheless, the findings have implications for attempts to design ICT for social inclusion, as designs involving a bigger chat group may not be adopted by Afghan women.

English-language instructions to register an account in the website of the EU project that was developing the chatbot, c) checking the email inbox (spam folder) and clicking the link therein for confirmation, d) filling in a profile, e) returning to the chatbot and accessing a link to agree to the data processing guidelines. Although the chatbot itself was translated into Dari, the user had to register an account following English-language instructions *before* being able to specify the language in their profile. None of the participants had English-language knowledge. In addition, some of the participants did not own an email address. The NGO worker and I registered email accounts for them so that they were able to use the chatbot at all. One participant expressed her frustration: "I'm talking for myself! I actually cannot use this. I cannot download or upload these requirements for this chatbot. I don't have an email address or Gmail. Then how can I ask for your help [in the chatbot]? You are doing all this for me."

Difficulty in navigating the chatbot partly stems from participants' lack of (digital) literacy, which can be attributed to the violence-induced lack of education in Afghanistan. Digital illiteracy is also observed in Steinbrink et al. (2021) who demonstrated that Afghan refugees in Germany have particularly low digital literacy compared to other nationalities. However, the design of the chatbot also *produced* barriers. The Latin alphabet, English language, and text-based communication in the onboarding process made usage of the chatbot difficult for participants (cf. Warschauer, 2004, p. 204). The design of the chatbot "Ask for Help" reflects designers' bias that users' preferred mode of communication is through written exchange. Dell and Kumar (2016, p. 2226) emphasized the importance of moving from text-based designs to voice and graphics when designing for low-literate populations. Almohamed et al. (2020, p. 81) also stress the advantage of having voice-based assistants for refugees to retrieve information and book appointments.

Another issue that came up in the study was the accessibility of the chatbot for Afghan refugee women who had just arrived in Germany. In a roleplaying exercise, participants enacted a conversation between a designer of the chatbot "Ask for Help" and a user. One participant playing the user observed: "When somebody is new to this country, she doesn't know about this chatbot." Another participant said: "There are times that the person does not have a Simcard." Several ideas were raised to circumvent these barriers: the chatbot could be pre-installed on any phone and immediately usable; it could be advertised at the immigration office and in reception camps; a phone having the chatbot installed could be shared with newcomers; a web version of the chatbot could be accessed via an Internet café or NGO computer. Hence, creativity in tapping the sharing economy and cooperating with technology providers, administrations, and public services might reduce barriers to accessing ICT.

Uncertainty and Insecurity

The study revealed that participants had a strong feeling of uncertainty and insecurity when interacting with the chatbot. One issue was uncertainty about the peers' availability. Not knowing whether another user would see and respond to a question sent to the chatbot made participants feel vulnerable. The chatbot matches the question of user A with a competent user B, who answers in their own time. A must then wait for B to answer the question. This delay in answering a question clashed with Afghan refugee women's everyday needs for real-time assistance like navigation. One participant described the problem: "I have moved to this country recently and I want to find an address. To find this address, I want to use my phone and this chatbot to ask for your help. Unfortunately, I will wait and wait for 5, 10, and 15 minutes and I don't receive any answers. Now I'm waiting in a square in the city and don't know what to do." The discontent with a service that offers delayed communication can be attributed to language barriers and cultural unfamiliarity with German infrastructures. While most individuals may prefer a quick answer to a question, Afghan women are dependent on real-time digital assistance in the case of navigation as they cannot ask a person in the street for help due to language barriers. Participants criticized that the chatbot design does not allow the user to see if B has already seen the question (like they can in other app services that provide a read confirmation via tick marks). The participants stressed that they preferred other online services (like WhatsApp) over the chatbot for urgent tasks. An ICT solution for Afghan refugee women's social inclusion should thus offer reassurance that the community of peers is reachable at an instance.

Another issue was insecurity caused by confusion around the features of the chatbot (also see previous point, "digital illiteracy"). Participants were unable to register in the chatbot and later had difficulty navigating the app, which caused frustration. In previous research, Talhouk et al. (2019, p. 1587) find that refugees experience frustration with failing technology. If technology does not work properly or refugees note that they are unable to use it, their self-confidence decreases and their motivation to participate in finding design solutions may be compromised. Given Afghan refugee women's experiences of uncertainty in critical areas of their lives (the survival of their family in Afghanistan, the outcome of their asylum application, financial insecurity), experiences of failure with the chatbot may compound psychological distress. ICT for Afghan refugee women should thus foster (even the smallest) experiences of success to support well-being and confidence.

Preferences for an ICT Solution for Afghan Refugee Women's Social Inclusion

Throughout the study, participants voiced their preferences for ICT-based solutions that support Afghan refugee women's integration into German society. The discussion of the shortcomings of the chatbot "Ask for Help" led to suggestions for improved design features. They are summarized and linked to their underlying value in table 2.

Supported value	Design feature	Explanation and examples
Reassurance	Real-time assistance	The ICT solution connects users to real people in an instance who can asssist them with (urgent) tasks
	Read confirmation	The ICT solution makes transparent if a request or message was seen by the respondent
Safety	Trusted respondents	The ICT solution connects users to trusted individuals, e.g., family and friends, volunteers, NGO workers
	Good privacy and security	Conversations in the ICT solution must be private, outsiders cannot contact the user
Independence Self-efficacy	Access without the help of others	The ICT solution can be set up without the help of others
	No English language, no Latin alphabet	The ICT solution uses arabic letters (if there is text at all) and the native language of users
	Audio feature	The ICT solution should be audio-based – or, if necessary – text-based and offer an audio option
	Easy to use	Any commands or features must be intuitive
	Straightforward registration and simple interface	The registration process must be straightforward; the ICT solution does not require an email address to create an account
	Experience of success	The ICT solution promotes moments of success through simple navigation and good results
Well-being	Friendly communication	The ICT solution ensures that communication is polite and respectful
	Positive feedback	The ICT solution provides reassurance when the ICT was used correctly to reduce insecurity of the user; “you did great”, “this worked”, “well done”

Table 2: Design features for an ICT solution that can support Afghan refugee women’s social inclusion

In a “storytelling” exercise, participants each completed one sentence of a story. In this story, an Afghan designer from Kabul was researching the needs of Afghan women for an app. Participants came up with the following features for an app: communication like (video) chatting and phone calls, but the conversations should be kept private; trustworthiness, “so that

the girls feel comfortable to talk in it;" German language training; informational and educational services; entertainment. In the exercise "draw a picture," the participants imagined real or fictional technologies for Afghan refugee women. Figure 4 shows a magic car that becomes invisible when crossing a border. Asylum seekers are not allowed to leave Germany until their case is decided. The car allows refugees to still visit families and friends across Europe. The magic glasses (figure 5) allow Afghan women to read any document in their native language. Moving through foreign countries and reading road signs but also reviewing letters from the job center in Germany, the glasses enable real-time translation.

Technologies closer to existing ones were also imagined. One participant developed the idea of an "Afghan app," which focuses on translation. Another participant extended this idea to include non-Afghans in the app. She imagined "a public or global app with different languages and like YouTube, with different pages for different purposes: for example, 'doctor page' [to find medical advice]." Such a general app should also help users find a lawyer, learn German, navigate abroad and in Germany, and find friends. One participant imagined an app "like a psychotherapist so that I can share my problems with [it]." However, not all imagind solutions were technology-based. One participant imagined an NGO from Afghan women for Afghan women: "[…] apart from translating the letters from the job center, we could help to solve more serious and fundamental problems of these women." The participant's idea stresses the importance of programs that are not ICT-based but rather connect Afghan refugee women with each other in person.

Discussion

Information and communication are key to social inclusion because they are the tools to bridge divides between the newcomer and the host society in terms of cultural and social practices, resources, and skills. Social inclusion is at risk when newcomers lack the information practices required to participate in communication (Lloyd et al., 2013, p. 122). The study results in this paper suggest that Afghan refugee women are challenged by navigating the digital mode of information and communication in Germany. Challenges in accessing ICT can be experienced by numerous groups including German elders, but reasons for a lack of access may differ. In the case of Afghan refugee women, experiences with technology are shaped by culturally specific gender configurations, their persecution by militant groups, and a severe deprivation of education. These particular experiences should be considered when designing ICT-based solutions for Afghan refugee women's social inclusion, as centering marginalized communities in design can be a step towards social justice (Costanza-Chock, 2020).

However, it would be short-sighted to merely focus on improving one-off design solutions. Rather, the design of Germany's migration management infrastructure produces barriers to social inclusion. Refugees must rely on the use of digital services to receive basic services. In some cities in Germany, making an appointment at the immigration office requires the use of a webbrowser, an email address, and filling in a long form. Sabie and Ahmed (2019) argue: "The political values that a state embeds in its infrastructure thus determine the quality of access of migrants to the critical services of their life" (Sabie & Ahmed, 2019, p. 219). The (infra)structural exclusion (as it presents itself for Afghan refugee women) is further exacerbated by the common narrative that refugees are heavy users of technology and benefit from ICT (AbuJarour et al., 2019; Fisher, 2018; Patil, 2019). To counter their exclusion, large-scale digital literacy classes need to be implemented for Afghan refugee women during their orientation in Germany.

This said, providing large-scale digital literacy training would present a one-sided integration of Afghan refugee women into German society. Afghan refugee women must adapt to German communication practices in a one-sided manner. Social inclusion is all too often seen as the adoption of White, male, and Christian practices by the cultural 'other' instead of coming from both sides and meeting in the middle (Berrey, 2015, p. 3). An alternative approach would be to reconfigure power relations and redistribute resources to ensure that those who are marginalized can become full members of society. In this vein, in-person services for low-literate refugees should be considered to foster social inclusion. These in-person services must be provided in a welcoming manner and defy any discriminatory treatment. Previous research suggests that in-person services can increase trust (Quirke, 2012, p. 536) and be meeting points for Afghan refugee women to connect with each other and socialize (Warschauer, 2004, p. 4).

Outlook

This paper has investigated the experiences of Afghan refugee women in interaction with technology. Ultimately, it has raised the following issues: While ICT solutions can support the social inclusion of Afghan refugee women in aspects like communication with peers and NGOs as well as navigation, they are not the only and not necessarily the most effective ways to foster inclusion. This is because experiences of marginalization, including a lack of (digital) literacy, rigid gender regimes, and the fear of persecution by militant groups shape Afghan refugee women's needs for trustworthy and simple-to-use ICT. To ensure that Afghan women can take advantage of ICT designed for refugees, the design must center their needs, for instance by offering audio features, intuitive commands and navigation, real-time assistance, and

registration without an email address. In addition, digital literacy classes are required to help Afghan refugee women adapt to the increasingly digital communication practices prevalent in German society, especially when it comes to interaction with administrations.

Finally, it is crucial to offer in-person services to Afghan refugee women and other low-literature refugees (and even low-literate German citizens, for that matter). In-person services can meet expectations of trust and safety but also allows for in-person contact between the host society and refugees as well as among the refugees themselves. In thinking about how German society may develop new strategies to meet the growing demand for social inclusion of newcomers, three avenues for social inclusion should be followed: ICT services that center the needs and assets of refugees in their design, free of cost digital literacy training, and in-person services based on trust and kindness. This three-fold approach can meet the diversity of practices and experiences among refugees and be a comprehensive blueprint for social inclusion with and without technology.

Chapter 8: The Contributions of Human-Computer Interaction to Rethinking Design: US-American and European Perspectives on Critical Diversity and Design

[This article was originally published in: Schelenz, Laura (2025): The Contributions of Human-Computer Interaction to Rethinking Design: US-American and European Perspectives on Critical Diversity and Design. *The Design Journal*, 1–18. https://doi.org/10.1080/14606925.2025.2573710.]

Design today is reckoning with the impact of social inequalities, historical and continuing injustices, and the unconscious reproduction of abusive power relations in the United States and Europe. This is the case for technology design, urban planning, and broader designs for our societies enacted by government policy and development work (Costanza-Chock, 2020; Dell & Kumar, 2016; Escobar, 2018; Fox et al., 2017; Goh et al., 2022; Piazzoni et al., 2022). In this moment of reckoning, works engaging a 'critical diversity′ perspective have become sources of inspiration for rethinking mainstream design. 'Critical diversity' refers to a research agenda that considers structural inequalities and questions of in/justice. The term is inspired by Critical Theory and Critical Diversity Studies (Anderson & Middleton, 2018; Collins, 2019; Steyn, 2015, 2018). One area of design geared towards computing and technology development – human-computer interaction (HCI) – has been particularly productive in developing alternative approaches that are critical of mainstream design (see for instance J. Bardzell & Bardzell, 2013). This paper investigates contemporary discourses in HCI that engage a critical perspective on question of diversity and design in the USA and Western/Northern Europe.

HCI sits at the intersection of computer science, (service) system design, and technology development, dealing with the design of user interfaces, user experiences, and the needs of diverse communities in their interaction with technology (see for instance the conferences ACM CHI, NordiCHI, and CSCW). Human-robot interaction is closely related to HCI as both fields are 'working with an interdisciplinary approach, target a continuously changing technology, and aim to improve the usability of interactive systems' (Huang, 2015, p. 2). They share similar concerns, too, such as gendered stereotypical representations in prototypes (Winkle et al., 2023). The area of design centred in this paper is thus a particular realm of technology design, namely, the design of interactions between humans and computers or robots.[39] The focus on HCI was chosen as this field has not only heavily engaged with critical perspectives in recent

[39] In the following, I will refer to HCI solely as it is the earlier and more encompassing field.

years but has further shaped critical perspectives in the broader design community. For instance, urban planners have referenced ‘design justice’ as an inspiring framework, which stems from the work of Costanza-Chock in technology design (Costanza-Chock, 2020; Goh et al., 2022, p. 6).

The USA and Western/Northern Europe are comparable because they share a tradition of liberal democracy, the value of a free economy, and a rhetoric of human rights. This said, the USA and European countries are situated in different histories and political, socio-economic, and cultural experiences. The US-based HCI discourse on critical diversity and design is reconstructed *via* a literature review that includes 82 papers retrieved from a systematic search in the ACM Digital Library. The European discourse is less established and has been reconstructed *via* expert interviews with researchers and designers based in Europe. Subsequently, materials were analysed and coded using qualitative content analysis (QCA) and compared according to a priori defined categories: predominant themes, critiques, understanding of design, theories, and methods.

US-based HCI discourses on critical diversity and design offer a shift in perspective away from mainstream design practices. Researchers and designers promote a structural analysis and critique of power relations as well as bottom-up and participatory approaches. Similar tendencies are observed for the European context with differences in priority. Designers and researchers in Europe stress the importance of collaboration with the public sector including government administrations, and they focus on gender rather than race in their treatment of individual diversity ‘categories.’ Implications will be discussed for a transatlantic critical discourse on diversity and design in HCI and design.

Materials and methods

This paper looks at perspectives that are critical of design approaches from the vantage point of HCI, and analyses how these perspectives compare in the USA and Europe. Embodied and disembodied technologies have been subject to critique in design (J. Bardzell & Bardzell, 2013). An explicitly ‘critical’ diversity and design perspective questions established power relations among designers, in design and the design industry. It may follow the tradition of Critical Theory, an ‘approach to social, political, and cultural problems that aimed to join theory and empirical research’ (Collins, 2019, p. 59). The paper’s analysis centres discourses on critical diversity and design. A discourse is ‘a collection of narratives or accounts about something, such as a historical event or political struggle. Discourses of various kinds circulate as people describe and interpret these things’ (Hirst et al., 2023, p. 63).

US perspectives on critical diversity and design were reconstructed *via* a literature review following a search in the ACM Digital Library. The American Association of Computing Machinery (ACM) is the world's largest outlet for academia on computer science and informatics; its library contains the publications from all ACM conferences and journals. While it has a global reach, the ACM can be characterized as highly influenced by US academic traditions. As part of its ecosystem, ACM has several highly regarded communities on HCI, all represented with annual proceedings in the ACM Digital Library. The literature search included the terms 'human-computer interaction' AND 'design' AND 'critical' or 'justice' or 'diversity.' The search was restricted to academic papers before 2023. Individual relevant publications from previous research were added to the body of materials. All in all, 82 publications were selected for analysis.[40]

The European literature does not reveal an emergent 'field' of critical design in HCI to the same extent as the US-based literature, although there are increasingly critical perspectives (see for example Abdulla, 2014; Ober, 2023). For a reconstruction of the European discourse, I conducted in-depth interviews with 16 designers and researchers from nine different countries who held a position at a university, company, or non-profit organization in Western/Northern Europe.[41] The interviews were semi-structured to allow interview partners to shape the conversation, which means that questions were adapted depending on the conversation flow and prior knowledge of the interview partner (IP) (Brinkmann, 2020; Carol A. B. Warren, 2002). The general interview structure contained questions on 1) the role and current activities of the IP, 2) a current (technology) design project of the IP, 3) a reflection on the user group(s) relevant to this project, 4) how the IP gathered knowledge about users, 5) the IP's understanding of the experiences of 'privilege and oppression' of the users, 6) the social impact of the project, 7) the IP's motivation to create social change, 8) the IP's understanding of diversity and 'diversity-aware technology', 9) the IP's understanding of 'intersectionality'.

Data was collected between November 2021 and March 2023; three interviews were conducted in person and the rest online due to the Covid-19 pandemic. In total, I collected 382 min of conversation. All interview partners signed an informed consent form ensuring them rights and protections under the EU GDPR. Recordings were transcribed by hand and analysed using Qualitative Content Analysis (QCA) after Mayring (2022), using the online tool QCA

[40] The list of materials representing the US-based discourse is available at the following link: https://drive.google.com/file/d/1XXqLVwYLvGlIWaMHV46tTwYSPLOxh0Op/view?usp=sharing

[41] Since data on nationality of interviewees was not collected, the designers were not necessarily *from* a European country. However, their status as employees at a European organization suggests their engagement with European colleagues and discourses.

map.[42] Five categories were determined a priori to guide the systematic analysis of both the literature review and expert interviews: predominant themes, critiques, understanding of design, theories, and methods. An example of the coding is presented in Figure 1. The material was color-coded according to the five categories above and marked with a brief summary of the content in a separate file.

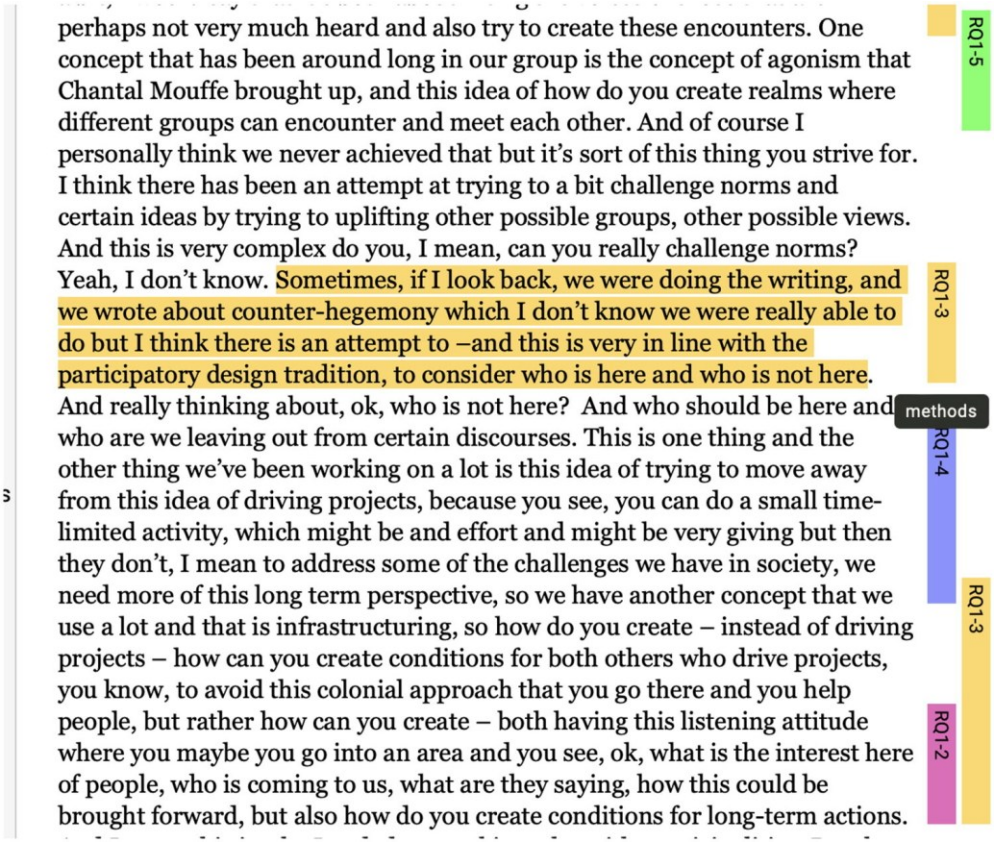

Figure 1: Coding example

As any study, this one also contains limitations. First, researchers based in Europe contribute to US publication outlets and vice versa, which makes it difficult to have a clear-cut identification of a 'regional' discourse. Yet, the contributions of designers and researchers to a particular discourse may not reflect their heritage but rather their socialization in a certain academic or practical tradition. European researchers and designers who publish in US outlets may thus be socialized in US-dominant narratives. A second limitation is that interview partners were based in Western and Northern European countries, which results in a blind spot regarding Central and Eastern European perspectives on critical diversity and design.

[42] https://www.qcamap.org/

A third limitation relates to the comparison of a literature review and expert interviews, because different epistemologies underlie these two methods. Literature is an outcome of a longer engagement with preconceived questions, whereas records from interviews represent spontaneous reactions to previously unknown questions. It could be argued that a literature review offers a deeper and more reliable understanding of a discourse than exploratory interviews. Given the diffusion of critical perspectives in European HCI, expert interviews seemed a reasonable compromise to produce a more structured picture of the discourse.

Results: Perspectives on critical diversity and design in comparison

The results from the comparative analysis will be clustered into overarching topics that guided the systematic analysis of the material. Table 1 shows the main features of the US-based and European discourses in the five categories with the most important feature described below.

Category	**US-based discourse**	**European discourse**
Pre-dominant theme	Diversity categories and classifications	The optimization of public services for marginalized groups
	Algorithmic bias and discrimination	Diversity features of embodied technology, bias and stereotyping
	Lack of diversity in the technology industry	Health, well-being, and inclusion of people with mental or physical constraints
	The role and positionality of the designer	
Critique	Lack of awareness among designers of structural discrimination in design	Lack of diversity in design; lack of awareness among designers about unequal power relations
		Lack of awareness of designers about social and political entanglements in design
	Modernist orientation in design	Modernist orientation in design
	Academic structures	Academic structures
	Neoliberalism in design	Neoliberalism in design
Under-standing of design	Design as deeply entangled with society and politics	Design as a participatory, democratic process
	Bottom-up activity of resistance	Design as a tool to challenge widely held assumptions
	Design as a process that is open to constant improvement	Collaboration with the public
Theories	Black feminism, Muslim feminism	Feminism, posthuman feminism, postcolonial feminism
	Intersectionality	Intersectionality
	Critical Race Theory	Critical Disability Studies

	Theories of justice and power	
Methods	Participatory Design	Participatory Design
	Bottom-up everyday 'here and now' practices	Policy engagement, institutional change
	Community engagement	Integrated/Embedded ethics
	Reflecting positionality	Reflecting positionality
	Speculative design and future imagination	Everyday people seen as design experts

Table 1: Main features of the US and European discourses on critical diversity and design.

Predominant theme

The US-based discourse in HCI that deals with critical diversity focuses primarily on diversity categories and the study of bias. Usually in mainstream technology design, classifications of 'difference' are encoded into technology to optimize the technology for different users with little regard for the risk of discrimination (Benjamin, 2019b, 17ff; 74). Critical work on diversity categories covers the binary conceptualization of gender in user representations (Schlesinger et al., 2017, p. 10; Stumpf et al., 2020) as well as stereotypical personae (Hankerson et al., 2016, p. 478). Other case studies have pointed to the systematic disadvantaging of Black and resource-constrained communities through biased algorithms in relevant applications (Eubanks, 2017). Another thematic focus is the diversity of positionalities, i.e. the experiences of privilege and discrimination of design researchers and practitioners, which also relates to a methodological approach and is discussed further in the subsection 'Methods'.

Like the US-American context, the European discourse shows interest in diversity features and the biases that emerge. IP11 is interested in the gendered design of robots:

> So what kinds of stereotypes occur when someone sees a robot that is completely human-like and then looks more like a woman than a man, right? And then I want to understand whether there are biases and discriminations that occur when we are interacting with a female-like agent but not human but rather robot.

IP15 is 'work[ing] with communities to identify the biases, stereotypes, and social-technical entanglements that we have with AI and specifically embodied AI'. The aim of this project is to develop methods that can help reflect on possible bias and improve the design of embodied technology. The most prominent theme in the European discourse, though, is the optimization of public services for different social groups. IP1 designs solutions for migrants who lack access to digital government services. Language is described as a barrier in making services available to migrant populations in different European countries. IP2 designs a chatbot for students at

their university: ‘We had a clear idea also from the user research that we did that students here have very often more, let’s say, psychological challenges, and things that are very difficult to talk about’.

Another important topic in the European discourse concerns the health, well-being, and inclusion of people with mental and physical disabilities or illnesses. IP9 designs socio-technical support systems for people living with HIV/AIDS and stresses the importance of focusing on their lived experiences:

> Quickly I think we kind of took a step back and used design, or critical design also, and participatory methods as a way to understand better the context and people’s lived experience, which is very complex. And it’s not just about designing yet another app [but] using design to understand the kind of challenges for them, to share personal health data, to live with a condition like HIV on an everyday basis.

IP16 looks at accessibility of data visualization for people with disabilities: ‘My research projects [are about] how does a crip technoscience view change the way we approach and design data visualization to make them accessible for people with visual impairments’. The European discourse on critical diversity and design thus heavily focuses on services for the marginalized, especially people with mental and physical challenges. The US-based discourse tends to centre questions of race and gender and encourages reflection of designers’ positionalities, but also shows awareness and activism around disability (Spiel, 2022), neurodiversity (Spiel et al., 2022), and crip technoscience (Hamraie & Fritsch, 2019).[43]

Critique

In the US-based discourse, designers and researchers have criticized a lack of awareness about structural discrimination in and through design. Benjamin (2019a, p. 8) argues that oppressive race relations are encoded into technology. D’Ignazio and colleagues see structural discrimination in the way that data is collected and used (D’Ignazio & Klein, 2020a, p. 438). Another concern is the domination of ‘White and Asian, white-collar, educated, urban’ (Tran O'Leary et al., 2019, p. 3) as well as male, heteronormative, and ableist perspectives in technology design (Wachter-Boettcher, 2017, p. 20). While the European discourse similarly reveals criticism of a lack of awareness about structural inequalities, the focus is less on structural racism per se. For instance, IP14 criticizes the inequality of power between designers and users overall:

[43] The exchange between European and US-based discourse may be most visible in the case of critical disability or critical access studies as authors contributing significantly to the US-based discourse are based at universities in Europe, such as Katta Spiel.

> I think this idea that there is someone who just uses things and that is designed based on an abstraction of their needs and characteristics in a way embodies a form of privilege from designers to have that knowledge. So I think we need to challenge that. We need to challenge the idea that the experts are on one side, and on the other side you just have someone who gets results.

IP 6 is conscious of the need to include various cultural backgrounds in designing technology-mediated experiences in memory institutions, yet noting that even then, inequalities in the representation of memory remain:

> So we worked with universities in [city A], in [country A], and universities in [country B], in [city B]. In [city A], the students are international, so they come from other Balkan countries that are not in the European Union […] we were looking here at that type of diversity, which is very different and not necessarily enough because for example we don't have what it means to have a European memory from the perspective of people with different abilities.

European designers also call for more consciousness about the political nature of design. IP4 calls for the amendment of computer science curricula to raise awareness about the social impact of design:

> I would like this part to be inserted in a technical university curriculum. So not only learning about the optimizing [of computer systems] and treating everything in numbers but also understanding what you are making has a social impact, a direct social impact on everyone.

IP13 criticizes that designers rarely think about the social impact of their products:

> [Designers] don't see themselves as social changers. I mean some of them do, of course. But yeah, in general I would say no, not really. Not even considering like environmental impact either. Because they sort of think that this is digital so it's free from the world somehow.

The need to be more conscious of political processes in design is not articulated in the same way in the US-based discourse since design is already understood as inherently entangled with politics (see section 'Understanding of design').

Further critique in both the US and European discourses relates to Modernism, solutionism, determinism, and the neoliberal orientation of mainstream design. Erete et al. (2023a, p. 5) argue that ordering processes of efficiency, e.g. bureaucracy, can hide racist and sexist practices of surveillance. Further critique relates to the perception that technology is value-neutral (Benjamin, 2019b, p. 41). In the European discourse, there is a concern especially about the neoliberal orientation of design practices with diversity at risk of appropriation, according to IP11: 'You always have the danger of people using and misusing [your] findings [of diversity-related research] to try to like pinkwash their companies, or rainbow-wash their companies.' IP13 recognizes the pressure in the design industry to '[bring] in the bucks' rather than focus on design for social good.

There is critique in both discourses of the academic structures in which design is developed. The US-based ACM has been called out for unequally distributing leadership positions and discouraging engagement with racism in academic work (Erete et al., 2020, 10, 33f). In the European discourse, similar observations were made. IP12 has experienced that publishing a critical perspective in computer science is difficult:

> So, in my first paper [of my PhD project], I had the struggle of everything had to be framed in ways that are targeting computer scientists, so they are not threatened by it […] In the third paper that I wrote, which is this autoethnographic paper or essay, I definitely use the words "privilege" and "ableism" and "sexism" and "cis normativity." So, I use these words and I don't care if the computer scientists are threatened by them because this is the theory that we need to use in order to really understand the problem.

In addition, IP7 explains that they have been careful not to take a too strong activist position because of the tenure-track system in academia, which requires constant evaluation of young professors: 'As an assistant professor, like in [country], tenure-track is pretty safe, but there is the tenure-track, and so I tried to make sure my work is both sides [activism and research]'.

Understanding of design

The US discourse treats design as deeply entangled with social and political processes. Costanza-Chock (2020, p. 4) invokes Langdon Winner's famous quote that 'artifacts have politics', pointing to the fact that design is never value-neutral. Carroll (2017) considers design an invisible force that creates structures including systems of inequality: 'You know that design is effective when it's not even there.' Design is also associated with speculation and resisting mainstream contemporary ways of being (S. Bardzell, 2018). Specific to the US context is that design is tied to community building and autonomy. Dombrowski (2017) links design to social justice, and S. Bardzell (2018) advocates an understanding of 'social design', where design is centred around the real needs of humans.

The European discourse largely perceives design as a participatory, democratic process, where the voices of different stakeholders are heard. IP5 talks about this in the context of robotic design solutions: 'What's important? What they [the users] think a robot could appropriately do? […] And yeah. Just taking their opinions into account and trying to help shape what we are doing with that'. IP7 who also works in human-robot interaction says: 'We should have diverse research participants because we are building robots for diverse people and everyone's voice should be heard.' IP14 stresses the importance of considering the diverse experiences and practices of everyday people in technology design:

> So if one of my arguments is that we need to better understand lived experiences and practices of technology as [they] happen, then obviously we need to be open to all lived experiences and

> practices. So let the diversity of practices be documented, be part of how we think of case studies, how we think of design concepts.

In both the US and European discourses, technology design is considered a community effort. Yet in Europe, designing with communities is done through collaborations with the public, especially institutionalised governmental and non-governmental actors. IP9 works with a non-profit organization that represents their stakeholders (people living with HIV/AIDS):

> The project was actually co-defined from the bid stage with charities, like HIV charities, and it's funded by [national institution], and the PIs [Principal Investigators] put a lot of work from the start to craft that space in the, in the proposal for having a peer researcher, so someone who lives with HIV.

Similarly, providing policy recommendations at the institutional level is important to designers in Europe. IP1 states that their goal is to institutionalise social innovation to achieve long-lasting impact with design:

> I mean what I am interested in is creating social innovation at the niche level but also how to generate mechanisms of institutionalization of social innovation, when what you have in the niche, you know, becomes part of the governance structure, policies, or institutions, and so on.

Such efforts place design at the centre of public engagement and tie it to formal structures rather than local, autonomous practice.

Theories

Most prominent among the theories promoted in the US discourse are different streams of feminism (Stumpf et al., 2020, p. 10). Black and Islamic Feminism are introduced as frameworks for the critical study of technology design (Erete et al., 2023a, p. 16; Rabaan & Dombrowski, 2022). Intersectionality is frequently advocated as a tool to make visible how power is configured in design processes. Ogbonnaya-Ogburu et al. (2020) introduce Critical Race Theory as a lens in human-computer interaction. Theories in the US-based HCI context further include Future Studies (Kozubaev et al., 2020), Activity Theory (Wong-Villacres et al., 2020), Afrofuturism (Winchester, 2018), and Black Utopianism (Tran O'Leary et al.).

In the European context, feminism and intersectionality have been noted as significant influences, e.g. in the works of IP4, IP9, IP10, IP11. IP4 states: 'So, in our workshop that we are doing, we mention a lot intersectionality, because we don't have just one identity, and women or men or whatever, we're not a homogenous group.' IP9 also reflects the different experiences of their research participants in an intersectional way:

> I can picture participants in my mind, I'm just really inspired by them, but yeah, even like a White gay man who lives in [the capital] will be very different from a mother here [in the North

> East of the country], who is from a black ethnicity, you know, and yeah, with two kids at home, living on her own.

A notable difference to the USA is that European perspectives invoke race as a diversity concept much less. With regard to human-robot interaction, IP5 observes: 'I think a lot of [the research] is focused on gender. There is not, in terms of intersectionality, there is not so much maybe with race or disability.' The lesser attention to 'race' is explained by IP7 who invokes limitations of dealing with race in computer science in Europe due to national data laws and practices:

> So, intersectionality is huge, right, from a feminist theory standpoint, intersectionality you cannot, you cannot get away from it. And right now, I'm really struggling with the fact that like, for example, [country] does not collect racial statistics, it's quite, they will collect like, you can know if someone is [native] or not for example, but there is no like Black, White, Hispanic, in the way that we do in the UK, and also in the US.

In addition, IP11 explains their own hesitation to work with concepts of race comes from a feeling that their privilege precludes them from tackling certain topics:

> I'm a White person, I come obviously from a privileged position, and I have not experienced the oppression that racialized people have endured. So one way I deal with that is that I am not touching on topics that I feel like I cannot fully grasp.

Thus, while feminism and intersectionality are more common in the US discourse, European works are increasingly engaging with them, yet in a different manner.

Methods

In both the US and European discourses, participatory design is very important. Participatory design 'democratizes design by involving those with a stake in its implications' (S. Bardzell, 2018, p. 3). Participatory design engages communities and smaller groups with a focus on the marginalized. Works in the US discourse acknowledge the limitations of participatory design, including the danger that the marginalized are subjugated into the position of 'other' (S. Bardzell, 2018, p. 18). Participatory design can also seem elitist as participants usually must adopt the language of designers (Harrington et al., 2019, p. 16). In Europe, where participatory design has its origin (in particular in Scandinavia), the method is widespread and there is less focus on the methods' risks.

Interview partners from the European context reported to engage in participatory design in the form of living labs (IP3), hackathons (IP1), interviews and user studies (IP2), workshops (IP5), participatory history-making *via*personal family archives (IP6), co-creation and community engagement (IP7, IP9), focus groups (IP8), diary studies (IP9), and action research

(IP12). Participants as everyday people are seen as experts of their life worlds. Note the juxtaposition of 'experts' and 'engineers' in the statement by IP7:

> The point is more like, how can [we] use participatory design plus specifically participatory automation where humans are teaching their robot what they want it to do. How can we make the whole robot development process participatory, so that it can be led by experts, rather than engineers.

This is also reflected in citizen science. According to IP8, 'we need new conventions in science in how to make science if we want to make citizen science. And this has to be attached to a political change or socio-political change or engagement.'

Another commonality in terms of methods in the US and European discourses is the focus on positionality statements. In the US context, researchers and designers strongly urge to reflect their own positionality. boyd (2019) has demanded an urgent reckoning with privilege in the technology industry. Costanza-Chock (2020, p. 9) acknowledge the advantages they have experienced due to Whiteness, education, ableism, and settler colonialism. More examples of positionality statements are given in D'Ignazio et al. (2020, p. 440), D'Ignazio and Klein (2020b, p. 25), and Kozubaev et al. (2020, p. 3).

Interview partners in Europe have shared their own take on engaging critically with one's positionality. IP15 reports that it is established practice for their team to reflect positionality 'with the disclaimer that we come from a position of power.' IP 3 expresses the need to think more about designers' own powerful position in society: 'You [meaning the researcher or designer] are in such a privileged position and sometimes you don't see that in many ways, because privilege is invisible to the ones who hold it.' Furthermore, as mentioned in section 'Theories', IP11 is highly conscious of their positionality by stating 'I'm a White person, I come obviously from a privileged position'. Thus, like the US context, European designers increasingly raise awareness about designers' roles in society.

Discussion

There are some notable commonalities but also differences between the US-based and European discourses on critical diversity and design. I will highlight three observations. First, there is a similarity between the US and European discourses in their critique of the orientation of mainstream design towards modern values including efficiency, objectivity, determinism, solutionism, and neoliberalism. Shared critical views in this regard can be explained with a shared history of modernity and an increased reckoning with the effects of extreme growth, capitalism, and technocratic institutions on the well-being of people. Another commonality between the US and European discourses is the critique of academic structures, which can

similarly be explained with shared experiences of a (neo)liberalisation of universities in the USA and Europe since the 1980s and 1990s (Morrish, 2020). There is enormous pressure on academics to conduct quick in-and-out of community projects and hastily publish results (Slaughter & Rhoades, 2000). This likely motivates the critique of worsening conditions for researching and practicing design in academia.

The second observation concerns commonalities and differences regarding the use of theories of situated knowledge. As in the USA, feminisms and intersectionality have been noted as significant influences on IPs in Europe. Reluctance to consider race may be explained with Europe's history of colonialism and national projects of 'ethnic hygiene', where 'race' was used as a construct to justify atrocities against minorities (Subramaniam, 2014). In Europe, race is often avoided and presented as an outdated concept, belonging to 'the problematic past of colonialism, scientific racism, and Nazi genocide' (M'charek et al., 2014, p. 462). At the same time, 'race' is avoided as a subject to reject US-centric categories of race (e.g. Black and White) that are deemed unsuitable to explain 'difference' in Europe (Balkenhol & Schramm, 2019, p. 587). Racism of course exists in European countries, e.g. in discussions of a national Leitkultur (Salem & Thompson, 2016). The US discourse on critical diversity and design seems far ahead of Europeans in addressing and making visible problematic concepts of race and their connection to racism in design.

The third observation relates to a notable difference between the US and European discourses regarding their relationship to institutions. In the European context, engaging the public sector including government is embedded in an understanding of design as a democratic process that requires the inclusion of all stakeholders. The focus on the public sector can also be explained with the funding structure in Europe. Many universities are funded by a local or federal state; design researchers are public employees. IP7 says:

> Even if it didn't make better research, we in universities are embedded within communities, I mean, in [country] I'm a public employee, like, universities are all publicly funded, so we kind of feel like, I feel that we have a responsibility to engage the local community in its diversity, in order to have a say in like the research that we are doing.

Furthermore, the priority of providing digital services to European citizens and immigrants may stem from continental Europe's history of having a strong government. Administrative structures existed even before liberalisation and democratisation in Europe (Kaufmann & Veit Wilson, 2015, p. 6). In contrast, the founding of the USA was informed by the ideals of individual freedoms (a strong private sphere) and a limited government (Appleby, 1986, p. 25). Different from the USA, it is considered the responsibility of European governments to be

involved in shaping social spaces, and designers in Europe accept the government as trustworthy partner.

In closing, both the US and European discourses in HCI move towards a more self-conscious, reflecting, and ultimately sensitive approach to power and ‘difference’ affecting design. There is potential to learn from each other. For instance, European designers and researchers can learn from the US discourse how race and structural ‘difference’ in the form of racism can be discussed more openly. At the same time, the particularities of each context preclude simple ‘translations’ of critical design approaches from one region to the other. In going forward, what can be at the heart of an integrated, transatlantic discourse on critical diversity and design is the desire to affect social change through engaging design students, questioning designer positionalities, and analysing the social impact of design with geographic, cultural, economic, educational, social, disciplinary, and political diversity in mind.

Chapter 9: Black Feminism and Artificial Intelligence: the Possibilities and Limitations of Contesting Discriminatory AI from a Critical Social Theory Perspective

[This article was originally published in: Schelenz, Laura (2025): Black feminism and Artificial Intelligence: the possibilities and limitations of contesting discriminatory AI from a critical social theory perspective. In: *Frontiers in Sociology* 10:1602947. Doi: 10.3389/fsoc.2025.1602947]

Artificial Intelligence (AI) has been subject to critical evaluation in computer science (see the field of Machine Learning Fairness, (see the field of Machine Learning Fairness, Oneto & Chiappa, 2020), but also in science and technology studies (STS), philosophy, ethics, sociology, and the humanities. AI has become an integral part of everyday life covering all aspects of society at global scale: AI may include prediction and decision-making tools as well as (generative) assistant technologies in education, hiring and recruitment, criminal justice, communication, personal (intimate) relationships, medicine, health, science, climate and all areas of technology development (Hacker, 2025). Given the permeation of AI in society in combination with concerns about bias in and through AI, critical scholarship on AI has gained traction in the humanities and social sciences (Hanemaayer, 2022). Particularly powerful critiques of AI include scholarly works inspired by US-American Black feminism (Schelenz, 2022a), a theory and practice developed by Black women in the United States to theorize and protest the oppression of Black women at the intersection of race, gender, and class (Collins, 2000a). This paper discusses an American Black feminist approach to contesting AI. The approach is interesting because it adds a structural, power-centric perspective to the debate around AI ethics. While other critical theories such as Marxism, decolonial studies or queer studies similarly consider power relations in society and technology, Black feminism is interesting because it enables an intersectional-type analysis that foreground AI's impact on technology stakeholders at the intersection of race, gender, class, sexuality, ability, and more. While a Black feminist intersectional-type approach to AI is certainly not the only feminist response to AI (see the wealth of feminist approaches in J. Browne et al., 2023), Black feminism is distinct because it builds on the structural experiences of Black women.

Black feminisms exist in plurality and have emerged in different regions in Africa, Latin America, the Caribbean, and Europe (Emejulu & Sobande, 2019; Norwood, 2013; Rodriguez et al., 2016). This paper focuses on US-centric Black feminism because US-based Black feminist works have increasingly targeted technology design and development as a site of

critical inquiry (I will come back to that in a moment). US-American Black feminism combines theory and practice of resistance to the oppression of Black women. Intersectionality is an important analytical tool developed by Black feminists to reveal intersecting experiences of marginalization (Cooper, 2016). The work that Black feminists do is grounded in a normative concern for social justice and has driven Black women's activism around social change since U.S. slavery (Guy-Sheftall, 1996).

Research inspired by Black feminism (in the following abbreviated as BF) has challenged AI-enabled technologies in recent years. This includes critiques of algorithmic bias and discrimination (Buolamwini, 2019a; Buolamwini & Gebru, 2018; Hampton, 2021; Howard & Borenstein, 2018; Noble, 2018), critiques of color-blindness in the technology industry (Daniels, 2015), accounts of racism against Black women in computer science (J. O. Thomas et al., 2018), studies of race as a system of power and its entanglement with technology (Benjamin, 2019b), frameworks for the design of technology (Costanza-Chock, 2020; Erete et al., 2023b; Kumar & Karusala, 2019), and frameworks for data collection and analysis (D'Ignazio & Klein, 2020b; Marshall, 2023). BF has become an important foundation of the critical study and design of society and technology, to the extent that the concepts and ideas of BF are also employed by White women and queer or transgender scholars.[44]

Given the increased popularity of BF in the critical study of AI-enabled technology, it is important to take a closer look at the potential and limitations of the theory. This paper therefore answers the following research questions: What tools of analysis and critical thinking can the theory and practice of Black feminism offer for the study and design of AI? What are limitations of Black feminism as a critical social theory in contesting AI? The paper focuses on the contestation of AI-based technologies that specifically work with concepts of diversity in their design (e.g., in the construction of datasets or features of an application) or that consider and cater to diverse groups of technology stakeholders. The focus on diversity and AI as the overarching use case was chosen as recent years saw not only an uptake in Black feminist critique but also more awareness of the lack of diversity in design (Erete et al., 2023b; Zou & Schiebinger, 2018).

Following a conceptual analysis of BF writings and their application to STS, the paper argues that US-American Black feminism as a critical social theory offers a rich foundation for

[44] I am myself a White European woman who has applied American Black feminist traditions to the study and design of "diversity-aware technology." While I acknowledge the limitations that my perspective brings to applying Black feminism, I believe that, if the theory is well-suited to analyze the topic at hand, a refusal to apply Black feminism due to the Whiteness of the scholar would further marginalize it. Promoting Black feminist tools of analysis from diverse positionalities may all in all contribute to better critical AI research and development.

contesting AI-based technologies that are harmful to, e.g., Black women. BF is particularly suited to critically evaluating diversity-related matters of AI, among other reasons because BF has a long tradition of analyzing notions of societal difference, which increasingly inform the design of automated personalization, recommendation, adaption, and population-based predictions (e.g., intersectional-type analysis, Crenshaw, 1989; Noble, 2016). Aside from arguing for the potential of contesting AI with BF, the paper also addresses its constraints. BF shows some limitations in critically studying AI: first, while BF focuses successfully on power systems, these are mostly limited to gender, race, and class. Second, while BF-inspired analyses are excellent at making visible the lived experiences of technology stakeholders (e.g., those subjected to algorithmic mediation), the way that those experiences are present may inadvertently essentialize them.

The third limitation points to a tension in BF about using established language and methods (including problematic constructs like “man” and “woman”) vs. refusing them and creating new ones in an analysis of AI bias. The last limitation shows overall constraints on critique from a critical social theory perspective, as the critique often comes from within the power systems which it tries to challenge.

The findings of the paper inspire a discussion of constraints on the critical study and design of AI more broadly. As stated above, BF aspires to create social change because Black women's consciousness of oppression animates their call for social justice (Collins, 2000a, p. 3). Yet when transformation toward socially just AI frameworks seems unreachable, because most critique operates from within rigid power structures that make it difficult to escape established ways of thinking and doing, where does this leave critical scholarship? In the discussion, I argue in favor of a continuum of action on AI, including structural critique that aims at transforming existing AI frameworks as well as smaller, strategic steps toward betterment.

Materials and methods

This paper contributes to the debate around contesting harmful AI applications and effects in relation to questions of diversity by discussing the potential and limitations of BF theory as a vantage point for critique. Focusing particularly on questions of diversity means that the analysis of the theory's potential and limitations considers arguments and use cases that are related to diversity. Some use cases chosen for the analysis fall into the broader category of diversity-aware or diversity-sensitive technology. To briefly define what I mean by diversity and diversity-awareness, diversity refers to a conceptual idea of difference (e.g., among users of technology) or a normative claim toward fairness, inclusion, and justice, (e.g., of different

stakeholders in the digital society, Schelenz et al., 2019). Diversity-aware technology is a technology that, in its design, draws on a concept of difference and/or a normative claim related to inclusion, representation, equality, and more (Schelenz, 2023a).

The focus on diversity-related arguments and use cases was chosen because more and more computer scientists and technology designers are interested in diversity. For instance, diversity is increasingly embedded as a concept and/or a value into algorithmic models, data collection processes, and the design of user interfaces: In human-computer interaction, diversity concepts of the user are employed to make the design of a system more responsive to the diverse needs of users (Himmelsbach et al., 2019). In AI-based recommendation systems like Netflix or Amazon, different items (products or content) to be recommended to the user are perceived as diverse, plus designers diversify recommendations to achieve fairness (Ekstrand et al., 2018; cf. Sonboli, Burke, et al., 2020) and user satisfaction (McNee et al., 2006). Furthermore, diversity is considered in balanced datasets that include information about different populations to counter data and algorithmic bias (cf. Buolamwini & Gebru, 2018; Kärkkäinen & Joo, 2019). With increased interest of designers in diversity, I observed in prior work that the way that diversity is leveraged in the technology industry often constitutes oversimplified, binary, and naturalizing operationalizations of (human) difference (Schelenz, 2022b). This is where a critique inspired by Black feminism comes in, and the potential and limitations of this critique are evaluated in the paper's analysis.

The analysis is structured as follows: First, I conducted a conceptual analysis of the key features and arguments of BF as well as its potential for producing comprehensive critiques of contemporary socio-technical practices. The conceptual analysis was conducted by reviewing writings on the theory and method of BF. The literature used for this exercise is reflected in a classic canon of Black feminist writings, represented in the anthology by (Guy-Sheftall, 1996). I thus relied on an older, classic body of literature represented by the likes of Kimberlé Crenshaw, Patricia Hill Collins, bell hooks, and Audré Lorde. More recent post-structuralist and queer Black feminist literature including by Jennifer Nash and Jasbir Puar was added to gain a full view of the theory's strands and evolution. The literature addressing specifically technology and AI using Black feminist argumentation includes among others Benjamin (2019a), Benjamin (2019b), Noble (2018), Noble (2016), Marshall (2023), Buolamwini (2023).

Second, BF tools and arguments were applied to different cases or examples of AI and diversity in order to identify the benefits and limitations of the theory. For instance, one use case is the study of algorithmic bias and discrimination in the field of AI ethics, which conducts research into how social groups are treated differently through algorithmic mediation. In recent

years, ethicists have pointed to numerous cases of profiling, misrepresentation, erasure and disparate treatment through AI, affecting core areas of life including education, job recruitment and hiring, and criminal justice (Fabris et al., 2025; Heesen et al., 2021). I have evaluated how BF adds to AI ethics or complements arguments of algorithmic discrimination. By way of example, my analysis determined that BF primarily promotes a structural perspective on societal relations rather than the focus on individual behavior or identity, which means that AI bias can be scrutinized from a BF perspective in a way that considers the broader social fabric. BF then allows AI-based applications to be assessed as to their impact on power relations: AI may alter and transform or reinforce and cement long-held gender and racial relations as well as injustices in society.

Another use case is the design and evaluation of AI-based technology that seeks to foster the social inclusion of migrants and refugees. In this context, diversity relates to social inclusion in the sense that migrants face different challenges than the majority society, e.g., with regard to language and access to computer systems. My own previous research has looked into how an (AI-based) chatbot can support Afghan refugee women in settling in Germany (Schelenz, 2023b). In the course of this work, I tested a prototype of the chatbot with Afghan refugee women and conducted a focus group with participants. This use case was chosen to evaluate how Black feminist arguments and approaches fare in a context outside the USA, because it helps make a claim about the applicability of the theory to technology-related use cases across regions. More examples for diversity-related technology development are included along the text to sharpen arguments and help the reader understand the particularities of a Black feminist critique of AI.

Much of the paper involves my own interpretation and judgment of how Black feminist arguments can enrich discussions and debate contesting AI. While my arguments are of course grounded in a rich foundation of literature (both in the field of Black feminist thinking but also the critical study of technology inspired by Black feminism), investigating, testing, and evaluating a theory remains a highly qualitative and at times subjective task of interpretation. I hope that scholars will critically engage with the soundness and substance of the arguments made in this paper and perhaps relate them to other use cases of AI or investigate a different critical theory (whether it is decolonial studies, queer studies, etc.) as to its potential and limitations in contesting AI.

Results

Part A: the benefits of Black feminism in contesting Artificial Intelligence

The results section first presents the benefits of BF in contesting AI.

Benefit 1: a structural perspective offers a multi-dimensional critical analysis of AI

An important feature of BF is its structural perspective on society and technology, which is inspired by an intersectional-type analysis. Intersectionality is one of the core products of Black feminist thought and moved into mainstream feminist politics around the 2000s (Yuval-Davis, 2006, p. 194). Established as a term by Crenshaw (1989, 1991), intersectionality reveals interlocking forms of oppression when systems of power such as race and gender interact (also see Collins, 2000a, p. 227; Collins & Bilge, 2016). Power systems are conglomerates of knowledge, narratives, and symbols that materialize in the creation and operation of institutions, policies, and practices. Power systems rely on the social construction of human difference. They structure society and technology by advancing certain norms and othering those people and practices that fall outside the norm.

Power systems then become systems of oppression for those who are deemed "Other" to the norm (Collins, 2000a, p. 4). To make explicit the hidden norms produced by power systems, BF considers the social groups that are structurally advantaged and disadvantaged in society: the norm of race is White, of gender is male, of class is middle or upper class, of sexuality is heterosexual, of ability is abled-bodied.[45] Black women as a group are structurally disadvantaged in society and technology as they are consistently "othered," whether as people or technology users (Noble, 2018) or designers (Twine, 2018). A structural perspective thus reveals the experiences of different groups (or "structural identities") in society.

The emphasis of a structural perspective in BF should not dismiss individual and personal influences on stakeholders' experiences with technology. However, an individual perspective often renders invisible that individuals are influenced by larger structures that shape their individual behavior (Collins, 2000a, p. 171). For example, racism is not merely an individual attitude but promoted through institutions that affect individuals (e.g., schools, administrations, etc. Collins, 2000a, p. 227; Feagin, 2006). Race as a system of power is produced and reproduced via institutions and the collective of individuals who are socialized in said institutions to internalize a series of "knowledge" (e.g., race science, Rusert, 2017; Subramaniam, 2014), narratives (of who is considered the norm, Reddy, 1998), and symbols (skin color, bodily appearance, hair, hooks, 2015b) that degrade Black women.

[45] Power systems that have received less attention in Black feminism include religion (the norm is Christianity), language (the norm is English) and wealth (the norm is wealth accumulation as opposed to redistribution). I will come back to these power systems later in this paper.

A structural perspective can also help theorize diversity and difference. Black feminists have demonstrated that Black women have structurally different experiences than White women. For example, sexual violence is experienced differently by White and Black women with Black women being turned away at the women's shelter or immigrants not finding support in their native language (Crenshaw, 1995). This is because of the co-constitution of systems of power and the intersecting oppression that Black women face (for a history of Black feminist intersectional thinking, see Brah & Phoenix, 2004). Structural difference or structural diversity is then tied to diverging experiences of oppression (and privilege, Collins, 2000a, 25f). The idea of structural diversity as the different experiences of groups is beneficial because it helps researchers and designers understand how AI-based technology affects people differently depending on their socio-economic-political standing in society.

In application to the study of AI-based technology, BF has the capacity to shed light on the lived experiences of technology stakeholders as they are shaped by power systems (Schelenz, 2022a). BF can reveal a technology stakeholder's position at the intersection of power systems such as gender, race, and class, and how this position materializes in specific types of oppression. For instance, Black women are impacted by design systems and institutions that have been found to exhibit bias against Black women due to an overrepresentation of White male perspectives in datasets and design teams (Criado Perez, 2019; Wachter-Boettcher, 2017, p. 20). Biased and derogatory cultural representations of Black women as "Mammies," "Jezebels," and "Sapphires" are disseminated at scale in AI-optimized search engines and chat bots, which act as knowledge gatekeepers and cement harmful stereotypes (Noble, 2018, p. 98; Salinas et al., 2024, 24ff).

Black women may further experience economic injustice as AI-based job applicant rankings can disadvantage Black women because of a combination of gender and racial biases: Hannák et al. (2017) show that users who are read as Black and female are ranked lower than candidates from other social groups including White women. As a consequence, Black female applicants do not appear on the radar of employers or they are seen as less qualified. This bias originates in ratings and reviews, as written feedback contains fewer and less positive adjectives for Black women than for other groups (Hannák et al., 2017). With the help of a structural perspective, these experiences of political, cultural, and economic marginalization entangled with the use of AI can be made visible and related to larger social and historical contexts.

Apart from hiring and recruitment, another prominent topic in AI ethics that foregrounds questions of diversity is the visual representation of humanoid AI systems. Cave and Dihal (2020) have argued that imaginations of AI in visual productions are overwhelmingly racialized

as White. Humanoid robots but also chatbots carry physical features that resemble a White face, and voice assistants use middle- class "White" American English instead of African-American Vernacular English. Black feminist analyses point to the not so subtle messaging behind these interpretations of AI identity. Benjamin (2019b, p. 57) discusses an advertisement from the 1960s which shows a robot as a dehumanized servant, whereas the text reads: "Slavery will be back" and "Slavery will be here to stay." In the past leading to the present, imaginations of robots by White people often involved a desire to dehumanize and command inferior "beings" (Benjamin, 2019b, p. 56). This resembles White attitudes during slavery, when Black women were forced to work not only in the cotton fields but also as servants in the household (hooks, 2015a, 24ff). Going a step further, the Whiteness of robots may also be an intentional move toward an imagined White utopia, in which Black women as domestic workers are entirely removed from the White family's home (Cave & Dihal, 2020, p. 94; Rhee, 2018, p. 94). A critical Black feminist analysis helps reveal those links and understand historically-grown dynamics of oppression which become entangled and embedded in the design of AI systems.

Benefit 2: Black feminist tools of analysis can be applied across regional and cultural contexts
Since power systems follow a similar logic (of creating a norm and an "Other"), Black feminism helps to theorize social difference across geographic and cultural contexts. Black feminists have stressed that systems of power are universal to the extent that they shape the experiences of Black women and women of color in all countries and cultures. There is an important clarification to be made here. The argument is not that experiences of marginalization are the same across geographic and cultural contexts. Rather, the systems of power (race, gender, class) shaping those experiences follow a similar logic in that they establish a norm and materialize as barriers for women's self-actualization (Dhamoon, 2015, 26ff).

The benefits of attending to local particularities is that power relations can be considered with regard to their materialization in a specific location. For instance, racism in Europe works differently than in the USA. It is usually expressed as cultural racism (especially against Muslim immigrants, Chin, 2017) and triggered by cultural, religious or ethnic symbols such as the hijab. Despite multicultural narratives, "white cultures are still posited as superior to other cultures and are thus normative whereas non-white cultures are doomed deviant and inferior" (Salem & Thompson, 2016, p. 13). In application to diversity and AI-based technology, this means that AI design must pay attention to how power materializes at the local level.

The case of a mobile phone application for Afghan refugee women in Germany illustrates the benefits of a BF approach (Schelenz, 2023b). According to the study results,

Afghan women experience insecurity in navigating their new environment in Germany, e.g., fear of getting lost and missing an appointment at the immigration office. Initially, the experience of uncertainty in a new environment seems universal, affecting well-situated travelers and refugees alike. Yet power systems (particularly gender, ethnicity, language, nation/state) produce a specific experience: due to cultural racism in Germany, Afghan women may be hesitant to approach people in the street for help in navigating the city.

An app may then be a welcome support system but requires real-time, on-the-go assistance from volunteers or peers, simple navigation, and audio communication, according to study participants (Schelenz, 2023b, p. 18). An obstacle is that technology is often offered in English and uses the Latin alphabet, both mostly foreign to Afghan refugee women (Sabie & Ahmed, 2019, p. 223). Afghan refugee women may learn German in Germany, the German government does not offer free language classes to Afghan refugees as they do to, e.g., Syrian refugees (Weibert et al., 2019, p. 4). The modalities of interaction with AI-assisted technology are thus influenced by geopolitically situated power dynamics. BF helps reveal those dynamics and contest the broader policy shift toward digital, AI-based services by the German migration management system, which are difficult to manage for many refugees.

Part B: limitations of Black feminism in contesting Artificial Intelligence

Despite the great potential of BF for the study and design of AI-based technology, especially in relation to questions of diversity and inclusion, there are also limitations.

Limitation 1: BF pays little attention to power systems like language, religion, and wealth but they are important elements of a critical analysis of AI

While Black feminism is an excellent theory to make sense of the work of power, how power relations come about, and how power systems interact, there is unequal attention to different power systems. Nash (2008, p. 9) criticizes the hegemony of gender and race in Black feminism:

> "Crenshaw's seminal analysis of the ways in which black women's experiences of sexual assault and domestic violence are mediated by both race and gender neglects the ways in which these experiences are also complicated by class, nationality, language, ethnicity, and sexuality."

Collins (2015, p. 2) suggests that the following systems of power are dominating intersectional analyses: "race, class, gender, sexuality, ethnicity, nation, ability, and age." Lorde (1980) suggests there is an overemphasis on gender and suggests to include "race, sexual preference, class, and age" in analyzing structural difference. Yuval-Davis (2006, p. 202) cites

a presentation by Helma Lutz who establishes 14 systems that structure social difference: “gender; sexuality; ‘race'/skin-color; ethnicity; nation/state; class; culture; ability; age; sedentariness/origin; wealth; North–South; religion; stage of social development.” The expansive view of Lutz points to the existence of power systems that have received little attention in Black feminism and may be more visible in post-colonial and decolonial feminisms that place emphasis on women's experience in the Global South, as immigrants, and transnational women (Dhamoon, 2015, p. 27; Khan, 2018; Mirza, 2009a, p. 7).

In the use case of the mobile phone app for Afghan refugee women, language and digital literacy play a significant role (Schelenz, 2023b, p. 17). The use case showcases how (socio-technical) power relations are created through “linguistic ideologies, language policies, […] and communicative practices” (Windle et al., 2020, p. 11). The dominance of the English language has shaped the design of technology from the start. (Warschauer, 2004, p. 203) observes that the American Standard Code for Information Interchange based on the English and Roman languages was used in the initial stages of the development of the Internet. Engineers outside of the Western context were thus unable to participate in shaping the Internet, giving American and European designers a head start before the world standard for text was switched to Unicode (Warschauer, 2004, p. 203).

On a different level, workers in call centers in the Global South that provide technology support to American customers are encouraged to fake a White English sounding name and accent to appear White to the customer (Poster, 2019, p. 152). Not only is English the dominant language in the global digital economy, but White middle-class English is preferred over African-American Vernacular English (also in voice assistants and chatbots, Cave & Dihal, 2020, p. 690; Windle et al., 2020, p. 15). BF traditionally exhibits a blind spot with regard to language as a structure shaping power relations. This becomes a challenge for contesting AI, its economic context and impact on low-literate and non-English-speaking populations. In light of the increased use of large language models, language (bias) may shape the future of digital search, information, and communication (Bender et al., 2021).

Religion is another system of power that has received less attention in Black feminism and intersectional analysis than race and gender. As Weber (2015, p. 23) observes, “discussions of intersectionality have been hesitant to engage faith and religion, other than to occasionally list religion as one in a list of relevant differences.” Yet religion as a system structures social relations and creates racialized and gendered hierarchies through “organized institutions that produce particular norms and forms of belonging” and exclusion (Weber, 2015, p. 23). The co-constitution of race and religion is traced by Robinson (2019), who draws on Black feminist

writer Wynter (2015), to the construction of the racialized human subject during modern colonial expansion. The idea of White European superiority is in large parts shaped by the transformation of the "human subject" as governed by god to the "human subject" as governed by the state. In other words, secularization established a notion of the human as subject to the state but also as "homo politicus," who governs the state. In this shift, man suddenly enacted the role of god: the European man (in particular) gains the status of the savior who is to "maintain the stability, order, and territorial expansion of the state" (Robinson, 2019, 259f). Furthermore, the idea of a universal monotheistic religion (as opposed to locally specific traditions, practices, beliefs) is exported with Western imperialism, and the religious norm of (White) Christianity is installed as part of colonization and nation-state formation (Robinson, 2019, 259f). To theorize religion as a system of power is not to say that Christianity, Judaism, or Islam are inherently oppressive. Religion, like race and gender, are macro-structures, and the way religious norms have been established, e.g., through male interpretations of religious texts, can be reframed to align with feminist notions (Day, 2016; Kirk-Duggan, 2014; Turman, 2016).

Coming back to the use case of Afghan refugee women's interaction with AI-based technology, religion plays a role. Afghan refugee women report that they have a high priority for privacy and safety when interacting with social media (Schelenz, 2023b, p. 13). This may be motivated by religious gender norms, as interpretations of Islam shape the expectation of women's and men's behavior online as much as offline. Privacy and safety of the woman (and by extension the family) are important concerns for Muslim women (Mustafa et al., 2020, p. 3). Afghan women are expected not to share photographs and real names (N. Ahmed et al., 2022) or interact with men outside the family in social media (Shahalimi, 2022b, p. 113). Yet these norms are rarely taken into account when designing technology. Instead, hegemonic Christian or secular gender norms dominate technology design. A power-centric analysis can reveal this bias and shift perspective to highlight the need for women-only online spaces, which have been used by Muslim women before (Piela, 2012). Regarding religion through the lens of power can thus be an efficient element in the toolbox to critically evaluate AI systems and their impact.

Finally, wealth as a power system has gained little attention in Black feminism vis à vis gender or race. Wealth or capital is a power system that normalizes the accumulation of wealth through the extraction of human and planetary resources. According to the logic of wealth, the redistribution or minimization of wealth is bad. Kelly (2023) calls this wealth or capital bias. Kelly argues that "wealth supremacy" is entangled with White supremacy as the exploitation of Black and Brown people serves to maximize the wealth of those who own it: White people (Kelly, 2023, 12; 41f). However, the way that intersectional-type analysis deals with economic

inequalities is through "class," e.g., pointing to the experiences of multiple burdens of Black working class women compared to White women or Black men (Collins, 2000b). Bender et al. (2021, p. 247) argue that class fails to capture the complex macrostructures of production and domination ("gesellschaftliche Produktions- und Herrschaftsverhältnisse," Garske, 2013, p. 248) that shape social life.

A focus on wealth or capital can account for dynamics of AI-supported data extraction that seek to maximize financial wealth, a phenomenon that Zuboff (2019) calls "surveillance capitalism." Surveillance capitalism uses behavioral data to know, predict, and nudge the behavior of people. Instead of producing or selling services or goods (this is a side effect), surveillance capitalism seeks maximum information (Zuboff, 2019, p. 513), and thereby maximizes wealth (advertising based on behavioral data is a source of immense income for the tech industry). Although Zuboff does not relate surveillance capitalism to capitalism's co-constitution with slavery and imperialism, the practice of massive data collection has been suggested to be violent to minorities (by "including" them under the pretext of algorithmic improvement in a harmful system of data exploitation, Hoffmann, 2020) and colonizing (by extracting data from African peoples but advancing sophisticated data protection regimes in Europe, Coleman, 2019).

Limitation 2: BF-inspired analyses may inadvertently essentialize experiences of technology stakeholders, making it difficult to contest AI in a way that gets to the origins of the discontent

As we have seen in the discussion of the benefits of BF, using an intersectional-type structural analysis and perspective can reveal technology stakeholders' particular socio-technical experiences interacting with AI. There is a significant benefit of presenting experiences rather than identities in contesting AI's impact on different groups. Focusing on identities in a social critique, also known as the additive approach of intersectionality, tends to present social identities as naturalized, implying a biological or genetic constitution of the identity group in question (Yuval-Davis, 2006, p. 199). Focusing on experiences is less essentializing in a way. This said, there has been criticism in BF itself of the way that experience may inadvertently have essentializing effects as well. Nash (2008, p. 12) argues that intersectionality's focus on experiences of oppression obscures differences within the totality of Black women's experiences; some women might experience privileges:

> "In painting black women, for example, as wholly oppressed and marginalized, intersectional theory can not attend to variations within black women's experiences that afford some black women greater privilege, autonomy, and freedom."

This becomes relevant in the realm of technology design and development as some Black women may be more privileged than others, e.g., in terms of having access to a device and the Internet. This privilege of access to technology is not self-evident for resource-constrained Black people, refugees, and people of color in some areas of the Global South (Alden, 2003; Khan, 2018). At the same time, Black women in computing—despite the multi-layered discrimination they experience (Y. A. Rankin et al., 2021) —can be considered privileged in terms of education and economic standing (Liang et al., 2021, p. 28). One risk of presenting experience is thus that it does injustice to a diversity of experiences within a set of experiences.

However, more importantly for contesting AI efficiently, while experiences of privilege and oppression are influenced by power systems, the context that determines experiences (i.e., the exact work of power systems/relations) is often removed from the narration of the experience, which makes the experience appear natural (Perpich, 2010, p. 14). This is problematic because experiences become perceived as facts rather than interpretations of social phenomena that point to underlying power structures, which *can* be changed. Mirza (2009a, p. 5) writes:

> "Appeals to experience risks obscuring regimes of power by naturalizing some experiences as normative, and others as not, leaving the processes that structure dominance in tact (Scott, 1992). A black and post-colonial feminist standpoint does not valorise experience as an explanation or justification in itself, but should be seen as an interpretation of the social world that needs explaining."

For a successful critique of problematic AI application and decision-making, the risk of essentializing experience becomes quite significant. Computer scientists who are already sensitive to human-centric and critical frameworks and engage in critical analyses of AI-based technologies increasingly emphasize the importance of accounting for structural experiences of technology stakeholders, e.g., through story-telling (Ogbonnaya-Ogburu et al., 2020), autobiography (Erete et al., 2023a) or ethnography (Wong-Villacres et al., 2018). Producing narratives of technology stakeholders' structural experiences is thus becoming a valued practice in contesting AI. However, stories that reflect experiences of racism through technology and in computer science may lack an immediate analysis of how such experiences came about (see for example Ogbonnaya-Ogburu et al., 2020). This is not to say that the stories do not matter or should not be presented. Rather, following Perpich (2010, p. 18), story-telling and autobiography to contest AI should be combined with the immediate historical and political contextualization of experience. Contextualizing experience conveys the idea that experiences are not inevitable but depend on how power is constructed, organized, reproduced, challenged, changed, or abandoned in and through technology and society.

Limitation 3: BF draws on modern Western concepts such as race and gender, thus contesting AI-based services for using related concepts appears inconsistent

Black feminist works have been criticized for overwhelmingly relying on the language and concepts (such as race and gender) that have been established in the course of modernity to subordinate Black women (Jung & Costa Vargas, 2021; McKittrick, 2021; McKittrick & Wynter, 2015; Puar, 2012, 2017). Using those concepts in contesting AI, the critique may become less effective or accidentally reinforce problematic ideas that stem from colonial times of modernity. According to Jerath (2021, p. 32), "modernity refers to social, political, and economic conditions and experiences that result from modernization and capitalism." Modern experience is shaped by industrialization, technological advancement, enlightenment, colonialism, slavery, the emergence of modern nation states, and the institutionalization of human rights in the USA and Europe as a response to the violent effects of modernity (Jerath, 2021). BF is tied to modernity because it has emerged in response to the brutal subordination of Black women during the transatlantic slave trade and its aftermath. BF has also used the language of rights and justice to fight for the recognition of Black women as subjects to human rights. Puar (2012, p. 54) criticizes that intersectionality draws on the language and concepts of modernity because this reinforces the systems of power that enable the subordination of Black women in the first place:

> "many of the cherished categories of the intersectional mantra—originally starting with race, class, gender, now including sexuality, nation, religion, age, and disability—are the products of modernist colonial agendas and regimes of epistemic violence, operative through a Western/Euro-American epistemological formation through which the notion of discrete identity has emerged."

Puar thus condemns that epistemologies established by Euro-American colonialists are the foundation of Black feminist argumentation against the violence enacted by settler colonialists or their descendants. Puar calls for different epistemologies and to reject subjectivity and identity. Puar proposes queer/assemblage theory as a vantage point from which identities can be theorized as flexible and temporal events that transcend boundaries (e.g., between human-animal or nature-culture, Puar, 2012, p. 58). Queer theory, as a product of post-modern and post-structuralist thought, has made important contributions to challenging established gender identities. Judith Butler has argued that feminism works against itself by remaining within the boundaries of a "woman" category which alleges female unity and naturalizes gender (Jagose, 2007, p. 83). Simply using the term "woman" can reproduce harmful gender roles associated with it. But breaking with woman as a category and building other categories walks into the

trap of defining new exclusive boundaries. “Queer” resists any definition and upholds flexibility and fluidity in social relations (Jagose, 2007, p. 98).

With regard to contesting AI, queer theory is crucial to reveal the harm experienced by technology stakeholders who defy established categories. Keyes (2019) condemns the need to identify oneself as female or male in a computerized system to gain access and receive services. But even when one is out and identified as trans or queer, this information may be used for profiling (such as targeted content, advertisement, and surveillance). Keyes (2019) describes data science as a profound threat to queer people because the reduction of complex identities to data points denies queer existence.

There have been different reactions to post-modern and post-structuralist perspectives on gender, race, and Black women's subjectivity. One line of response acknowledges that race and gender are socially conceived as part of modern discourses, but abandoning those constructs in analysis would render their effects (racism, sexism) invisible. This is expressed by Hankerson et al. (2016), when they argue that racial bias in technology has real effects on users. When technology does not work for Black people, they feel excluded, which has consequences for their psychological state. The problem is that diversity is often not considered in the creation of personae or avatars in computer games and thus the White male body may be automatically centered as the primary user of a computer system (Hankerson et al., 2016, p. 478).

Another line of response argues that leveraging problematic constructs in identity politics may be the only way to alleviate the crackdown on Black women's lives. Cooper (2016, p. 395) counters Puar's argument against subject formation by saying that it neglects the struggle of Black women in the American context, where the nation state grants rights to subjects. Only those who can establish themselves as subjects will be able to access legal protection (which, arguably, is difficult enough for Black women). Indeed, identity politics has successfully drawn attention to the discriminatory effects of contemporary AI systems. The activism of computer scientists, authors, and activists Joy Buolamwini and Timnit Gebru, who explicitly promote a Black feminist perspective while also raising awareness about the experiences of trans users of technology, have shaped the regulatory efforts around facial recognition technology (Kantayya, 2020). Buolamwini has engaged the central institutions of the American state by testifying in Congress about racial and gender bias in facial recognition technology (Buolamwini, 2019b). Such an approach formulates Black women technology stakeholders as subjects of rights and accepts the American government as a necessary point of contact in recognizing those rights. While this approach alleviates the worst discrimination,

underlying ways of thinking/designing that cause systemic marginalization remain unchallenged.

Discussion

Black feminism is one of the most sophisticated frameworks for social justice. Working with Black feminist theory to contest AI-based algorithmic decision-making or the design of AI-based systems and their implications for different groups of technology stakeholders has enormous potential. Yet there are limitations to challenging data science or AI-based recommendation, matching, classification, evaluation and so on. This is because most of the critique is constrained by the unequal power relations that are firmly embedded in Western language and methods, including methods of programming, designing, and classifying people into subjects of algorithmic mediation (Bonilla-Silva & Zuberi, 2008). This inspires the consideration of constraints in critical research and design more broadly and how researchers and designers can deal with these constraints. In order to successfully, efficiently and lastingly contest discriminatory AI applications, I suspect that a continuum is necessary: from structural critique that aims at the transformation of AI-provoked injustice to smaller, strategic steps toward betterment. The latter may seem unsatisfying for a critical theory but being patient and carving out opportunities within constraints can gradually improve the lived reality of marginalized technology stakeholders.

How can researchers and designers contest problematic AI use considering the constraints that they face? In the process of developing a plurality of strategies toward AI justice, it is important to recognize and name the constraints that researchers and designers are facing. For instance, one constraint is the silencing and dismissal of critical scholars in mainstream discourses, which may be motivated by racism and misogyny. A strategy of working within this constraint is to wait for changed conditions in public discourse. An example of this is the 2023 rise in concern about large language models and generative AI. Leaders in the technology industry and economy have started to warn of the risk of such technology, e.g., Geoffrey Hinton and Joseph Stiglitz (Bushwick, 2023; Taylor & Hern, 2023). Black women had warned much earlier about the issue and have been discredited for naming the dangers of "stochastic parrots" (cf. Bender et al., 2021; Turner et al., 2021). However, the fact that, in 2023, AI and economic leaders publicly voiced concern can be an opportunity for scholars working with critical approaches to reiterate their arguments and move their critique of unregulated AI into the mainstream.

Another constraint is the economic context within which critical research on AI is produced, as academic and industrial spaces are shaped by neoliberal capitalism. Given this constraint, alternative business models and infrastructures for design have been proposed (cf. Scholz & Schneider, 2017). Smyth and Dimond (2014, p. 70) propose cooperatives instead of companies to make the work environment where design is facilitated anti-oppressive so that the design can be anti-oppressive as well. Smyth and Dimond (2014, p. 71) write:

> "Worker co-ops tend to create long-term, stable jobs and a concern for community benefit. Many espouse a 'multiple bottom line,' wherein the business's objectives are not limited to financial returns and include other values such as environmental sustainability, community impact, and worker happiness."

Examples of existing worker cooperatives in the technology sector are Sassafras Tech Collective[46] and Research Action Design.[47] In the context of online platforms, Poster (2019, p. 163) suggests "platform cooperativism" as a new model for more just online environments. They propose to look to the history of African-American cooperatives for models of platform infrastructures where ownership and governance is in users' hands (e.g., inspiration can be found in W.E.B DuBois'1907 text "Economic Cooperation among Negro Americans"). While these initiatives are promising, the reality is that most research and design takes place in contexts where neoliberal capitalism dictates design decisions.

This is not only true for big technology companies that demand the alignment of their products with company interests. Funding from governments, intergovernmental organizations (such as the European Union) and foundations can similarly put constraints on researchers and designers, who may further engage in self-censorship to avoid agitating potential funders (Wolf et al., 2022, p. 444). There is also a (perceived) need to scale new AI-based products or processes as a high adoption rate is a measure for success (Wolf et al., 2022, p. 444). Additionally, there is time pressure. Research projects are usually funded for a few years (if at all), pressuring researchers to go in and out of communities with little concern for their long-term development. Time is entangled with the "publish or perish culture" in academia where a researcher is incentivized to produce numerous written outputs in a short amount of time (cf. Erete et al., 2023b, 26f). Little or no funding is offered to community members who participate in the research or design project (Tran O'Leary et al., 2019, p. 8).

In light of the neoliberal framework for contesting harmful AI, researchers may focus on small-scale, local, and community-driven activities to consider how power relations materialize in a specific location/history. Such local design processes and the resulting

[46] https://sassafras.coop
[47] https://rad.cat

technologies thrive not because of scale but because they build on the needs and assets of a particular community. Another possibility to engage in critical work on AI is to reflect the meaning of alleged "diversity" concepts that are widely perceived as self-evident in data science (e.g., race, gender). Reflection (which comes down to theoretical work) may seem unsatisfying as it creates neither immediate change nor a product or a solution. However, changing the minds (or at least destabilizing established views) is quite radical in the literal meaning of the word as it goes to the roots of the dominant value and knowledge systems which shape socio-technical relations. Finally, working within but also stretching the boundaries of constraints can best be done in cooperation because this gives more weight to the work of researchers and designers using critical approaches. For example, there can be cooperation between different feminist movements with a focus on "designing across difference" and uplifting different critical approaches at the same time.

Conclusion

This paper has shown that a critical social theory like American Black feminism has tremendous potential but also limitations to contest harmful AI applications and their implications. It has highlighted that there are constraints on the critical scholarship and design of AI-based technologies, which come back to unequal power relations in society, between individuals and institutions, and biased knowledge systems deeply engrained in scholarly consciousness. Still, lots can be done within those constraints to challenge discriminatory AI-based systems and harmful methods of classification, sorting, and profiling via algorithmic mediation. This includes the strategic use of established (harmful) language and methods as well as strategic essentialization to get the attention of the mainstream discourse but then transforming concepts and ideas that have a history of colonization and exclusion. Further promising approaches include (a) critical reflection of notions of diversity or difference that are increasingly embedded in AI but risk reinforcing societal inequalities, (b) historical contextualization of AI applications and their effects (e.g., the case of the "Whiteness" of AI-systems' humanoid presentations), as well as (c) intersectional-type analysis of the works of power in and through AI, including the effects of race and gender regimes in combination with less recognized power systems like language, religion, and wealth. This paper hopefully has inspired the application and testing of the above strategies and motivated the use of American Black feminism in studies of AI more broadly. Increasing the number of critical studies on AI systems and their effects on diverse groups of technology stakeholders helps not only make these (computer) systems

better but also learn more about the practice of contesting powerful socio-technical systems beyond AI.

Concluding Remarks

This dissertation has investigated the meaning and design of diversity-aware technology from a critical perspective by leveraging arguments, prior work, and actions inspired by Black feminism. In the last chapter (chapter 9), the dissertation looks at the theory and practice of contesting harmful technology itself and the possibilities and limitations of working within constraints that arise during the design of technology. To offer some concluding thoughts on the issue and avenues for further research, I would like to bring Black feminism in conversation with Islamic feminism, and suggest a plurality of action based on cooperation across 'difference.'

A promising theory for thinking about how to move forward within constraints is Islamic feminism. The movement of Islamic feminism has been recognized in the 1990s (Badran, 2005, p. 8), although Muslim women's interpretation of the Qur'an – an important activity in Islamic feminism – likely goes back centuries. Islamic feminism differs from secular feminisms (which have gained prominence in the Middle East before Islamic feminism, Badran, 2005, p. 9) in that it understands the Qur'an as the guiding source for the struggle for gender equality.[48] Contrary to Black feminists, Islamic feminists do not see religion as a power structure that can be oppressive. To be clear, Black feminists do not claim that Islam is an oppressive system but that religion as such can produce unequal power relations. Religion in Islamic feminism is analyzed from a different angle, understood as a positive, affirmative value for Muslim women's identities. This does not mean that Islamic feminists are ignorant of injustice, as they have challenged male-dominated (mis)interpretations of the Qur'an (Salem, 2013; Wadud, 1999). Whereas Black feminists rely on a language of power, oppression, and resistance, Islamic feminists center the notion of agency and possibilities for feminist action 'within constraints.' Before exploring the notion of women's agency 'within constraints,' it is important to note why Islamic feminism rejects a language of oppression/resistance.

First, the skepticism towards oppression/resistance narratives may have to do with Muslim women's experiences of racism from White feminist movements that declare Islam as inherently oppressive. As part of Islamic feminist works, Muslim women counter racializing and colonizing narratives of the Muslim woman as unemancipated. The emphasis on agency tries to convey that Muslim women veil or follow conservative gender norms as a choice. Singh (2015, p. 661) states:

[48] For a history on the emergence of secular and Islamic feminism in the Middle East, see Badran (2005).

> "The literature on religious women's agency confronts and challenges the 'false consciousness' thesis (Bracke 2003; Bilge 2010) of many 'second-wave' and modern liberal secularist feminisms: that women who participate in patriarchal religious traditions are acting against their own objective interests, and are therefore simply the passive and brainwashed victims, dupes, or doormats of men and their patriarchal institutions (Avishai 2008; Burke 2012)."

Second, the skepticism towards oppression/resistance narratives can be traced to a different epistemology in Islamic thought. Mahmood (2011) has argued that the oppression/resistance narrative is grounded in the Western idea of an autonomous subject, an individual agentic identity that strives towards liberation. Black feminism has adopted the idea of a modern subject in pursuit of emancipation. An Islamic feminist view of subjectivity sees the human (and by implication human agency) tied to the subordination to a transcendent will (Singh, 2015, 662f). Agency can thus be exercised from within sites of subordination. This shift in perspective on the nature of human autonomy can help consider approaches to the critical study and design of technology that do not require the destruction of oppressive (infra)structures around technology but create opportunities for marginalized technology stakeholders within constraints.

Rabaan and Dombrowski (2022) have advanced Islamic feminism in human-computer interaction as a framework that can offer inspiration to computer scientists for the design of technology within political, social or cultural constraints. "Working within constraints" is defined by Rabaan et al. (2020, p. 16) as "the practice of finding feasible paths [of women] within their social, cultural, economic, and political conditions that work towards diminishing adverse impacts of ongoing or anticipated abuse." The idea of 'working within constraints' is elaborated in the context of domestic violence and Muslim women's strategic, patient, and calculated responses to such experiences. Constraints that Muslim women face in dealing with abuse include gender norms that prohibit Muslim women to go to work, a dominant father, the family's fear of social stigma and shame, and a lack of financial independence or support through social structures outside the family. Rabaan et al. (2020, p. 16) show that Muslim women may not leave an abusive husband but negotiate, wait, and carve out windows of freedom.[49] Working (or designing) within constraints means that constraining conditions (e.g. conservative gender norms) do not have to be broken first to empower women (or stakeholders

[49] One example presented in the study by Rabaan et al. (2020, p. 16) is that the husband of a study participant had forced her to transfer her income to his account as he claimed to be the "head of the household" and make all financial decisions. While the participant could no longer benefit from her income the way she wanted, she still engaged in the work because it gave her a sense of accomplishment. She thus reframed the benefits of the job for her to gain empowerment.
An important note: The examples here should not suggest that all Muslim women stay with an abusive husband. Furthermore, also White and secular women or Black and Brown women may stay with an abusive husband. Abuse as a 'global' experience confronts women of all kind with the difficulty to navigate this experience. Although the constraints may differ depending on a woman's context, women of all kind have varying responses to abuse.

of technology). Rabaan and Dombrowski (2022, p. 7) argue: "By using an Islamic feminist approach, we strive to provide implications and design within the context of entangled political and cultural norms, rather than eradicate those norms or perceive our user as a passive victim."

From a Black feminist perspective that centers the analysis of power systems/relations, constraints may be perceived as euphemism for oppression. The choice of the term 'constraint' may go back to Islamic feminists' skepticism towards the oppression/resistance narrative (explained above). The strength of Black feminism lies in naming constraints as the result of oppressive power structures, stressing that these constraints are not natural and that they *can* be changed. The strength of Islamic feminism is to show that power structures do not *have* to be broken for women to have agency, autonomy, and liberation. Bringing Black feminism in conversation with Islamic feminism in the context of technology design and development is helpful to consider a diversity of pathways towards social justice through the critical study and design of diversity-aware technology. Some Black feminists are outspoken against oppression, have suggested the abolition of harmful computer systems (Benjamin, 2019a), and successfully campaigned for moratoria on facial recognition technology (Kantayya, 2020). Some Islamic feminists suggest the possibility to be agentic and autonomous within abusive systems and carve out spaces of freedom for technology stakeholders (Mustafa et al., 2020; Rabaan & Dombrowski, 2022; Rabaan et al., 2020).

For a critical study and design of diversity-aware technology, a continuum of approaches from radical transformation of unjust social relations to small steps towards betterment is necessary. This is because structural changes may come slowly and technology stakeholders deserve relief today from algorithmic discrimination, hate speech and violence online, and the exploitation of their data through the extractive part of the technology industry. While radical social change can be an important ideal that motivates researchers and designers, being patient and carving out opportunities within constraints can gradually improve the lived reality of marginalized technology stakeholders. Moving forward within constraints does not mean that researchers and designers must give up on bold, visionary projects that seek the transformation of socio-technical relations. Moving forward within constraints means that, given the constraints that researchers and designers face, the critical study and design of diversity-aware technology *also* requires small, strategic moves.

Stretching the boundaries of constraints can best be done in cooperation because this gives more weight to the work of researchers and designers using critical approaches. On the one hand, there can be cooperation between different feminist movements with a focus on 'designing across difference.' This means uplifting different critical design approaches at the

same time and seeing them as mutually beneficial. For instance, Black feminisms and Islamic feminisms are two families of critical theory that can benefit each other through complementary approaches to social change but also through dialogue that foregrounds 'difference' (such as different structural experiences and related epistemologies). In line with the practice of "transversal politics" (Collins, 2000a, 245f; Yuval-Davis, 2011, p. 116), 'difference' should remain central to the dialogue between feminist movements. 'Difference'[50] can thus inform debate, while coalitions can be formed around a shared value or goal. "The transversal coming together should be not with the members of the other group en bloc, but with those who, in their different rooting, share values and goals compatible with one's own" (Yuval-Davis, 2011, p. 130). The idea of transversal politics can inspire the design of diversity-aware technology by focusing on coalition-building through dialogue without having to dissolve differences. What a transversal practice of 'designing across difference' can look like in practice should be subject to empirical studies in the future.

On the other hand, there can be cooperation in the form of solidarity with researchers and designers in the Global South. The distribution of funding opportunities for researchers and designers in the Global North and Global South is not equal (Lazem et al., 2022, p. 180). Solidarity can mean to transfer funding to partners in the Global South to ensure their participation at eye level in transnational research and design projects. It is true that researchers and designers in the Global North are working within neoliberal and capitalist constraints. Still, they probably have access to larger sums of funding because of the wealth in their countries. This aspect needs consideration in future work to ponder how a critical diversity-aware design praxis can look like across geographic and social contexts.

A plurality of methods for the research and design of diversity and technology may include many other forms. It can help move towards more just socio-technical futures. From speculative art, worker collectives, protest, and activism to conference publications, reforms of disciplinary curricula, and interdisciplinary cooperation – these activities can all contribute in one way or another to questioning established norms about diversity in relation to technology and ultimately creating improved socio-technical relations.

[50] 'Difference' in a coalition between Black and Islamic feminists might include different attitudes towards the oppression/resistance narrative, cf. Singh (2015), or the diverging conceptualizations of human autonomy in Black and Islamic feminism, see Mahmood (2011).

Literature

References

Abascal, J., Gardeazabal, X., Pérez, J. E., Valencia, X., Arbelaitz, O., Muguerza, J., & Yera, A. (2019). Personalizing the User Interface for People with Disabilities. In M. Augstein, E. Herder, W. Wörndl, & E. Yigitbas (Eds.), *Proceedings of the 23rd International Workshop on Personalization and Recommendation on the Web and Beyond* (p. 29). ACM. https://doi.org/10.1145/3345002.3349292

Abdelhady, D., Gren, N., & Joormann, M. (Eds.). (2020). *Refugees and the Violence of Welfare Bureaucracies in Northern Europe*. Manchester University Press.

Abdelkarim, S. (2021, October 9). Afghan Women and Resistance to the War on Terror. *Verfassungsblog*. https://verfassungsblog.de/os1-afghan-women/

Abdulla, D. (2014). A Manifesto of Change or Design Imperialism? A Look at the Purpose of the Social Design Practice. In C. Coletta, S. Colombo, P. Magaudda, A. Mattozzi, L. L. Parolin, & L. Rampino (Eds.), *A Matter Of Design. Making Society Through Science And Technology: Proceedings of the 5th STS Italia Conference* (245-260). STS Italia Publishing.

AbuJarour, S [Safa'a], Wiesche, M., Díaz Andrade, A., Fedorowicz, J., Krasnova, H., Olbrich, S., Tan, C.-W., Urquhart, C., & Venkatesh, V. (2019). ICT-enabled Refugee Integration: A Research Agenda. *Communications of the Association for Information Systems*, *44*(40), 874–891. https://doi.org/10.17705/1CAIS.04440

Ahmadi, F. (2006). Islamic Feminism in Iran: Feminism in a New Islamic Context. *Journal of Feminist Studies in Religion*, *22*(2), 33–53.

Ahmed, N., Tasmin, M., & Nasim Ibrahim, S. M. (2022). Technology for Empowerment: Context of Urban Afghan Women. *Technology in Society*, *70*, 1–6. https://doi.org/10.1016/j.techsoc.2022.102058

Ahmed, S. (2009). Embodying Diversity: Problems and Paradoxes for Black Feminists. *Race Ethnicity and Education*, *12*(1), 41–52. https://doi.org/10.1080/13613320802650931

Ahmed, S. (2012). *On Being Included: Racism and Diversity in Institutional Life*. Duke University Press.

Akca, A. A. (2020). *Moscheeleben in Deutschland: Eine Ethnographie zu Islamischem Wissen, Tradition und religiöser Autorität* (1st ed.). *Globaler lokaler Islam*. Transcript Verlag.

Alam, K., & Imran, S. (2015). The Digital Divide and Social Inclusion Among Refugee Migrants. *Information Technology & People*, *28*(2), 344–365. https://doi.org/10.1108/ITP-04-2014-0083

Alden, C. (2003). Let Them Eat Cyberspace: Africa, the G8 and the Digital Divide. *Millennium: Journal of International Studies*, *32*(3), 457–476. https://doi.org/10.1177/03058298030320030601

Alexander, M., & West, C [Cornel]. (2012). *The New Jim Crow: Mass Incarceration in the Age of Colorblindness*. The New Press.

Allard, D., & Blanchard, E. G. (Eds.). (2011). *Handbook of Research on Culturally-Aware Information Technology: Perspectives and Models*. IGI Global. http://services.igi-global.com/resolvedoi/resolve.aspx?doi=10.4018/978-1-61520-883-8

Allen, A. (2016). Feminist Perspectives on Power. *Stanford Encyclopedia of Philosophy Archive*. https://plato.stanford.edu/archives/fall2016/entries/feminist-power

Almohamed, A., Vyas, D., & Zhang, J. (2018). Designing for Refugees: Insights from Design Workshop. In G. Buchanan & D. Stevenson (Eds.), *Proceedings of the 30th Australian Conference on Computer-Human Interaction* (pp. 92–96). ACM. https://doi.org/10.1145/3292147.3292196

Almohamed, A., Zhang, J., & Vyas, D. (2020). Magic Machines for Refugees. In N.A. (Ed.), *Proceedings of the 3rd ACM SIGCAS Conference on Computing and Sustainable Societies* (pp. 76–86). ACM. https://doi.org/10.1145/3378393.3402256

Ålund, A. (1999). Ethnicity, Multiculturalism and the Problem of Culture. *European Societies*, *1*(1), 105–116. https://doi.org/10.1080/14616696.1999.10749927

Alves, P., Saraiva, P., Carneiro, J., Campos, P., Martins, H., Novais, P., & Marreiros, G. (2020). Modeling Tourists' Personality in Recommender Systems. In T. Kuflik, I. Torre, R. Burke, & C. Gena (Eds.), *Proceedings of the 28th ACM Conference on User Modeling, Adaptation and Personalization* (pp. 4–13). ACM. https://doi.org/10.1145/3340631.3394843

(2020, May 28).

American Civil Liberties Union. (2021, March 16). *Racial Profiling*. https://www.aclu.org/issues/racial-justice/race-and-criminal-justice/racial-profiling

Ammicht Quinn, R. (2006). Kulturethik. In M. Düwell, C. Hübenthal, & M. H. Werner (Eds.), *Handbuch Ethik* (2. aktualisierte und erw. Aufl., pp. 264–269). J.B. Metzler.

Anderson, S. K., & Middleton, V. A. (Eds.). (2018). *Explorations in Diversity: Examining the Complexities of Privilege, Discrimination, and Oppression* (Third edition). Oxford University Press.

Angwin, J., Larson, J., Mattu, S., & Kirchner, L. (2016, May 23). Machine Bias. *ProPublica*. https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing

Appleby, J. (1986). Republicanism in Old and New Contexts. *The William and Mary Quarterly*, *43*(1), 20–34. https://doi.org/10.2307/1919355

Arun, C. (2020). AI and the Global South. In M. D. Dubber, F. Pasquale, S. Das, & C. Arun (Eds.), *The Oxford Handbook of Ethics of AI* (pp. 587–606). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780190067397.013.38

Badran, M. (2005). Between Secular and Islamic Feminism/s: Reflections on the Middle East and Beyond. *Journal of Middle East Women's Studies*, *1*(1), 6–28.

Balkenhol, M., & Schramm, K. (2019). Doing Race in Europe: Contested Pasts and Contemporary Practices. *Social Anthropology*, *27*(4), 585–593. https://doi.org/10.1111/1469-8676.12721

Banskota, A., & Ng, Y.-K. (2020). Recommending Video Games to Adults with Autism Spectrum Disorder for Social-Skill Enhancement. In T. Kuflik, I. Torre, R. Burke, & C. Gena (Eds.), *Proceedings of the 28th ACM Conference on User Modeling, Adaptation and Personalization* (pp. 14–22). ACM. https://doi.org/10.1145/3340631.3394867

Baranoff, J., Gonzales, R. I., Liu, J., Yang, H., & Zheng, J. (2015). Lantern: Empowering Refugees Through Community-Generated Guidance Using Near Field Communication. In B. Begole (Ed.), *Chi 2015 : Extended Abstracts Publication of the 33rd annual CHI Conference on Human Factors in Computing Systems* (pp. 7–12). ACM. https://doi.org/10.1145/2702613.2726950

Barber, W., & Badre, A. (2006). *Cuturability: The Merging of Culture and Usability.* Graphics, Visualization & Usability Center, Georgia Institute of Technology. https://www.usj.edu.lb/moodle/stephane.bazan/obs_interculturelle/Barber%20and%20Badre.pdf

Bardzell, J., & Bardzell, S. (2013). What is "Critical" About Critical Design? In W. E. Mackay (Ed.), *ACM Digital Library, Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 3297–3306). ACM. https://doi.org/10.1145/2470654.2466451

Bardzell, S. (2018). Utopias of Participation: Feminism, Design, and the Futures. *ACM Transactions on Computer-Human Interaction*, *25*(1), 1–24. https://doi.org/10.1145/3127359

Bartneck, C., Yogeeswaran, K., Ser, Q. M., Woodward, G., Sparrow, R., Wang, S., & Eyssel, F. (2018). Robots And Racism. In T. Kanda, S. Ŝabanović, G. Hoffman, & A. Tapus (Eds.), *HRI'18: Proceedings of the 2018 ACM/IEEE International Conference on Human Robot Interaction* (pp. 196–204). ACM. https://doi.org/10.1145/3171221.3171260

Barz, M., Altmeyer, K., Malone, S., Lauer, L., & Sonntag, D. (2020). Digital Pen Features Predict Task Difficulty and User Performance of Cognitive Tests. In T. Kuflik, I. Torre, R. Burke, & C. Gena (Eds.), *Proceedings of the 28th ACM Conference on User Modeling, Adaptation and Personalization* (pp. 23–32). ACM. https://doi.org/10.1145/3340631.3394839

Bauer, C. (2019). *Allowing for Equal Opportunities for Artists in Music Recommendation*. https://arxiv.org/pdf/1911.05395

Beal, F. (1996). Double Jeopardy: To Be Black and Female. In B. Guy-Sheftall (Ed.), *Words of Fire: An Anthology of African-American Feminist Thought* (2nd edition, pp. 146–156). The New Press.

Die Beauftragte der Bundesregierung für Migration. (2023). *Lagebericht: Rassismus in Deutschland*. https://www.integrationsbeauftragte.de/resource/blob/1864320/2157012/77c8d1dddeea760bc13dbd87ee9a415f/lagebericht-rassismus-komplett-data.pdf?download=1

Bender, E. M., Gebru, T., McMillan-Major, A., & Shmitchell, S. (2021). On the Dangers of Stochastic Parrots: Can Language Models Be Too Big? In *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency* (pp. 610–623). ACM. https://doi.org/10.1145/3442188.3445922

Bendl, R., Bleijenbergh, I., Henttonen, E., & Mills, A. J. (Eds.). (2016). *The Oxford Handbook of Diversity in Organizations*. Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199679805.013.3

Benjamin, R. (Ed.). (2019a). *Captivating Technology: Race, Carceral Technoscience, and Liberatory Imagination in Everyday Life*. Duke University Press.

Benjamin, R. (2019b). *Race after Technology : Abolitionist Tools for the New Jim Code*. Polity Press.

Berg, M. (2022). Women Refugees' Media Usage: Overcoming Information Precarity in Germany. *Journal of Immigrant & Refugee Studies*, *20*(1), 125–139. https://doi.org/10.1080/15562948.2021.1907642

Berndsen, J., Smyth, B., & Lawlor, A. (2019). Pace My Race: Recommendations for Marathon Running. In T. Bogers, A. Said, P. Brusilovsky, & D. Tikk (Eds.), *13th ACM Conference on Recommender Systems* (pp. 246–250). ACM. https://doi.org/10.1145/3298689.3346991

Berrey, E. (2015). *The Enigma of Diversity: The Language of Race and the Limits of Racial Justice*. The University of Chicago Press.

Biselli, A. (2020, January 9). Automatisiertes Misstrauen: Dialektanalyse bei Geflüchteten. *Netzpolitik.Org*. https://netzpolitik.org/2020/automatisiertes-misstrauen/

Biselli, A., & Meister, A. (2019, July 19). Asylbehörde sucht mit Künstlicher Intelligenz nach auffälligen Geflüchteten. *Netzpolitik.Org*. https://netzpolitik.org/2019/asylbehoerde-sucht-mit-kuenstlicher-intelligenz-nach-auffaelligen-gefluechteten/

Block, H., & Riesewieck, M. (Directors). (2018). *The Cleaners*.

Bloor, M., Frankland, J., Thomas, M., & Robson, K. (2001). *Focus Groups in Social Research*. *Introducing Qualitative Methods*. SAGE Publications Ltd. https://doi.org/10.4135/9781849209175

Bock, J. G., Haque, Z., & McMahon, K. A. (2020). Displaced and Dismayed: How ICTs Are Helping Refugees and Migrants, and How We Can Do Better. *Information Technology for Development*, *26*(4), 670–691. https://doi.org/10.1080/02681102.2020.1727827

Bon, A., Dittoh, F., Lô, G., Pini, M., Bwana, R., WaiShiang, C., Kulathuramaiyer, N., & Baart, A. (2022). Decolonizing Technology and Society: A Perspective from the Global South. In H. Werthner, E. Prem, E. A. Lee, & C. Ghezzi (Eds.), *Perspectives on Digital Humanism.* Springer International Publishing.

Bonilla-Silva, E. (2001). *White Supremacy and Racism in the Post-Civil Rights Era*. Lynne Rienner Publishers.

Bonilla-Silva, E. (2015, February 27). *Why We Cannot Just Get Along? Race Matters in the Color-Blind Racial Moment.* Brown University. Race Today: A Symposium on Race in America, Providence, RI, USA. https://www.youtube.com/watch?v=T9ChSyjxjUI

Bonilla-Silva, E. (2018). *Racism Without Racists: Color-blind Racism and the Persistence of Racial Inequality in America* (Fifth edition). Rowman & Littlefield.

Bonilla-Silva, E., & Zuberi, T. (2008). *White Logic, White Methods: Racism and Methodology*. Rowman & Littlefield Publishers.

Borowska, K. (2020, December 15). The Monopoly On Technology And How To Defeat It. *Forbes*. https://www.forbes.com/sites/kasiaborowska/2020/12/15/the-monopoly-on-technology-and-how-to-defeat-it/?sh=24e316cb21af

boyd, d. (2019, September 13). Facing the Great Reckoning Head-On. *Medium*. https://onezero.medium.com/facing-the-great-reckoning-head-on-8fe434e10630

Bozdag, E. (2013). Bias in Algorithmic Filtering and Personalization. *Ethics and Information Technology*, *15*(3), 209–227. https://doi.org/10.1007/s10676-013-9321-6

Brah, A., & Phoenix, A. (2004). Ain't I A Woman? Revisiting Intersectionality. *Journal of International Women's Studies*, *5*(3), 75–86.

Bridges, K. M. (2017). *The Poverty of Privacy Rights*. Stanford Law Books, Stanford University Press.

Brinkmann, S. (2020). Unstructured and Semistructured Interviewing. In P. Leavy (Ed.), *The Oxford Handbook of Qualitative Research* (Second edition). Oxford University Press.

Brown, D., & Grinter, R. E. (2016). Designing for Transient Use: A Human-in-the-loop Translation Platform for Refugees. In J. Kaye, A. Druin, C. Lampe, D. Morris, & J. P. Hourcade (Eds.), *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems* (pp. 321–330). ACM. https://doi.org/10.1145/2858036.2858230

Browne, J., Cave, S., Drage, E., & McInerney, K. (Eds.). (2023). *Feminist AI : Critical Perspectives on Data, Algorithms and Intelligent Machines*. Oxford University Press.

Browne, S. (2015). *Dark Matters: On the Surveillance of Blackness*. Duke University Press.

Buolamwini, J. (2019a). *Response: Racial and Gender Bias in Amazon Rekognition - Commercial AI System for Analyzing Faces.* Medium. https://medium.com/@Joy.Buolamwini/response-racial-and-gender-bias-in-amazon-rekognition-commercial-ai-system-for-analyzing-faces-a289222eeced

Buolamwini, J. (2019b, June 26). *Artificial Intelligence: Societal and Ethical Implications: Testimony at United States House Committee on Science, Space and Technology*. https://www.congress.gov/event/116th-congress/house-event/109688

Buolamwini, J. (2023). *Unmasking AI: My Mission to Protect what is Human in a World of Machines* (First edition). Random House.

Buolamwini, J., & Gebru, T. (2018). Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification. *Proceedings of Machine Learning Research*, *81*, 1–15. http://proceedings.mlr.press/v81/buolamwini18a/buolamwini18a.pdf

Burke, R. (2017). *Multisided Fairness for Recommendation*. https://arxiv.org/pdf/1707.00093

Burke, R., Sonboli, N., & Ordonez-Gauger, A. (2018). Balanced Neighborhoods for Multi-sided Fairness in Recommendation. *Proceedings of Machine Learning Research*, *81*, 1–13. http://proceedings.mlr.press/v81/burke18a/burke18a.pdf

Bushwick, S. (2023, August 1). Unregulated AI Will Worsen Inequality, Warns Nobel-Winning Economist Joseph Stiglitz. *Scientific American*. https://www.scientificamerican.com/article/unregulated-ai-will-worsen-inequality-warns-nobel-winning-economist-joseph-stiglitz/

Busso, M., Acosta, R. C. A., & Götzen, A. de. (2023, June 16). *A Research Infrastructure for Generating and Sharing Diversity-aware Data*. https://arxiv.org/pdf/2306.09759

Bustamante Duarte, A. M., Brendel, N., Degbelo, A., & Kray, C. (2018). Participatory Design and Participatory Research: An HCI Case Study with Young Forced Migrants. *ACM Transactions on Computer-Human Interaction*, *25*(1), 1–39. https://doi.org/10.1145/3145472

Cahn, N., Carbone, J., & Levit, N. (2019). Discrimination by Design. *Arizona State Law Journal*, *51*(1), 1–69.

Capurro, R. (2019). Enculturating Algorithms. *NanoEthics*, *13*, 131–137. https://doi.org/10.1007/s11569-019-00340-9

Carol A. B. Warren. (2002). Qualitative Interviewing. In J. F. Gubrium & J. A. Holstein (Eds.), *Handbook of Interview Research* (pp. 83–102). Sage.

Carroll, A. (2017). *Designing for Justice* (TEDxHerndon). https://www.ted.com/talks/antionette_carroll_designing_for_justice

Castells, P., Hurley, N. J., & Vargas, S. (2015). Novelty and Diversity in Recommender Systems. In F. Ricci, L. Rokach, & B. Shapira (Eds.), *Recommender Systems Handbook* (pp. 881–918). Springer US. https://doi.org/10.1007/978-1-4899-7637-6_26

Cave, S., & Dihal, K. (2020). The Whiteness of AI. *Philosophy & Technology*, *33*, 685–703. https://doi.org/10.1007/s13347-020-00415-6

Chen, J [Jay], Mankoff, J., & Gomes, C. (Eds.) (2019). *Proceedings of the 2nd ACM SIGCAS Conference on Computing and Sustainable Societies*. ACM.

Chen, Z., Lu, Y., Nieminen, M. P., & Lucero, A. (2020). Creating a Chatbot for and with Migrants: Chatbot Personality Drives Co-Design Activities. In R. Wakkary, K. Andersen, W. Odom, A. Desjardins, & M. G. Petersen (Eds.), *Proceedings of the 2020 ACM Designing Interactive Systems Conference* (pp. 219–230). ACM. https://doi.org/10.1145/3357236.3395495

Cheney, I., & Shattuck, S. (Directors). (2020). *Picture a Scientist*.

Cherry, M. A. (2016). Virtual Work and Invisible Labor. In M. G. Crain, W. Poster, & M. A. Cherry (Eds.), *Invisible Labor: Hidden Work in the Contemporary World.* University of California Press.

Chi, N., Lurie, E., & Mulligan, D. K. (2021). Reconfiguring Diversity and Inclusion for AI Ethics. In M. Fourcade, B. Kuipers, S. Lazar, & D. Mulligan (Eds.), *Proceedings of the 2021 AAAI/ACM Conference on AI, Ethics, and Society* (pp. 447–457). ACM. https://doi.org/10.1145/3461702.3462622

Chin, R. (2017). *The Crisis of Multiculturalism in Europe: A History*. Princeton University Press. https://doi.org/10.1515/9781400884902

Choudhury, N. (2007). Constraint Spaces for Islamic Feminism: Women's Rights and the 2004 Constitution of Afghanistan. *Yale Journal of Law and Feminism*, *19*(155), 154–199.

Clark, A. D., Dantzler, P. A., & Nickels, A. E. (2018). Black Lives Matter: (Re)Framing the Next Wave of Black Liberation. *Research in Social Movements, Conflicts and Change*, *42*, 145–172. https://doi.org/10.1108/S0163-786X20180000042006

Clifton, B., Lavigne, S., & Tseng, F. (n.d.). Predicting Financial Crime: Augmenting the Predictive Policing Arsenal. https://whitecollar.thenewinquiry.com/static/whitepaper.pdf

Code to Inspire. (2015, May 11). *Code to Inspire - Empowering Women in Afghanistan with Coding*. https://www.youtube.com/watch?v=8p9eX4JDWqc

Coleman, D. (2019). Digital Colonialism: The 21st Century Scramble for Africa through the Extraction and Control of User Data and the Limitations of Data Protection Laws. *Michigan Journal of Race and Law*, *24*, 417–439. https://repository.law.umich.edu/cgi/viewcontent.cgi?article=1294&context=mjrl

Collins, P. H. (2000a). *Black Feminist Thought: Knowledge, Consciousness, and the Politics of Empowerment* (2nd ed.). *Perspectives on gender*. Routledge.

Collins, P. H. (2000b). Gender, Black Feminism, and Black Political Economy. *The Annals of the American Academy of Political and Social Science*, *568*, 41–53.

Collins, P. H. (2007). *Fighting Words: Black Women and the Search for Justice*. *Contradictions of modernity: Vol. 7*. University of Minnesota Press.

Collins, P. H. (2015). Intersectionality's Definitional Dilemmas. *Annual Review of Sociology*, *41*(1), 1–20. https://doi.org/10.1146/annurev-soc-073014-112142

Collins, P. H. (2019). *Intersectionality as Critical Social Theory*. Duke University Press.

Collins, P. H., & Bilge, S. (2016). *Intersectionality: Key Concepts*. Polity Press.

Colucci, E. (2007). "Focus Groups Can be Fun": The Use of Activity-Oriented Questions in Focus Group Discussions. *Qualitative Health Research*, *17*(10), 1422–1433. https://doi.org/10.1177/1049732307308129

Colucci, E. (2008). On the Use of Focus Groups in Cross-Cultural Research. In A. C. Michalos, E. Diener, W. Glatzer, T. Moum, M. A. G. Sprangers, J. Vogel, R. Veenhoven, & P. Liamputtong (Eds.), *Social Indicators Research Series. Doing Cross-Cultural Research* (Vol. 34, pp. 233–252). Springer Netherlands. https://doi.org/10.1007/978-1-4020-8567-3_15

Combahee River Collective. (1996). A Black Feminist Statement. In B. Guy-Sheftall (Ed.), *Words of Fire: An Anthology of African-American Feminist Thought* (2nd edition, pp. 232–240). The New Press.

Constantinides, A., Pietron, A. M., Belk, M., Fidas, C., Han, T., & Pitsillides, A. (2020). A Cross-cultural Perspective for Personalizing Picture Passwords. In T. Kuflik, I. Torre, R. Burke, & C. Gena (Eds.), *Proceedings of the 28th ACM Conference on User Modeling, Adaptation and Personalization* (pp. 43–52). ACM. https://doi.org/10.1145/3340631.3394859

Cooper, B. C. (2016). Intersectionality. In L. J. Disch & M. E. Hawkesworth (Eds.), *The Oxford Handbook of Feminist Theory* (pp. 385–406). Oxford University Press.

Costa, F. S. d., & Dolog, P. (2019). Collective Embedding for Neural Context-aware Recommender Systems. In T. Bogers, A. Said, P. Brusilovsky, & D. Tikk (Eds.), *13th ACM Conference on Recommender Systems* (pp. 201–209). ACM. https://doi.org/10.1145/3298689.3347028

Costanza-Chock, S. (2020). *Design Justice: Community-led Practices to Build the Worlds We Need*. *Information Policy*. The MIT Press.

Crawley, R. (2006). Diversity and the Marginalisation of Black Women's Issues. *Policy Futures in Education*, *4*(2), 172–184. https://doi.org/10.2304/pfie.2006.4.2.172

Crenshaw, K. (1988). Race, Reform, and Retrenchment: Transformation and Legitimation in Antidiscrimination Law. *Harvard Law Review*, *101*(7), 1331–1387.

Crenshaw, K. (1989). Demarginalizing the Intersection of Race and Sex: A Black Feminist Critique of Antidiscrimination Doctrine, Feminist Theory and Antiracist Politics. *University of Chicago Legal Forum*, *1989*(1), 139–167. https://chicagounbound.uchicago.edu/cgi/viewcontent.cgi?article=1052&context=uclf

Crenshaw, K. (1991). Mapping the Margins: Intersectionality, Identity Politics, and Violence against Women of Color. *Stanford Law Review*, *43*(6), 241–1299.

Crenshaw, K. (1995). Mapping the Margins: Intersectionality, Identity Politics, and Violence Against Women of Color. In K. Crenshaw, N. Gotanda, G. Peller, & K. Thomas (Eds.), *Critical Race Theory: The Key Writings that Formed the Movement* (pp. 357–383). New Press; Distributed by W.W. Norton & Co.

Crenshaw, K., Gotanda, N., Peller, G., & Thomas, K. (1995). Introduction. In K. Crenshaw, N. Gotanda, G. Peller, & K. Thomas (Eds.), *Critical Race Theory: The Key Writings that Formed the Movement* (pp. xiii–xxxii). New Press; Distributed by W.W. Norton & Co.

Crenshaw, K., Harris, L. C., HoSang Martinez, D., & Lipsitz, G. (Eds.). (2019). *Seeing Race Again: Countering Colorblindness Across the Disciplines*. University of California Press.

Criado Perez, C. (2019). *Invisible Women. Data Bias in a World Designed for Men*. Abrams.

D'Ignazio, C., & Klein, L. F. (2020a). *Data feminism / Catherine D'Ignazio and Lauren F. Klein*. * ideas series*. The MIT Press.

Daniels, J. (2009). *Cyber Racism: White Supremacy Online and the New Attack on Civil Rights*. *Perspectives on a multiracial America series*. Rowman & Littlefield Publishers.

Daniels, J. (2015). "My Brain Database Doesn't See Skin Color": Color-Blind Racism in the Technology Industry and in Theorizing the Web. *American Behavioral Scientist*, *59*(11), 1377–1393. https://doi.org/10.1177/0002764215578728

Daniels, J., Nkonde, M., & Mir, D. (2019). *Advancing Racial Literacy: Why Ethics, Diversity in Hiring, and Implicit Bias Training Aren't Enough.* Data & Society. https://datasociety.net/wp-content/uploads/2019/05/Racial_Literacy_Tech_Final_0522.pdf

Davis, A. Y. (2011). *Women, Race, and Class*. Random House US.

Day, K. (2016). *Religious Resistance to Neoliberalism: Womanist and Black Feminist Perspectives*. Palgrave Macmillan.

Dell, N., & Kumar, N. (2016). The Ins and Outs of HCI for Development. In J. Kaye, A. Druin, C. Lampe, D. Morris, & J. P. Hourcade (Eds.), *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems* (pp. 2220–2232). ACM. https://doi.org/10.1145/2858036.2858081

Delmont, M. F. (2016). *Why Busing Failed: Race, Media, and the National Resistance to School Desegregation*. *American Crossroads: Vol. 42*. University of California Press.

Desmond, M., & Emirbayer, M. (2009). WHAT IS RACIAL DOMINATION? *Du Bois Review: Social Science Research on Race*, *6*(2), 335–355. https://doi.org/10.1017/S1742058X09990166

Dhamoon, R. (2010). *Identity/Difference Politics: How Difference is Produced, and Why it Matters*. UBC Press.

Dhamoon, R. (2011). Considerations on Mainstreaming Intersectionality. *Political Research Quarterly*, *64*(1), 230–243. https://doi.org/10.1177/1065912910379227

Dhamoon, R. (2015). A Feminist Approach to Decolonizing Anti-Racism: Rethinking Transnationalism, Intersectionality, and Settler Colonialism, 20–37. https://feralfeminisms.com/wp-content/uploads/2015/12/ff_A-Feminist-Approach-to-Decolonizing-Anti-Racism_issue4.pdf

D'Ignazio, C., Graeff, E., Harrington, C. N., & Rosner, D. K. (2020). Toward Equitable Participatory Design: Data Feminism for CSCW amidst Multiple Pandemics. In M. Bietz & A. Wiggins (Eds.), *Conference Companion Publication of the 2020 on Computer Supported Cooperative Work and Social Computing* (pp. 437–445). ACM. https://doi.org/10.1145/3406865.3418588

D'Ignazio, C., & Klein, L. F. (2020b). *Data Feminism*. The MIT Press.

Dingemanse, M., & Liesenfeld, A. (2022). From Text to Talk: Harnessing Conversational Corpora for Humane and Diversity-aware Language Technology. In S. Muresan, P. Nakov, & A. Villavicencio (Chairs), *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics.* Symposium conducted at the meeting of Association for Computational Linguistics, Dublin, Ireland. https://aclanthology.org/2022.acl-long.385.pdf

Dombrowski, L. (2017). Socially Just Design and Engendering Social Change. *Interactions*, *24*(4), 63–65. https://doi.org/10.1145/3085560

Dos Santos Tavares, A. P., Sahay, S., Joia, L. A., & Fornazin, M. (2022). Connecting the Unconnected in the Global South: Bridging the Digital Divide from a Critical Perspective. *IFIP Advances in Information and Communication Technology*, *657*, 490–503. https://doi.org/10.1007/978-3-031-19429-0_29

Dudzik, B., Hung, H., Neerincx, M., & Broekens, J. (2020). Investigating the Influence of Personal Memories on Video-Induced Emotions. In T. Kuflik, I. Torre, R. Burke, & C. Gena (Eds.), *Proceedings of the 28th ACM Conference on User Modeling, Adaptation and Personalization* (pp. 53–61). ACM. https://doi.org/10.1145/3340631.3394842

Dunbar-Hester, C. (2020). *Hacking Diversity: The Politics of Inclusion in Open Technology Cultures*. *Princeton Studies in Culture and Technology*. Princeton University Press.

Eggers, M. M., Kilomba, G., Piesche, P., & Arndt, S. (Eds.). (2017). *Mythen, Masken und Subjekte: Kritische Weißseinsforschung in Deutschland* (3. Auflage). Unrast.

Ekstrand, M. D., Burke, R., & Diaz, F. (2019). Fairness and Discrimination in Retrieval and Recommendation. In B. Piwowarski, M. Chevalier, E. Gaussier, Y. Maarek, J.-Y. Nie, & F. Scholer (Eds.), *Proceedings of the 42nd International ACM SIGIR Conference on Research and Development in Information Retrieval* (pp. 1403–1404). ACM. https://doi.org/10.1145/3331184.3331380

Ekstrand, M. D., Tian, M., Kazi, M. R. I., Mehrpouyan, H., & Kluver, D. (2018). Exploring Author Gender in Book Rating and Recommendation. In S. Pera (Ed.), *Proceedings of the 12th ACM Conference on Recommender Systems* (pp. 242–250). ACM. https://doi.org/10.1145/3240323.3240373

Ellis, B. H., Kia-Keating, M., Yusuf, S. A., Lincoln, A., & Nur, A. (2007). Ethical Research in Refugee Communities and the Use of Community Participatory Methods. *Transcultural Psychiatry*, *44*(3), 459–481. https://doi.org/10.1177/1363461507081642

Emejulu, A., & Sobande, F. (2019). *To Exist is to Resist: Black Feminism in Europe* (1st ed.). Pluto Press.

Emerson, A., Geden, M., Smith, A., Wiebe, E., Mott, B., Boyer, K. E., & Lester, J. (2020). Predictive Student Modeling in Block-Based Programming Environments with Bayesian Hierarchical Models. In T. Kuflik, I. Torre, R. Burke, & C. Gena (Eds.), *Proceedings of the 28th ACM Conference on User Modeling, Adaptation and Personalization* (pp. 62–70). ACM. https://doi.org/10.1145/3340631.3394853

Erete, S. (2021). Using Black Feminist Epistemologies and Activist Frameworks to Counter Structural Racism in Design. *Interactions*, *28*(5), 56–59. https://doi.org/10.1145/3479981

Erete, S., Israni, A., & Dillahunt, T. (2018). An Intersectional Approach to Designing in the Margins. *Interactions*, *25*(3), 66–69. https://doi.org/10.1145/3194349

Erete, S., Rankin, Y. A., & Thomas, J. O. (2020). I Can't Breathe: Reflections from Black Women in CSCW and HCI. In J. Nichols (Ed.), *Proceedings of the ACM on Human-Computer Interaction* (4 (CSCW3), pp. 1–23). ACM.

Erete, S., Rankin, Y., & Thomas, J. (2023a). A Method to the Madness: Applying an Intersectional Analysis of Structural Oppression and Power in HCI and Design. *ACM*

Transactions on Computer-Human Interaction, *30*(2), 1–45. https://doi.org/10.1145/3507695

Erete, S., Rankin, Y. A., & Thomas, J. O. (2023b). A Method to the Madness: Applying an Intersectional Analysis of Structural Oppression and Power in HCI and Design. *ACM Transactions on Computer-Human Interaction*, *30*(2), 1–45. https://doi.org/10.1145/3507695

Escobar, A. (2012). *Encountering Development: The Making and Unmaking of the Third World* (Paperback reissue). Princeton University Press.

Escobar, A. (2018). *Designs for the Pluriverse: Radical Interdependence, Autonomy, and the Making of Worlds*. *New ecologies for the twenty-first century*. Duke University Press.

Eskandanian, F., & Mobasher, B. (2020). Using Stable Matching to Optimize the Balance between Accuracy and Diversity in Recommendation. In T. Kuflik, I. Torre, R. Burke, & C. Gena (Eds.), *Proceedings of the 28th ACM Conference on User Modeling, Adaptation and Personalization* (pp. 71–79). ACM. https://doi.org/10.1145/3340631.3394858

Eskandanian, F., Mobasher, B., & Burke, R. (2017). A Clustering Approach for Personalizing Diversity in Collaborative Recommender Systems. In M. Bielikova, E. Herder, F. Cena, & M. Desmarais (Eds.), *Proceedings of the 25th Conference on User Modeling, Adaptation and Personalization* (pp. 280–284). ACM. https://doi.org/10.1145/3079628.3079699

Eubanks, V. (2017). *Automating Inequality: How High-tech Tools Profile, Police, and Punish the Poor*. St. Martin's Press.

European Data Protection Board. (2021, June 21). *EDPB & EDPS Call for Ban on Use of AI for Automated Recognition* [Press release]. https://edpb.europa.eu/news/news/2021/edpb-edps-call-ban-use-ai-automated-recognition-human-features-publicly-accessible_de

European Parliament. (2016). *General Data Protection Regulation (REGULATION (EU) 2016/679)* [Legislative Act]. https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32016R0679&qid=1572627841585&from=EN

Fabris, A., Baranowska, N., Dennis, M. J., Graus, D., Hacker, P., Saldivar, J., Zuiderveen Borgesius, F., & Biega, A. J. (2025). Fairness and Bias in Algorithmic Hiring: A Multidisciplinary Survey. *ACM Transactions on Intelligent Systems and Technology*, *16*(1), 1–54. https://doi.org/10.1145/3696457

Fan, W., Ma, Y., Yin, D., Wang, J., Tang, J., & Li, Q. (2019). Deep Social Collaborative Filtering. In T. Bogers, A. Said, P. Brusilovsky, & D. Tikk (Eds.), *13th ACM Conference on Recommender Systems* (pp. 305–313). ACM. https://doi.org/10.1145/3298689.3347011

Feagin, J. R. (2006). *Systemic Racism: A Theory of Oppression*. Routledge.

Feagin, J. R. (2013). *The White Racial Frame: Centuries of Racial Framing and Counter-framing* (2nd ed.). Routledge.

Fellows, M. L., & Razack, S. (1998). The Race to Innocence: Confronting Hierarchical Relations among Women. *Gender Race & Justice*, *1*, 335–352. https://scholarship.law.umn.edu/cgi/viewcontent.cgi?article=1278&context=faculty_articles

Fisher, K. E. (2018). Information Worlds of Refugees. In C. F. Maitland (Ed.), *Information policy series. Digital Lifeline? ICTs for Refugees and Displaced Persons* (n.a.). The MIT Press. https://doi.org/10.7551/mitpress/10843.003.0010

Fisher, K. E., Yefimova, K., & Yafi, E. (2016). Future's Butterflies: Co-Designing ICT Wayfaring Technology with Refugee Syrian Youth. In J. C. Read & P. Stenton (Eds.), *Proceedings of the The 15th International Conference on Interaction Design and Children* (pp. 25–36). ACM. https://doi.org/10.1145/2930674.2930701

Ford, R. T. (2009). *Racial Culture: A Critique*. Princeton University Press.

Fox, S., Dimond, J., Irani, L., Hirsch, T., Muller, M., & Bardzell, S. (2017). Social Justice and Design: Power and Oppression in Collaborative Systems. In C. P. Lee (Ed.), *Companion of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing* (pp. 117–122). ACM. https://doi.org/10.1145/3022198.3022201

Fraser, N. (1996). Social Justice in the Age of Identity Politics: Redistribution, Recognition, and Participation. *Tanner Lectures on Human Values*. https://tannerlectures.utah.edu/_documents/a-to-z/f/Fraser98.pdf (Stanford University).

Friedman, B., & Hendry, D. (2019). *Value Sensitive Design: Shaping Technology with Moral Imagination*. The MIT Press.

Friedman, B., & Nissenbaum, H. (1996). Bias in Computer Systems. *ACM Transactions on Information Systems*, *14*(3), 330–347. https://doi.org/10.1145/230538.230561

Frolov, E., & Oseledets, I. (2019). HybridSVD: When Collaborative Information Is Not Enough. In T. Bogers, A. Said, P. Brusilovsky, & D. Tikk (Eds.), *13th ACM Conference on Recommender Systems* (pp. 331–339). ACM. https://doi.org/10.1145/3298689.3347055

Fu, Z., Xian, Y., Gao, R., Zhao, J [Jieyu], Huang, Q., Ge, Y., Xu, S., Geng, S., Shah, C., Zhang, Y [Yongfeng], & Melo, G. d. (2020, June 3). *Fairness-Aware Explainable Recommendation over Knowledge Graphs*. http://arxiv.org/pdf/2006.02046v2

Gammage, M. M. (2017). *Representations of Black Women in the Media: The Damnation of Black Womanhood*. *Routledge Transformations in Race and Media: Vol. 5*. Routledge.

Garske, P. (2013). Intersektionalität als Herrschaftskritik? Die Kategorie ‚Klasse' und das gesellschaftskritische Potenzial der Intersektionalitäts-Debatte. In V. Kallenberg, J. Meyer, & J. M. Müller (Eds.), *Intersectionality und Kritik* (pp. 245–263). Springer Fachmedien Wiesbaden.

Gildemeister, R. (2008). Doing Gender: Soziale Praktiken der Geschlechterunterscheidung. In R. Becker & B. Kortendiek (Eds.), *Handbuch Frauen- und Geschlechterforschung* (pp. 137–145). VS Verlag für Sozialwissenschaften. https://doi.org/10.1007/978-3-531-91972-0_17

Gillard, H., Howcroft, D., Mitev, N., & Richardson, H. (2010). "Missing Women": Gender, ICTs, and the Shaping of the Global Economy. *Information Technology for Development*, *14*(4), 262–279. https://doi.org/10.1002/itdj.20098

Goffman, E. (1990). *Stigma: Notes on the Management of Spoiled Identity*. *Penguin Psychology*. Penguin.

Goh, K., Loukaitou-Sideris, A., & Mukhija, V. (Eds.). (2022). *Urban and Industrial Environment*. *Just Urban Design: The Struggle for a Public City*. The MIT Press.

Gray, M. L., & Suri, S. (2019). *Ghost Work: How to Stop Silicon Valley from Building a New Global Underclass*. Houghton Mifflin Harcourt.

Grutter v. Bollinger 539 US 306, 2003 (United States Court of Appeals for the Sixth Circuit June 23, 2003). https://www.oyez.org/cases/2002/02-241

Guy-Sheftall, B. (Ed.). (1996). *Words of Fire: An Anthology of African-American Feminist Thought* (2nd edition). The New Press.

Hacker, P. (2025). *Oxford Intersections: AI in Society*. Oxford University Press. https://doi.org/10.1093/9780198945215.001.0001

Haegele, J. A., & Hodge, S. (2016). Disability Discourse: Overview and Critiques of the Medical and Social Models. *Quest*, *68*(2), 193–206. https://doi.org/10.1080/00336297.2016.1143849

Hampton, L. M. (2021). Black Feminist Musings on Algorithmic Oppression. In *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency* (pp. 1–12). ACM. https://doi.org/10.1145/3442188.3445929

Hamraie, A., & Fritsch, K. (2019). Crip Technoscience Manifesto. *Catalyst: Feminism, Theory, Technoscience*, *5*(1), 1–33. https://doi.org/10.28968/cftt.v5i1.29607

Hanemaayer, A. (Ed.). (2022). *Social and Cultural Studies of Robots and AI*. *Artificial Intelligence and Its Discontents*. Springer International Publishing. https://doi.org/10.1007/978-3-030-88615-8

Hankerson, D., Marshall, A. R., Booker, J., El Mimouni, H., Walker, I., & Rode, J. A. (2016). Does Technology Have Race? In J. Kaye, A. Druin, C. Lampe, D. Morris, & J. P. Hourcade (Eds.), *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems* (pp. 473–486). ACM. https://doi.org/10.1145/2851581.2892578

Hanna, A., Denton, E., Smart, A., & Smith-Loud, J. (2020). Towards a Critical Race Methodology in Algorithmic Fairness. In M. Hildebrandt, C. Castillo, E. Celis, S. Ruggieri, L. Taylor, & G. Zanfir-Fortuna (Eds.), *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency* (pp. 501–512). ACM. https://doi.org/10.1145/3351095.3372826

Hannák, A., Wagner, C., Garcia, D., Mislove, A., Strohmaier, M., & Wilson, C. (2017). Bias in Online Freelance Marketplaces. In C. P. Lee, S. Poltrock, L. Barkhuus, M. Borges, & W. Kellogg (Eds.), *Proceedings of the 2017 ACM International Conference on Computer-Supported Cooperative Work and Social Computing* (pp. 1914–1933). ACM. https://doi.org/10.1145/2998181.2998327

Harrington, C., Erete, S., & Piper, A. M. (2019). Deconstructing Community-Based Collaborative Design: Towards More Equitable Participatory Design Engagements. *Proceedings of the ACM on Human-Computer Interaction*, *3*(CSCW), 1–25. https://doi.org/10.1145/3359318

Harris-Perry, M. V. (2013). *Sister Citizen: Shame, Stereotypes, and Black Women in America*. Yale University Press.

Haslanger, S. A. (2004). Oppressions: Racial and Other. In M. P. Levine & T. Pataki (Eds.), *Racism in Mind* (pp. 97–123). Cornell University Press.

Haslanger, S. A. (2012). *Resisting Reality: Social Construction and Social Critique*. Oxford University Press.

Heesen, J., Reinhardt, K., & Schelenz, L. (2021). Diskriminierung durch Algorithmen vermeiden: Analysen und Instrumente für eine demokratische digitale Gesellschaft. In G. Bauer, M. Kechaja, S. Engelmann, & L. Haug (Eds.), *Gesellschaft der Unterschiede: Vol. 60. Diskriminierung und Antidiskriminierung: Beiträge aus Wissenschaft und Praxis* (1. Auflage). transcript; Transcript Verlag. https://doi.org/10.14361/9783839450819-008

Helberger, N. (2019). On the Democratic Role of News Recommenders. *Digital Journalism*, *7*(8), 993–1012. https://doi.org/10.1080/21670811.2019.1623700

Hellman, D. (2018). Indirect Discrimination and the Duty to Avoid Compounding Injustice. *Virginia Public Law and Legal Theory Research Paper*, *2017*(53), 1–16. https://ssrn.com/abstract=3033864

Hellman, D. (2020). Measuring Algorithmic Fairness. *Virginia Law Review*, *106*(4), 811–866.

Herbig, N., Düwel, T., Helali, M., Eckhart, L., Schuck, P., Choudhury, S., & Krüger, A. (2020). Investigating Multi-Modal Measures for Cognitive Load Detection in E-Learning. In T. Kuflik, I. Torre, R. Burke, & C. Gena (Eds.), *Proceedings of the 28th ACM Conference on User Modeling, Adaptation and Personalization* (pp. 88–97). ACM. https://doi.org/10.1145/3340631.3394861

Herzog, D., & Wörndl, W. (2019). User-centered Evaluation of Strategies for Recommending Sequences of Points of Interest to Groups. In T. Bogers, A. Said, P. Brusilovsky, & D.

Tikk (Eds.), *13th ACM Conference on Recommender Systems* (pp. 96–100). ACM. https://doi.org/10.1145/3298689.3346988

Hill, K. (2020, June 24). Wrongfully Accused by an Algorithm. *The New York Times*. https://www.nytimes.com/2020/06/24/technology/facial-recognition-arrest.html

Himmelsbach, J., Schwarz, S., Gerdenitsch, C., Wais-Zechmann, B., Bobeth, J., & Tscheligi, M. (2019). Do We Care About Diversity in Human Computer Interaction: A Comprehensive Content Analysis on Diversity Dimensions in Research. In S. Brewster, G. Fitzpatrick, A. Cox, & V. Kostakos (Eds.), *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (pp. 1–16). ACM. https://doi.org/10.1145/3290605.3300720

Hinton, E. (2021). *America on Fire: The Untold History of Police Violence and Black Rebellion Since the 1960s*. Liveright Publishing Corp.

Hirskyj-Douglas, I., & Lucero, A. (2019). On the Internet, Nobody Knows You're a Dog… Unless You're Another Dog. In S. Brewster, G. Fitzpatrick, A. Cox, & V. Kostakos (Eds.), *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (pp. 1–12). ACM. https://doi.org/10.1145/3290605.3300347

Hirst, A., Merich, D. de, Hoover, J., & Roccu, R. (2023). *Global Politics: Myths and Mysteries*. Oxford University Press.

Hoffmann, A. L. (2019). Where Fairness Fails: Data, Algorithms, and the Limits of Antidiscrimination Discourse. *Information, Communication & Society*, *22*(7), 900–915. https://doi.org/10.1080/1369118X.2019.1573912

Hoffmann, A. L. (2020). Terms of Inclusion: Data, Discourse, Violence. *New Media & Society*, *23*(12), 3539-3556. https://doi.org/10.1177/1461444820958725

Holmes, A. (2015, October 27). Has ‘Diversity’ Lost Its Meaning? *The New York Times Magazine*. https://www.nytimes.com/2015/11/01/magazine/has-diversity-lost-its-meaning.html

hooks, b. (2015a). *Ain't I A Woman: Black Women and Feminism* (Second edition). Routledge.

hooks, b. (2015b). *Black Looks: Race and Representation* [New ed.]. Routledge.

Hossfeld, K. (1995). Why Aren't High-Tech Workers Organized? Lessons in Gender, Race, and Nationality from Silicon Valley. In D. A. Cornford (Ed.), *Working People of California* (pp. 405–432). University of California Press.

Howard, A., & Borenstein, J. (2018). The Ugly Truth About Ourselves and Our Robot Creations: The Problem of Bias and Social Inequity. *Science and Engineering Ethics*, *24*(5), 1521–1536. https://doi.org/10.1007/s11948-017-9975-2

Huang, W. (2015). *When HCI Meets HRI: the Intersection and Distinction: Available via ResearchGate.* Virginia Tech.

Hudson-Weems, C. (1989). The Tripartite Plight of African-American Women as Reflected in the Novels of Hurston and Walker. *Journal of Black Studies*, *20*(2), 192–207.

Hull, G. T., Bell Scott, P., & Smith, B. (Eds.). (2010). *All the Women Are White, All the Blacks Are Men, But Some of Us Are Brave: Black Women's Studies* (13. repr [original: 1982]. Feminist Press at the City University of New York.

Hussain, F., & Amin, S. N. (2018). ‘I Don’t Care about their Reactions’: Agency and ICTs in Women’s Empowerment in Afghanistan. *Gender & Development*, *26*(2), 249–265. https://doi.org/10.1080/13552074.2018.1475924

Hyphen-Labs. (2019). *Rewire Your Brain in VR | NeuroSpeculative AfroFeminism.* Tribeca Film Institute. https://www.youtube.com/watch?v=43GUtyKPRm0

Iqbal, M., Aryafar, K., & Anderton, T. (2019). Style Conditioned Recommendations. In T. Bogers, A. Said, P. Brusilovsky, & D. Tikk (Eds.), *13th ACM Conference on Recommender Systems* (pp. 128–136). ACM. https://doi.org/10.1145/3298689.3347007

Jack, G. (2016). Advancing Postcolonial Approaches in Critical Diversity Studies. In R. Bendl, I. Bleijenbergh, E. Henttonen, & A. J. Mills (Eds.), *The Oxford Handbook of Diversity in Organizations* (pp. 153–174). Oxford University Press.

Jackson, S. J., Bailey, M., & Foucault Welles, B. (2020). *#HashtagActivism: Networks of Race and Gender Justice*. The MIT Press.

Jagose, A. (2007). *Queer Theory: An Introduction* (Reprinted). New York University Press.

Jauhiainen, J. S., Eyvazlu, D., Junnila, J., & Virnes, A. (2022). Digital Divides, the Internet and Social Media Uses among Afghans in Iran. *Technology in Society*, *70*, 1–10. https://doi.org/10.1016/j.techsoc.2022.102057

Jerath, K. S. (2021). *Science, Technology and Modernity: An Interdisciplinary Approach* (1st ed.). Springer International Publishing.

Jimenez, A., & Roberts, T. (2019). Decolonising Neo-Liberal Innovation: Using the Andean Philosophy of 'Buen Vivir' to Reimagine Innovation Hubs. In P. Nielsen & H. C. Kimaro (Eds.), *Information and Communication Technologies for Development: Strengthening Southern-driven Cooperation as a Catalyst for ICT4D: 15th IFIP WG 9.4 International Conference on Social Implications of Computers in Developing Countries, Proceedings: Part II.* Springer.

Joya, M., & Mallett, D. (2009). *Ich erhebe meine Stimme: Eine Frau kämpft gegen den Krieg in Afghanistan*. Piper.

Joya, M., & O'Keefe, D. (2010). *Raising My Voice: The Extraordinary Story of the Afghan Woman Who Dares to Speak Out*. Rider.

Jung, M.-K., & Costa Vargas, J. H. (2021). *Antiblackness*. Duke University Press.

Kallaway, K. (2015, August 30). Citizen: An American Lyric by Claudia Rankine Review – the Ugly Truth of Racism. *The Guardian*. https://www.theguardian.com/books/2015/aug/30/claudia-rankine-citizen-american-lyric-review

Kallenberg, V., Meyer, J., & Müller, J. M. (Eds.). (2013). *Intersectionality und Kritik*. Springer Fachmedien Wiesbaden. https://doi.org/10.1007/978-3-531-93168-5

Kaminskas, M., & Bridge, D. (2017). Diversity, Serendipity, Novelty, and Coverage: A Survey and Empirical Analysis of Beyond-Accuracy Objectives in Recommender Systems. *ACM Transactions on Interactive Intelligent Systems*, *7*(1), 1–42. https://doi.org/10.1145/2926720

Kantayya, S. (Director). (2020). *Coded Bias*. Documentary.

Karenga, T., & Tembo, C. (2012). Kawaida Womanism: African Ways of Being Woman in the World. *Western Journal of Black Studies*, *36*(1), 33–47.

Kärkkäinen, K., & Joo, J. (2019, August 14). *FairFace: Face Attribute Dataset for Balanced Race, Gender, and Age*. https://arxiv.org/pdf/1908.04913

Kaufmann, F.-X., & Veit Wilson, J. H. (2015). *European Foundations of the Welfare State*. Berghahn Books.

Kelly, M. (2023). *Wealth Supremacy: How the Extractive Economy and the Biased Rules of Capitalism Drive Today's Crises*. Berrett-Koehler Publishers.

Keshavarz, M. (2020). Violent Compassions: Humanitarian Design and the Politics of Borders. *Design Issues*, *36*(4), 20–32. https://doi.org/10.1162/desi_a_00611

Keyes, O. (2019). Counting the Countless: Why Data Science is a Profound Threat for Queer People, *Real Life Magazine*. https://reallifemag.com/counting-the-countless/

Khan, F. M. (2018). Access and Beyond: An Intersectional Approach to Women's Everyday Experiences with ICTs. *Networking Knowledge: Journal of the MeCCSA Postgraduate Network*, *11*(2), 5–20. https://doi.org/10.31165/nk.2018.112.533

Khan-Cullors, P., & bandele, a. (2017). *When They Call You a Terrorist: A Black Lives Matter Memoir*. Canongate Books Ltd.

Kind, C., Ward, C., Ochu, E., Fubara-Manuel, I., Chander, S., & Boudiaf, Y. (2021, June 24). *Racial Justice*. JUST AI Fellows. Ada Lovelace Institute. Prototyping AI Ethics Futures. https://www.adalovelaceinstitute.org/event/prototyping-ai-ethics-futures-racial-justice/

King, D. K. (1988). Multiple Jeopardy, Multiple Consciousness: The Context of a Black Feminist Ideology. *Signs: Journal of Women in Culture and Society*, *14*(1), 42–72. https://doi.org/10.1086/494491

Kirk-Duggan, C. (2014). Womanist Theology as a Corrective to African American Theology. In A. B. Pinn & K. G. Cannon (Eds.), *The Oxford Handbook of African American Theology.* Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199755653.013.0019

Korhonen, A., Traum, D., & Màrquez, L. (Eds.) (2019). *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*. Association for Computational Linguistics.

Kozubaev, S., Elsden, C., Howell, N., Søndergaard, M. L. J., Merrill, N., Schulte, B., & Wong, R. Y. (2020). Expanding Modes of Reflection in Design Futuring. In R. Bernhaupt, F. Mueller, D. Verweij, J. Andres, J. McGrenere, A. Cockburn, I. Avellino, A. Goguey, P. Bjørn, S. Zhao, B. P. Samson, & R. Kocielnik (Eds.), *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (pp. 1–15). ACM. https://doi.org/10.1145/3313831.3376526

Kravčík, M., Schmid, K., & Igel, C. (2019). Towards Requirements for Intelligent Mentoring Systems. In M. Augstein, E. Herder, W. Wörndl, & E. Yigitbas (Eds.), *Proceedings of the 23rd International Workshop on Personalization and Recommendation on the Web and Beyond* (pp. 19–21). ACM. https://doi.org/10.1145/3345002.3349290

Kreß, L.-M., & Kutscher, N. (2016). Flucht und digitale Medien. In B. W. Nikles & Bundesarbeitsgemeinschaft Kinder- und Jugendschutz (Eds.), *Kinder- und Jugendschutz in Wissenschaft und Praxis. Lebenssituation und Bedürfnisse von minderjährigen Flüchtlingen* (pp. 88–92). Bundesarbeitsgemeinschaft Kinder- und Jugendschutz e.V.

Kumar, N., & Karusala, N. (2019). Intersectional Computing. *Interactions*, *26*(2), 50–54. https://doi.org/10.1145/3305360

Kundinger, T., & Riener, A. (2020). The Potential of Wrist-Worn Wearables for Driver Drowsiness Detection: A Feasibility Analysis. In T. Kuflik, I. Torre, R. Burke, & C. Gena (Eds.), *Proceedings of the 28th ACM Conference on User Modeling, Adaptation and Personalization* (pp. 117–125). ACM. https://doi.org/10.1145/3340631.3394852

Ladson-Billings, G. (2009). 'Who You Callin' Nappy-headed?' A Critical Race Theory Look at the Construction of Black Women. *Race Ethnicity and Education*, *12*(1), 87–99. https://doi.org/10.1080/13613320802651012

Lakehal, A., Alti, A., Laborie, S., & Philippe, R. (2018). Ontology-based Context-aware Recommendation Approach For Dynamic Situations Enrichment. In S. Ilarri (Ed.), *13th International Workshop on Semantic and Social Media Adaptation and Personalization* (pp. 81–86). IEEE.

Landoni, M., Pera, M. S., Murgia, E., & Huibers, T. (2020). Inside Out: Exploring the Emotional Side of Search Engines in the Classroom. In T. Kuflik, I. Torre, R. Burke, & C. Gena (Eds.), *Proceedings of the 28th ACM Conference on User Modeling, Adaptation and Personalization* (pp. 136–144). ACM. https://doi.org/10.1145/3340631.3394847

Lashbrook, A. (2018, August 16). AI-Driven Dermatology Could Leave Dark-Skinned Patients Behind. *The Atlantic*. https://www.theatlantic.com/health/archive/2018/08/machine-learning-dermatology-skin-color/567619/

Lazem, S., Giglitto, D., Nkwo, M. S., Mthoko, H., Upani, J., & Peters, A. (2022). Challenges and Paradoxes in Decolonising HCI: A Critical Discussion. *Computer Supported Cooperative Work (CSCW)*, *31*(2), 159–196. https://doi.org/10.1007/s10606-021-09398-0

Leavy, S. (2018). Gender Bias in Artificial Intelligence: the Need for Diversity and Gender Theory in Machine Learning. In E. Abraham, E. Di Nitto, & R. Mirandola (Eds.), *Proceedings of the 1st International Workshop on Gender Equality in Software Engineering* (pp. 14–16). ACM. https://doi.org/10.1145/3195570.3195580

Leonhardt, J., Anand, A., & Khosla, M. (2018). User Fairness in Recommender Systems. In P.-A. Champin, F. Gandon, M. Lalmas, & P. G. Ipeirotis (Eds.), *Companion Proceedings of The Web Conference 2018* (pp. 101–102). ACM. https://doi.org/10.1145/3184558.3186949

Leurs, K., & Smets, K. (2018). Five Questions for Digital Migration Studies: Learning From Digital Connectivity and Forced Migration In(to) Europe. *Social Media + Society*, *4*(1), 1-16. https://doi.org/10.1177/2056305118764425

Lewis, D. (2001). Introduction: African Feminisms. *Agenda: Empowering Women for Gender Equity*, *50*, 4–10.

Lewis, H., Dwyer, P., Hodkinson, S., & Waite, L. (2015). Hyper-precarious Lives: Migrants, Work and Forced Labour in the Global North. *Progress in Human Geography*, *39*(5), 580–600. https://doi.org/10.1177/0309132514548303

Liang, C. A., Munson, S. A., & Kientz, J. A. (2021). Embracing Four Tensions in Human-Computer Interaction Research with Marginalized People. *ACM Transactions on Computer-Human Interaction*, *28*(2), 1–47. https://doi.org/10.1145/3443686

Lloyd, A., Anne Kennan, M., Thompson, K. M., & Qayyum, A. (2013). Connecting with New Information Landscapes: Information Literacy Practices of Refugees. *Journal of Documentation*, *69*(1), 121–144. https://doi.org/10.1108/00220411311295351

Loecherbach, F., Welbers, K., Moeller, J., Trilling, D., & van Atteveldt, W. (2021). Is This a Click Towards Diversity? Explaining When and Why New Users Make Diverse Choices. In ACM (Ed.), *13th ACM Web Science Conference* (pp. 282–290). ACM. https://doi.org/10.1145/3447535.3462506

Lorde, A. (1980). *Age, Race, Class and Sex: Women Redefining Difference: Paper Delivered at the Copeland Colloquium, Amerst College, MA, USA*. https://www.colorado.edu/odece/sites/default/files/attached-files/rba09-sb4converted_8.pdf

Luebke, D. M., & Milton, S. (1994). Locating the Victim: An Overview of Census-taking, Tabulation Technology and Persecution in Nazi Germany. *IEEE Annals of the History of Computing*, *16*(3), 25–39. https://doi.org/10.1109/MAHC.1994.298418

Luft, R. E., & Ward, J. (2009). Toward an Intersectionality Just Out of Reach: Confronting Challenges to Intersectional Practice. In V. Demos & M. Texler Segal (Eds.), *Advances in Gender Research. Perceiving Gender Locally, Globally, and Intersectionally* (Vol. 13, pp. 9–37). Emerald Group Publishing Limited. https://doi.org/10.1108/S1529-2126(2009)0000013005

Lumsden, K., & Harmer, E. (2019). *Online Othering: Exploring Digital Violence and Discrimination on the Web*. *Palgrave Studies in Cybercrime and Cybersecurity*. Springer International Publishing. https://doi.org/10.1007/978-3-030-12633-9

Lyon, D. (Ed.). (2003). *Surveillance as Social Sorting: Privacy, Risk, and Digital Discrimination*. Routledge.

M'charek, A., Schramm, K., & Skinner, D. (2014). Technologies of Belonging: The Absent Presence of Race in Europe. *Science, Technology, & Human Values*, *39*(4), 459–467. https://doi.org/10.1177/0162243914531149

Mackenzie, C., McDowell, C., & Pittaway, E. (2007). Beyond 'Do No Harm': The Challenge of Constructing Ethical Relationships in Refugee Research. *Journal of Refugee Studies*, *20*(2), 299–319. https://doi.org/10.1093/jrs/fem008

Mahmood, S. (2011). *Politics of Piety: The Islamic Revival and the Feminist Subject*. Princeton University Press.

Malik, M. (2018). *Bias and Beyond in Digital Trace Data* [Doctoral dissertation], Carnegie Mellon University. http://reports-archive.adm.cs.cmu.edu/anon/isr2018/abstracts/18-105.html

Malik, M. (2019, April 25). Can Algorithms Themselves Be Biased? *Medium*. https://medium.com/berkman-klein-center/can-algorithms-themselves-be-biased-cffecbf2302c

Mancini, T., Sibilla, F., Argiropoulos, D., Rossi, M., & Everri, M. (2019). The Opportunities and Risks of Mobile Phones for Refugees' Experience: A Scoping Review. *PLOS ONE*, *14*(12), 1-24. https://doi.org/10.1371/journal.pone.0225684

Marda, V., Constanza-Schock, S., Crider, C., & Dunn, A. (2021, March 10). *The Future is Up for Grabs: Panel Discussion.* ACM Conference on Fairness, Accountability, Transparency (FAccT), Virtual event.

Marshall, B. H. (2023). *Data Conscience: Algorithmic Siege on Our Humanity*. Wiley.

Martinez Demarco, S. (2023). From Digital Inclusion to IT Appropriation: Gendered Aspects of Appropriation Imaginary and Practices. *GENDER – Zeitschrift Für Geschlecht, Kultur Und Gesellschaft*, *15*(1), 72–86. https://doi.org/10.3224/gender.v15i1.06

Matzner, T. (2019). The Human Is Dead – Long Live the Algorithm! Human-Algorithmic Ensembles and Liberal Subjectivity. *Theory, Culture & Society*, *16*(7), 1-22. https://doi.org/10.1177/0263276418818877

Mauro, N., Ardissono, L., & Cena, F. (2020). Personalized Recommendation of PoIs to People with Autism. In T. Kuflik, I. Torre, R. Burke, & C. Gena (Eds.), *Proceedings of the 28th ACM Conference on User Modeling, Adaptation and Personalization* (pp. 163–172). ACM. https://doi.org/10.1145/3340631.3394845

Mayring, P. (2000). Qualitative Content Analysis. *Forum Qualitative Sozialforschung / Forum: Qualitative Social Research, 1* (2), 1–10.

Mayring, P. (2022). *Qualitative Inhaltsanalyse: Grundlagen und Techniken* (13. Auflage). Beltz.

McIntosh, P. (1988). *White Privilege and Male Privilege: A Personal Account of Coming to See Correspondences through Work in Women's Studies*. Working Paper 189. Wellesley Centers for Women, Wellesley, MA. https://www.nationalseedproject.org/images/documents/White_Privilege_and_Male_Privilege_Personal_Account-Peggy_McIntosh.pdf

McKittrick, K. (2015). Yours in the Intellectural Struggle: Sylvia Wynter and the Realization of the Living. In K. McKittrick & S. Wynter (Eds.), *Sylvia Wynter: On Being Human as Praxis.* Duke University Press.

McKittrick, K. (2021). *Dear Science and Other Stories*. *Errantries*. Duke University Press.

McKittrick, K., & Wynter, S. (Eds.). (2015). *Sylvia Wynter: On Being Human as Praxis*. Duke University Press.

McMillan Cottom, T. (2017). Black Cyberfeminism: Ways Forward for Intersectionality. In J. Daniels, K. Gregory, & T. M. Cottom (Eds.), *Digital Sociologies.* Policy Press.

McNee, S. M., Riedl, J., & Konstan, J. A. (2006). Being Accurate Is Not Enough. In G. Olson & R. Jeffries (Eds.), *CHI '06 Extended Abstracts on Human Factors in Computing Systems* (pp. 1097–1101). ACM. https://doi.org/10.1145/1125451.1125659

Millecamp, M., Htun, N. N., Conati, C., & Verbert, K. (2020). What's in a User? Towards Personalising Transparency for Music Recommender Interfaces. In T. Kuflik, I. Torre, R. Burke, & C. Gena (Eds.), *Proceedings of the 28th ACM Conference on User*

Modeling, Adaptation and Personalization (pp. 173–182). ACM. https://doi.org/10.1145/3340631.3394844

Mirza, H. S. (Ed.). (1997). *Black British Feminism: A Reader*. Routledge.

Mirza, H. S. (2009a). Plotting a History: Black and Postcolonial Feminisms in 'New Times'. *Race Ethnicity and Education*, *12*(1), 1–10. https://doi.org/10.1080/13613320802650899

Mirza, H. S. (2009b). Postcolonial Subjects, Black Feminism, and the Intersectionality of Race and Gender in Higher Education. *Counterpoints*, *369*, 233–248 (Postcolonial Challenges in Education).

Mirza, H. S. (2013). 'A Second Skin': Embodied Intersectionality, Transnationalism and Narratives of Identity and Belonging Among Muslim Women in Britain. *Women's Studies International Forum*, *36*, 5–15. https://doi.org/10.1016/j.wsif.2012.10.012

Misselhorn, C. (2018). Artificial Morality. Concepts, Issues and Challenges. *Society*, *55*(2), 161–169. https://doi.org/10.1007/s12115-018-0229-y

Mittelstadt, B. D., Allo, P., Taddeo, M., Wachter, S., & Floridi, L. (2016). The Ethics of Algorithms: Mapping the Debate. *Big Data & Society*, *3*(2), 1-21. https://doi.org/10.1177/2053951716679679

Mohamed, S., Png, M.-T., & Isaac, W. (2020). Decolonial AI: Decolonial Theory as Sociotechnical Foresight in Artificial Intelligence. *Philosophy & Technology*, *33*(4), 659–684. https://doi.org/10.1007/s13347-020-00405-8

MONITOR. (2020). *Gesichtserkennung in Deutschland*. ARD. https://www.youtube.com/watch?v=3lA_HxacA50

Morgan, K. P. (2019). Describing the Emperor's New Clothes: Three Myths of Educational (In-)Equity. In A. Diller, B. Houston, K. P. Morgan, & M. Ayim (Eds.), *The Gender Question in Education: Theory, Pedagogy, and Politics.* Routledge.

Morrish, L. (2020). Academic Freedom and the Disciplinary Regime in the Neoliberal University. In M. Lenormand & S. Dawes (Eds.), *Neoliberalism in Context: Governance, Subjectivity and Knowledge* (pp. 235–253). Palgrave Macmillan. https://doi.org/10.1007/978-3-030-26017-0_13

Muller, M. J., & Kuhn, S. (1993). Participatory Design. *Communications of the ACM*, *36*(6), 24–28. https://doi.org/10.1145/153571.255960

Mustafa, M., Lazem, S., Alabdulqader, E., Toyama, K., Sultana, S., Ibtasam, S., Anderson, R., & Ahmed, S. I. (2020). IslamicHCI: Designing with and within Muslim Populations. In R. Bernhaupt, F. Mueller, D. Verweij, J. Andres, J. McGrenere, A. Cockburn, I. Avellino, A. Goguey, P. Bjørn, S. (. Zhao, B. P. Samson, & R. Kocielnik (Eds.), *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems* (pp. 1–8). ACM. https://doi.org/10.1145/3334480.3375151

Nader, Z. (2020). Can Facebook Usher in Political Inclusion for Afghan Women? *Asia Research Brief*, *2020*(38), 1–2. https://yorkspace.library.yorku.ca/xmlui/bitstream/handle/10315/38056/YCAR_ARB0038-Nader.pdf?sequence=1&isAllowed=y

Nakamura, L. (2002). *Cybertypes: Race, Ethnicity, and Identity on the Internet*. Routledge.

Nash, J. C. (2008). Re-Thinking Intersectionality. *Feminist Review*, *89*(1), 1–15. https://doi.org/10.1057/fr.2008.4

Nash, J. C. (2019). *Black Feminism Reimagined: After Intersectionality*. *Next Wave*. Duke University Press.

Neve, J., & Palomares, I. (2019). Latent Factor Models and Aggregation Operators for Collaborative Filtering in Reciprocal Recommender Systems. In T. Bogers, A. Said, P. Brusilovsky, & D. Tikk (Eds.), *13th ACM Conference on Recommender Systems* (pp. 219–227). ACM. https://doi.org/10.1145/3298689.3347026

Noble, S. U. (2016). A Future for Intersectional Black Feminist Technology Studies. *The Scholar and Feminist Online*. http://sfonline.barnard.edu/traversing-technologies/safiya-umoja-noble-a-future-for-intersectional-black-feminist-technology-studies/

Noble, S. U. (2018). *Algorithms of Oppression: How Search Engines Reinforce Racism*. New York University Press.

Noble, S. U., & Tynes, B. M. (Eds.). (2016). *Digital Formations: vol. 105*. *The Intersectional Internet: Race, Sex, Class and Culture Online*. Peter Lang Publishing.

Nordland, R., Faizi, F., & Abed, F. (2019, January 27). Afghan Women Fear Peace With Taliban May Mean War on Them. *The New York Times*. https://www.nytimes.com/2019/01/27/world/asia/taliban-peace-deal-women-afghanistan.html

Norman, K. L., & Kirakowski, J. (Eds.). (2018). *The Wiley Handbook of Human Computer Interaction*. John Wiley & Sons. https://doi.org/10.1002/9781118976005

Norwood, C. (2013). Perspective in Africana Feminism; Exploring Expressions of Black Feminism/Womanism in the African Diaspora. *Sociology Compass*, *7*(3), 225–236. https://doi.org/10.1111/soc4.12025

Ober, M. (2023). *depatriarchise design*. https://depatriarchisedesign.com/about/

Ogbonnaya-Ogburu, I. F., Smith, A. D., To, A., & Toyama, K. (2020). Critical Race Theory for HCI. In R. Bernhaupt, F. Mueller, D. Verweij, J. Andres, J. McGrenere, A. Cockburn, I. Avellino, A. Goguey, P. Bjørn, S. Zhao, B. P. Samson, & R. Kocielnik (Eds.), *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (pp. 1–16). ACM. https://doi.org/10.1145/3313831.3376392

Okonkwo, S. (2019, July 12). *Intersectionality Will Save the Future of Science*. TEDxUCLA. https://www.youtube.com/watch?v=JOayfGNRw2o

Olteanu, A., Castillo, C., Diaz, F., & Kıcıman, E. (2019). Social Data: Biases, Methodological Pitfalls, and Ethical Boundaries. *Frontiers in Big Data*, *2*(13). https://doi.org/10.3389/fdata.2019.00013

O'Neil, C. (2016). *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy*. Allen Lane.

Oneto, L., & Chiappa, S. (2020). Fairness in Machine Learning. In L. Oneto, N. Navarin, A. Sperduti, & D. Anguita (Eds.), *Studies in Computational Intelligence. Recent Trends in Learning From Data* (Vol. 896, pp. 155–196). Springer International Publishing. https://doi.org/10.1007/978-3-030-43883-8_7

Osman, W. (2019). Media and Imperialism in the Global Village: A Case Study of Four Malalais. In A. Punathambekar & S. Mohan (Eds.), *JSTOR Open Access Monographs. Global Digital Cultures: Perspectives from South Asia* (pp. 280–296). University of Michigan Press.

Palmero Aprosio, A., Menini, S., & Tonelli, S. (2020). Adaptive Complex Word Identification through False Friend Detection. In T. Kuflik, I. Torre, R. Burke, & C. Gena (Eds.), *Proceedings of the 28th ACM Conference on User Modeling, Adaptation and Personalization* (pp. 192–200). ACM. https://doi.org/10.1145/3340631.3394857

Pandey, S., & Ilavarasan, P. V. (2019). People, Information and Culture: Locating Forms of Capital by Afghan Sikh Refugees in India through ICTs. *Technological Forecasting and Social Change*, *146*, 331–338. https://doi.org/10.1016/j.techfore.2019.06.005

Papamitsiou, Z., Sharma, K., & Giannakos, M. N. (2020). On the Dependence Structure Between Learners' Response-time and Knowledge Mastery: If Not Linear, Then What? In T. Kuflik, I. Torre, R. Burke, & C. Gena (Eds.), *Proceedings of the 28th ACM Conference on User Modeling, Adaptation and Personalization* (pp. 201–210). ACM. https://doi.org/10.1145/3340631.3394865

Patil, A. (2019). The Role of ICTs in Refugee Lives. In R. Chandwani, P. Singh, N. Kumar, & R. Veeraraghavan (Eds.), *Proceedings of the Tenth International Conference on Information and Communication Technologies and Development* (pp. 1–6). ACM. https://doi.org/10.1145/3287098.3287144

PBS Newshour. (2022). *Afghan Women, Girls Struggle for Rights Amid Taliban Control: Aired 15 August 2022*. https://www.pbs.org/video/one-year-on-sot-1660596988/

Peeters, M. M. M., van Diggelen, J., van den Bosch, K., Bronkhorst, A., Neerincx, M. A., Schraagen, J. M., & Raaijmakers, S. (2021). Hybrid Collective Intelligence in a Human–AI Society. *AI & SOCIETY*, *36*(1), 217–238. https://doi.org/10.1007/s00146-020-01005-y

Pei, C., Zhang, Y [Yi], Zhang, Y [Yongfeng], Sun, F., Lin, X., Sun, H., Wu, J., Jiang, P., Ge, J., Ou, W., & Pei, D. (2019). Personalized Re-ranking for Recommendation. In T. Bogers, A. Said, P. Brusilovsky, & D. Tikk (Eds.), *13th ACM Conference on Recommender Systems* (pp. 3–11). ACM. https://doi.org/10.1145/3298689.3347000

Peller, G. (1995). Race-Consciousness. In K. Crenshaw, N. Gotanda, G. Peller, & K. Thomas (Eds.), *Critical Race Theory: The Key Writings that Formed the Movement.* New Press; Distributed by W.W. Norton & Co.

Perpich, D. (2010). Black Feminism, Poststructuralism, and the Contested Character of Experience. In D.-D. L. Marcano, K. T. Gines, & M. Del Guadalupe Davidson (Eds.), *SUNY Series in Gender Theory. Convergences: Black Feminism and Continental Philosophy* (pp. 13–34). SUNY Press.

Piazzoni, F., Poe, J., & Santi, E. (2022). What Design for Urban Design Justice? *Journal of Urbanism: International Research on Placemaking and Urban Sustainability*, *17*, Article 379–400, 1–22. https://doi.org/10.1080/17549175.2022.2074522

Picchi, M. (2017). Egypt in Transition: What Future for Islamic Feminism? In A. M. Di Tolla & E. Francesca (Eds.), *Studi Maġrebini: 14-15. Emerging Actors in Post-Revolutionary North-Africa: Gender Mobility and Social Activism* (pp. 285–322). Università degli Studi di Napoli "L'Orientale".

Piela, A. (2012). *Muslim Women Online: Faith and Identity in Virtual Space*. *Routledge Islamic Studies Series*. Routledge.

Poster, W. (2019). Racialized Surveillance in the Digital Service Industry. In R. Benjamin (Ed.), *Captivating Technology: Race, Carceral Technoscience, and Liberatory Imagination in Everyday Life* (pp. 133–169). Duke University Press.

Puar, J. K. (2012). “I would rather be a cyborg than a goddess": Becoming-Intersectional in Assemblage Theory. *PhiloSOPHIA*, *2*(1), 49–66. jasbirkpuar.com/wp-content/uploads/2018/08/JKP_Cyborg-Goddess.pdf

Puar, J. K. (2017). *Terrorist Assemblages: Homonationalism in Queer Times*. *Next Wave*. Duke University Press.

Putnam, K. A. (2021). Information Technology Usage Among Afghan Women Business Leaders. *Communications of the IIMA*, *19*(1), 1–21. https://scholarworks.lib.csusb.edu/ciima/vol19/iss1/4

Quirke, L. (2012). Information Practices in Newcomer Settlement: A Study of Afghan Immigrant and Refugee Youth in Toronto. In J.-E. Mai (Ed.), *Proceedings of the 2012 iConference* (pp. 535–537). ACM. https://doi.org/10.1145/2132176.2132278

Rabaan, H., & Dombrowski, L. (2022). Breaking Stereotypes: Islamic Feminism and HCI. *Interactions*, *29*(6), 6–7. https://doi.org/10.1145/3565256

Rabaan, H., Young, A. L., & Dombrowski, L. (2020). Daughters of Men: Saudi Women's Sociotechnical Agency Practices in Addressing Domestic Abuse. In J. Nichols (Ed.), *Proceedings of the ACM on Human-Computer Interaction* (Vol. 44 (CSCW3), pp. 1–31). ACM. https://doi.org/10.1145/3432923

Raju, J. (2020). Diversity, Inclusion, and Social Justice in the Information Context. *The International Journal of Information, Diversity, & Inclusion (IJIDI)*, *4*(3/4). https://doi.org/10.33137/ijidi.v4i3/4.34974

Rankin, Y., & Thomas, J. (2019). Straighten Up and Fly Right: Rethinking Intersectionality in HCI Research. *Interactions*, *26*(6), 64–68. https://doi.org/10.1145/3363033

Rankin, Y. A., Thomas, J. O., & Erete, S. (2021). Real Talk: Saturated Sites of Violence in CS Education. In M. Sherriff (Ed.), *Proceedings of the 52nd ACM Technical Symposium on Computer Science Education* (pp. 802–808). ACM. https://doi.org/10.1145/3408877.3432432

Rawls, J. (1971). *A Theory of Justice: Original Edition*. Harvard University Press.

Reddy, M. T. (1998). Invisibility/Hypervisibility: The Paradox of Normative Whiteness. *Transformations: The Journal of Inclusive Scholarship and Pedagogy*, *9*(2), 55–64 (SPECIAL ISSUE: "How Could You Not Hear It?" Writings on Race · Color · Whiteness).

Reuters (2018, October 11). Amazon Ditched AI Recruiting Tool that Favored Men for Technical Jobs. *The Guardian*. https://www.theguardian.com/technology/2018/oct/10/amazon-hiring-ai-gender-bias-recruiting-engine

Rhee, J. (2018). *The Robotic Imaginary: The Human and the Price of Dehumanized Labor*. University of Minnesota Press.

Rhim, S., Lee, U., & Han, K. (2020). Tracking and Modeling Subjective Well-Being Using Smartphone-Based Digital Phenotype. In T. Kuflik, I. Torre, R. Burke, & C. Gena (Eds.), *Proceedings of the 28th ACM Conference on User Modeling, Adaptation and Personalization* (pp. 211–220). ACM. https://doi.org/10.1145/3340631.3394855

Ricci, F., Rokach, L., & Shapira, B. (Eds.). (2015). *Recommender Systems Handbook*. Springer US. https://doi.org/10.1007/978-1-4899-7637-6

Ritchie, A. J. (2017). *Invisible No More: Police Violence Against Black Women and Women of Color*. Beacon Press.

Roberge, J., & Seyfert, R. (2017). Was sind Algorithmuskulturen? In J. Roberge & R. Seyfert (Eds.), *Algorithmuskulturen: Über die rechnerische Konstruktion der Wirklichkeit* (pp. 7–40). Transcript Verlag.

Robinson, B. G. (2019). Racialization and Modern Religion: Sylvia Wynter, Black Feminist Theory, and Critical Genealogies of Religion. *Critical Research on Religion*, *7*(3), 257–274. https://doi.org/10.1177/2050303219848065

Rocca, B. (2019, June 3). Introduction to Recommender Systems. *Towards Data Science*. https://towardsdatascience.com/introduction-to-recommender-systems-6c66cf15ada

Rodriguez, C. R., Tsikata, D., & Adomako Ampofo, A. (Eds.). (2016). *Transatlantic Feminisms: Women and Gender Studies in Africa and the Diaspora*. Lexington Books.

Roughgarden, J. (2013). *Evolution's Rainbow: Diversity, Gender, and Sexuality in Nature and People*. University of California Press.

Rusert, B. (2017). *Fugitive Science: Empiricism and Freedom in Early African American Culture*. *America and the Long 19th Century*. New York University Press.

Sabbagh, D. (2020, September 28). Trump 2016 Campaign 'Targeted 3.5m Black Americans to Deter them from Voting'. *The Guardian,* 2020. https://www.theguardian.com/us-news/2020/sep/28/trump-2016-campaign-targeted-35m-black-americans-to-deter-them-from-voting

Sabie, D., & Ahmed, S. I. (2019). Moving into a Technology Land. In J. Chen, J. Mankoff, & C. Gomes (Eds.), *Proceedings of the 2nd ACM SIGCAS Conference on Computing and Sustainable Societies* (pp. 218–233). ACM. https://doi.org/10.1145/3314344.3332481

Sabie, D., Talhouk, R., Ekmekcioglu, C., Maitland, C., Wulf, V., Yafi, E., Sabie, S., Almohamed, A., AbuJarour, S [Safa'a], Le Louvier, K., Hussain, F., & Ahmed, S. I. (2021). Migration and Mobility in HCI: Rethinking Boundaries, Methods, and Impact. In Y. Kitamura, A. Quigley, K. Isbister, & T. Igarashi (Eds.), *Extended Abstracts of the 2021 CHI Conference on Human Factors in Computing Systems* (pp. 1–6). ACM. https://doi.org/10.1145/3411763.3441352

Sacharidis, D. (2020). Building User Trust in Recommendations via Fairness and Explanations. In T. Kuflik, I. Torre, R. Burke, & C. Gena (Eds.), *Adjunct Publication of the 28th ACM Conference on User Modeling, Adaptation and Personalization* (pp. 313–314). ACM. https://doi.org/10.1145/3386392.3399995

Sacharidis, D., Mukamakuza, C. P., & Werthner, H. (2020). Fairness and Diversity in Social-Based Recommender Systems. In T. Kuflik, I. Torre, R. Burke, & C. Gena (Eds.), *Adjunct Publication of the 28th ACM Conference on User Modeling, Adaptation and Personalization* (pp. 83–88). ACM. https://doi.org/10.1145/3386392.3397603

Salem, S. (2013). Feminist Critique and Islamic Feminism: The Question of Intersectionality. *The Postcolonialist*, *1*(1), 2–15.

Salem, S. (2018). Intersectionality and its Discontents: Intersectionality as Traveling Theory. *European Journal of Women's Studies*, *25*(4), 403–418. https://doi.org/10.1177/1350506816643999

Salem, S., & Thompson, V. (2016). Old Racisms, New Masks: On the Continuing Discontinuities of Racism and the Erasure of Race in European Contexts. *Nineteen Sixty Nine: An Ethnic Studies Journal*, *3*(1), 1–23. https://escholarship.org/uc/item/98p8q169#main

Salinas, A., Haim, A., & Nyarko, J. (2024, February 21). *What's in a Name? Auditing Large Language Models for Race and Gender Bias*. http://arxiv.org/pdf/2402.14875v3

Sanz-Cruzado, J., & Castells, P. (2018). Enhancing Structural Diversity in Social Networks by Recommending Weak Ties. In S. Pera (Ed.), *Proceedings of the 12th ACM Conference on Recommender Systems* (pp. 233–241). ACM. https://doi.org/10.1145/3240323.3240371

Schelenz, L. (2021). Diversity-aware Recommendations for Social Justice? Exploring User Diversity and Fairness in Recommender Systems. In J. Masthoff, E. Herder, N. Tintarev, & M. Tkalčič (Eds.), *Adjunct Proceedings of the 29th ACM Conference on User Modeling, Adaptation and Personalization* (pp. 404–410). ACM. https://doi.org/10.1145/3450614.3463293

Schelenz, L. (2022a). Artificial Intelligence Between Oppression and Resistance: Black Feminist Perspectives on Emerging Technologies. In A. Hanemaayer (Ed.), *Social and Cultural Studies of Robots and AI. Artificial Intelligence and Its Discontents* (pp. 225–249). Springer International Publishing. https://doi.org/10.1007/978-3-030-88615-8_11

Schelenz, L. (2022b). Diversity Concepts in Computer Science and Technology Development: A Critique. *Science, Technology, & Human Values*, *48*(5), 1054-1079. https://doi.org/10.1177/01622439221122549

Schelenz, L. (2023a). Diversity and Social Justice in Technology Design: Reflections on Diversity-aware Technology. *International Journal of Critical Diversity Studies*, *5*(2). https://doi.org/10.13169/intecritdivestud.5.2.0033

Schelenz, L. (2023b). Technology, Power, and Social Inclusion: Afghan Refugee Women's Interaction with ICT in Germany. *The International Journal of Information, Diversity, & Inclusion (IJIDI)*, *7*(3/4), 1–31. https://doi.org/10.33137/ijidi.v7i3/4.40292

Schelenz, L., Bison, I., Busso, M., Götzen, A. de, Gatica-Perez, D., Giunchiglia, F., Meegahapola, L., & Ruiz-Correa, S. (2021). The Theory, Practice, and Ethical Challenges of Designing a Diversity-Aware Platform for Social Relations. In M.

Fourcade, B. Kuipers, S. Lazar, & D. Mulligan (Eds.), *Proceedings of the 2021 AAAI/ACM Conference on AI, Ethics, and Society* (pp. 905–915). ACM. https://doi.org/10.1145/3461702.3462595

Schelenz, L., Reinhardt, K., & Gjuraij, D. (2019). *Developing a Conceptual and Ethical Framework of Diversity: Deliverable 9.1 for the project WeNet- The Internet of Us.* International Center for Ethics in the Sciences and Humanities (IZEW). https://internetofus.eu/wp-content/uploads/sites/38/2019/07/D9.1-EKUT-Ethical-Framework-final.pdf

Schlesinger, A., Edwards, K., & Grinter, R. E. (2017). Intersectional HCI: Engaging Identity Trough Gender, Race, and Class. In G. Mark, S. Fussell, C. Lampe, m. schraefel, J. P. Hourcade, C. Appert, & D. Wigdor (Eds.), *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems* (pp. 5412–5427). ACM. https://doi.org/10.1145/3025453.3025766

Scholz, T., & Schneider, N. (Eds.). (2017). *Ours to Hack and to Own: The Rise of Platform Cooperativism, A New Vision for the Future of Work and a Fairer Internet*. OR Books.

Schopp, K., Schelenz, L., Heesen, J., & Pawelec, M. (2019). Ethical Questions of Digitalization in the Global South: Perspective on Justice and Equality. *TATUP - Zeitschrift Für Technikfolgenabschätzung in Theorie Und Praxis*, *28*(2). https://www.tatup.de/index.php/tatup/article/view/213

Schreieck, M., Wiesche, M., & Krcmar, H. (2017). Governing Nonprofit Platform Ecosystems – An Information Platform for Refugees. *Information Technology for Development*, *23*(3), 618–643. https://doi.org/10.1080/02681102.2017.1335280

Schreieck, M., Zitzelsberger, J., Siepe, S., Wiesche, M., & and Krcmar, H. (2017). Supporting Refugees in Everyday Life - Intercultural Design Evaluation of an Application for Local Information. *Twenty First Pacific Asia Conference on Information Systems (PACIS) 2017*, 1–12.

Scott, J. W. (1992). Experience. In J. Butler & J. W. Scott (Eds.), *Feminists Theorize the Political* (pp. 22–40). Routledge.

Sertkan, M., Neidhardt, J., & Werthner, H. (2020). Eliciting Touristic Profiles: A User Study on Picture Collections. In T. Kuflik, I. Torre, R. Burke, & C. Gena (Eds.), *Proceedings of the 28th ACM Conference on User Modeling, Adaptation and Personalization* (pp. 230–238). ACM. https://doi.org/10.1145/3340631.3394868

Seyfert, R., & Roberge, J. (Eds.). (2016). *Routledge Advances in Sociology*. *Algorithmic Cultures: Essays on Meaning, Performance and New Technologies*. Routledge.

Shahalimi, N. (Ed.). (2022a). *We Are Still Here: Afghan Women on Courage, Freedom, and the Fight to be Heard*. Foreword by Margaret Atwood (First American edition). Plume, an imprint of Penguin Random House LLC.

Shahalimi, N. (Ed.). (2022b). *Wir sind noch da! Mutige Frauen aus Afghanistan* (2. Auflage). Elisabeth Sandmann Verlag.

Shen, Y. (2020). Transdisciplinary Convergence: Intelligent Infrastructure for Sustainable Development. *Data Intelligence*, *3*(2), 1–13. https://doi.org/10.1162/dint_a_00063

Singh, J. (2015). Religious Agency and the Limits of Intersectionality. *Hypatia*, *30*(4), 657–674.

Skenderi, E., Olshannikova, E., Olsson, T., Huhtamäki, J., Koivunen, S., Yao, P., & Huttunen, H. (2019). Investigation of Egocentric Social Structures for Diversity-Enhancing Followee Recommendations. In G. A. Papadopoulos, G. Samaras, S. Weibelzahl, D. Jannach, & O. C. Santos (Eds.), *Adjunct Publication of the 27th Conference on User Modeling, Adaptation and Personalization* (pp. 257–261). ACM. https://doi.org/10.1145/3314183.3323460

Slaughter, S., & Rhoades, G. (2000). The Neo-Liberal University. *New Labor Forum*, *6*, 73–79.

Smyth, T., & Dimond, J. (2014). Anti-Oppressive Design. *Interactions*, *21*(6), 68–71. https://doi.org/10.1145/2668969

Sonboli, N., Burke, R., Liu, Z., & Mansoury, M. (2020). Fairness-aware Recommendation with librec-auto. In N.A. (Ed.), *Fourteenth ACM Conference on Recommender Systems* (pp. 594–596). ACM. https://doi.org/10.1145/3383313.3411525

Sonboli, N., Eskandanian, F., Burke, R., Liu, W., & Mobasher, B. (2020). Opportunistic Multi-aspect Fairness through Personalized Re-ranking. In T. Kuflik, I. Torre, R. Burke, & C. Gena (Eds.), *Proceedings of the 28th ACM Conference on User Modeling, Adaptation and Personalization* (pp. 239–247). ACM. https://doi.org/10.1145/3340631.3394846

Spiel, K. (2022). Disability and Play - What to consider beyond accessibility. *The Annual Symposium on Computer-Human Interaction in Play*, 324–325. https://doi.org/10.1145/3505270.3558356

Spiel, K., Hornecker, E., Williams, R. M., & Good, J. (2022). ADHD and Technology Research – Investigated by Neurodivergent Readers. In S. Barbosa, C. Lampe, C. Appert, D. A. Shamma, S. Drucker, J. Williamson, & K. Yatani (Eds.), *CHI Conference on Human Factors in Computing Systems* (pp. 1–21). ACM. https://doi.org/10.1145/3491102.3517592

Spivak, G. C. (2018, September 4). *Margins and Marginalities.* Department of English and Comparative Literature. Central University of Kerala, Kerala, India. https://www.youtube.com/watch?v=SessF1WOWwc

Statistisches Bundesamt. (2022). *Schutzsuchende nach Schutz-status, Regionen und Herkunftsländern* [Schutzsuchende nach Schutzstatus am 31.12.2021]. https://www.destatis.de/DE/Themen/Gesellschaft-Umwelt/Bevoelkerung/Migration-Integration/Tabellen/schutzsuchende-staatsangehoerigkeit-schutzstatus.html

Steel, D., Fazelpour, S., Gillette, K., Crewe, B., & Burgess, M. (2018). Multiple Diversity Concepts and Their Ethical-Epistemic Implications. *European Journal for Philosophy of Science*, *8*(3), 761–780. https://doi.org/10.1007/s13194-018-0209-5

Steinbrink, E., Reichert, L., Mende, M., & Reuter, C. (2021). Digital Privacy Perceptions of Asylum Seekers in Germany: An Empirical Study about Smartphone Usage during the Flight. *Proceedings of the ACM on Human-Computer Interaction*, *5*(CSCW2), 1–24. https://doi.org/10.1145/3479526

Steyn, M. (2015). Critical Diversity Literacy: Essentials for the 21st Century. In S. Vertovec (Ed.), *Titolo Collana. Routledge International Handbook of Diversity Studies.* Routledge.

Steyn, M. (2018). Introduction. *International Journal of Critical Diversity Studies*, *1*(1), 6–10. https://doi.org/10.13169/intecritdivestud.1.1.0006

Strohmayer, A., Bellini, R., Meissner, J., Mitchell Finnigan, S., Alabdulqader, E., Toombs, A., & Balaam, M. (2018). #CHIversity: Implications for Equality, Diversity, and Inclusion Campaigns. In E. Sari (Ed.), *Asian CHI Symposium 2018* (pp. 1–10). ACM. https://doi.org/10.1145/3170427.3188396

Stumpf, S., Peters, A., Bardzell, S., Burnett, M., Busse, D., Cauchard, J., & Churchill, E. (2020). Gender-Inclusive HCI Research and Design: A Conceptual Review. *Now Foundations and Trends*, *13*(1), 1–69. https://doi.org/10.1561/1100000056

Subramaniam, B. (2014). *Ghost Stories for Darwin: The Science of Variation and the Politics of Diversity*. University of Illinois Press. https://doi.org/10.5406/j.ctt6wr5ch

Sue, D. W. (2010). *Microaggressions in Everyday Life: Race, Gender and Sexual Orientation*. John Wiley.

Sue, D. W., Capodilupo, C. M., Torino, G. C., Bucceri, J. M., Holder, A. M. B., Nadal, K. L., & Esquilin, M. (2007). Racial Microaggressions in Everyday Life: Implications for Clinical Practice. *The American Psychologist*, *62*(4), 271–286. https://doi.org/10.1037/0003-066X.62.4.271

Suerbaum, M. (2023). Embodying Legal Precarity: Living with Ongoing Short-term Protection in Germany. *International Migration*, *61*(3), 25–38. https://doi.org/10.1111/imig.12903

Susser, D., Roessler, B., & Nissenbaum, H. F. (2019). Online Manipulation: Hidden Influences in a Digital World. *Georgetown Law Technology Review*, *4*(1), 1–45. https://doi.org/10.2139/ssrn.3306006

Talhouk, R., Balaam, M., Toombs, A. L., Garbett, A., Akik, C., Ghattas, H., Araujo-Soares, V., Ahmad, B., & Montague, K. (2019). Involving Syrian Refugees in Design Research. In S. Harrison, S. Bardzell, C. Neustaedter, & D. Tatar (Eds.), *Proceedings of the 2019 on Designing Interactive Systems Conference* (pp. 1583–1594). ACM. https://doi.org/10.1145/3322276.3322335

Taylor, J., & Hern, A. (2023, May 2). 'Godfather of AI' Geoffrey Hinton Quits Google and Warns Over Dangers of Misinformation. https://www.theguardian.com/technology/2023/may/02/geoffrey-hinton-godfather-of-ai-quits-google-warns-dangers-of-machine-learning

Thomas, J. O., Joseph, N., Williams, A., Crum, C., & Burge, J. (2018). Speaking Truth to Power: Exploring the Intersectional Experiences of Black Women in Computing. In N.A. (Ed.), *2018 Research on Equity and Sustained Participation in Engineering, Computing, and Technology (RESPECT)* (pp. 1–8). IEEE. https://doi.org/10.1109/RESPECT.2018.8491718

Thomas, S., Sauer, M., & Zalewski, I. (2018). *Unbegleitete Minderjährige Geflüchtete: Ihre Lebenssituationen und Perspektiven in Deutschland*. Transcript Verlag.

Tian, Y., Zhou, K., Lalmas, M., Liu, Y., & Pelleg, D. (2020). Cohort Modeling Based App Category Usage Prediction. In T. Kuflik, I. Torre, R. Burke, & C. Gena (Eds.), *Proceedings of the 28th ACM Conference on User Modeling, Adaptation and Personalization* (pp. 248–256). ACM. https://doi.org/10.1145/3340631.3394849

Tintarev, N [Nina]. (2017, August 31). *Presenting Diversity Aware Recommendations: Making Challenging News Acceptable.* The FATREC Workshop on Responsible Recommendation. https://scholarworks.boisestate.edu/fatrec/2017/1/9/ https://doi.org/10.18122/B2HQ41

Tran O'Leary, J., Zewde, S., Mankoff, J., & Rosner, D. K. Who Gets to Future? In (pp. 1–13). https://doi.org/10.1145/3290605.3300791

Tran O'Leary, J., Zewde, S., Mankoff, J., & Rosner, D. K. (2019). Who Gets to Future? Race, Representation, and Design Methods in Africatown. In S. Brewster, G. Fitzpatrick, A. Cox, & V. Kostakos (Eds.), *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (pp. 1–13). ACM. https://doi.org/10.1145/3290605.3300791

Trillò, T. (2018). Can The Subaltern Tweet? Reflections on Twitter as a Space of Appearance and Inequality in Accessing Visibility. *Studies on Home and Community Science*, *11*(2), 116–124. https://doi.org/10.1080/09737189.2017.1420404

Trompenaars, F., & Hampden-Turner, C. (2012). *Riding the Waves of Culture: Understanding Diversity in Global Business* (Rev. and updated 3. ed.). Brealey.

Turman, E. M. (2016). *Toward a Womanist Ethic of Incarnation: Black Bodies, the Black Church, and the Council of Chalcedon*. Palgrave Macmillan.

Turner, K., Wood, D., & D'Ignazio, C. (2021). *The Abuse and Misogynoir Playbook* [originally appeared in the January Issue of the Quarterly State of AI Ethics Report

published by the Montreal AI Ethics Institute]. MIT Media Lab. https://www.media.mit.edu/publications/abuse-and-misogynoir-playbook/

Tuzcu, P. (2020). Cyberkolonialismus und dekoloniale feministische Applikationen. In B. Hoffarth, E. Reuter, & S. Richter (Eds.), *Hildesheimer Geschlechterforschung: Band 1. Geschlecht und Medien: Räume, Deutungen, Repräsentationen* (pp. 126–147). Campus Verlag.

Twine, F. W. (2018). Technology's Invisible Women: Black Geek Girls in Silicon Valley and the Failure of Diversity Initiatives. *International Journal of Critical Diversity Studies*, *1*(1), 58–79. https://doi.org/10.13169/intecritdivestud.1.1.0058

UNESCO. (2020). *Interview: "Literacy Rate in Afghanistan Increased to 43 per cent": Interview with Mohammad Yasin Samim, Senior Technical Advisor to the Minister at the Ministry of Education of Afghanistan.* Institute for Lifelong Learning. https://uil.unesco.org/interview-literacy-rate-afghanistan-increased-43-cent

Convention Relating to the Status of Refugees, 1951. https://www.unhcr.org/3b66c2aa10.html

United States Holocaust Memorial Museum. (2021). *Persecution of Roma (Gypsies) in Prewar Germany, 1933-1939*. https://encyclopedia.ushmm.org/content/en/article/persecution-of-roma-gypsies-in-prewar-germany-1933-1939

Vertovec, S. (2012). “Diversity” and the Social Imaginary. *European Journal of Sociology*, *53*(3), 287–312. https://doi.org/10.1017/S000397561200015X

Vertovec, S. (Ed.). (2015). *Titolo Collana*. *Routledge International Handbook of Diversity Studies*. Routledge.

Vincent, J. (2020, August 18). NYPD Used Facial Recognition to Track Down Black Lives Matter Activist. *The Verge*. https://www.theverge.com/2020/8/18/21373316/nypd-facial-recognition-black-lives-matter-activist-derrick-ingram

Wachter-Boettcher, S. (2017). *Technically Wrong: Sexist Apps, Biased Algorithms, and Other Threats of Toxic Tech*. W.W. Norton & Company.

Wadud, A. (1999). *Qur’an and Woman: Rereading the Sacred Text from a Woman's Perspective*. *ACLS Humanities E-Book*. Oxford University Press.

Wakunuma, K., & Masika, R. (2017). Cloud Computing, Capabilities and Intercultural Ethics: Implications for Africa. *Telecommunications Policy*, *41*(7-8), 695–707. https://doi.org/10.1016/j.telpol.2017.07.006

Walker, A. (2004). *In Search of Our Mothers' Gardens: Womanist Prose*. Recording for the Blind & Dyslexic.

Wang, N., Chen, L., & Yang, Y. (2020). The Impacts of Item Features and User Characteristics on Users' Perceived Serendipity of Recommendations. In T. Kuflik, I. Torre, R. Burke, & C. Gena (Eds.), *Proceedings of the 28th ACM Conference on User Modeling, Adaptation and Personalization* (pp. 266–274). ACM. https://doi.org/10.1145/3340631.3394863

Warschauer, M. (2004). *Technology and Social Inclusion: Rethinking the Digital Divide*. The MIT Press.

Weber, B. M. (2015). Gender, Race, Religion, Faith? Rethinking Intersectionality in German Feminisms. *European Journal of Women's Studies*, *22*(1), 22–36. https://doi.org/10.1177/1350506814552084

Weibert, A., Krüger, M., Aal, K., Salehee, S. S., Khatib, R., Randall, D., & Wulf, V. (2019). Finding Language Classes: Designing a Digital Language Wizard with Refugees and Migrants. *Proceedings of the ACM on Human-Computer Interaction*, *3*(CSCW), 1–23. https://doi.org/10.1145/3359218

WeNet project. (2023). *The WeNet Manifesto: Building Virtual Communities*. https://www.internetofus.eu/wp-content/uploads/sites/38/2023/03/The-WeNet-Consortiums-Manifesto-2023-v8.pdf

West, C [Carolyn]. (2018). Mammy, Sapphire, Jezbel and the Bad Girls of Reality Television. In J. C. Chrisler & C. Golden (Eds.), *Lectures on the Psychology of Women* (Fifth edition, pp. 139–158). Waveland Press.

Wilkinson, M. D., Dumontier, M., Aalbersberg, I. J. J., Appleton, G., Axton, M., Baak, A., Blomberg, N., Boiten, J.-W., Da Silva Santos, L. B., Bourne, P. E., Bouwman, J., Brookes, A. J., Clark, T., Crosas, M., Dillo, I., Dumon, O., Edmunds, S., Evelo, C. T., Finkers, R., . . . Mons, B. (2016). The FAIR Guiding Principles for Scientific Data Management and Stewardship. *Scientific Data*, *3*, 1–9. https://doi.org/10.1038/sdata.2016.18

Williams, J. (2016). Pluralism, the English School and the Challenge of Normative Theory. *E-International Relations*. https://www.e-ir.info/2016/02/11/pluralism-the-english-school-and-the-challenge-of-normative-theory/

Williams, S. (2015). Digital Defense: Black Feminists Resist Violence With Hashtag Activism. *Feminist Media Studies*, *15*(2), 341–344. https://doi.org/10.1080/14680777.2015.1008744

Williams, S. (2016). #SayHerName: Using Digital Activism to Document Violence Against Black Women. *Feminist Media Studies*, *16*(5), 922–925. https://doi.org/10.1080/14680777.2016.1213574

Winchester, W. W. (2018). Afrofuturism, Inclusion, and the Design Imagination. *Interactions*, *25*(2), 41–45. https://doi.org/10.1145/3182655

Windle, J. A., Jesus, D. de, & Bartlett, L. (2020). *The Dynamics of Language and Inequality in Education*. Multilingual Matters. https://doi.org/10.21832/9781788926959

Winecoff, A. A., Brasoveanu, F., Casavant, B., Washabaugh, P., & Graham, M. (2019). Users in the Loop: A Psychologically-informed Approach to Similar Item Retrieval. In T. Bogers, A. Said, P. Brusilovsky, & D. Tikk (Eds.), *13th ACM Conference on Recommender Systems* (pp. 52–59). ACM. https://doi.org/10.1145/3298689.3347047

Winkle, K., Lagerstedt, E., Torre, I., & Offenwanger, A. (2023). 15 Years of (Who)man Robot Interaction: Reviewing the H in Human-Robot Interaction. *ACM Transactions on Human-Robot Interaction*, *12*(3), 1–28. https://doi.org/10.1145/3571718

Witteborn, S. (2015). Becoming (Im)Perceptible: Forced Migrants and Virtual Practice. *Journal of Refugee Studies*, *28*(3), 350–367. https://doi.org/10.1093/jrs/feu036

Wolf, C. T., Asad, M., & Dombrowski, L. S. (2022). Designing within Capitalism. In F. `. Mueller, S. Greuter, R. A. Khot, P. Sweetser, & M. Obrist (Eds.), *Designing Interactive Systems Conference* (pp. 439–453). ACM. https://doi.org/10.1145/3532106.3533559

Womack, Y. (2013). *Afrofuturism: The World of Black Sci-fi and Fantasy Culture*. Chicago Review Press.

Wong-Villacres, M., DiSalvo, C., Kumar, N., & DiSalvo, B. (2020). Culture in Action: Unpacking Capacities to Inform Assets-Based Design. In R. Bernhaupt, F. Mueller, D. Verweij, J. Andres, J. McGrenere, A. Cockburn, I. Avellino, A. Goguey, P. Bjørn, S. Zhao, B. P. Samson, & R. Kocielnik (Eds.), *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (pp. 1–14). ACM. https://doi.org/10.1145/3313831.3376329

Wong-Villacres, M., Kumar, A., Vishwanath, A., Karusala, N., DiSalvo, B., & Kumar, N. (2018). Designing for Intersections. In I. Koskinen, Y. Lim, T. Cerratto-Pargman, K. Chow, & W. Odom (Eds.), *Proceedings of the 2018 Designing Interactive Systems Conference* (pp. 45–58). ACM. https://doi.org/10.1145/3196709.3196794

Wörmann, S. (2003). *Afghanische Frauen zwischen Islam und Sozialismus: Gesellschaftliche Realitäten von 1920 bis 2001*. Tectum.

Wu, G., Luo, K., Sanner, S., & Soh, H. (2019). Deep Language-based Critiquing for Recommender Systems. In T. Bogers, A. Said, P. Brusilovsky, & D. Tikk (Eds.), *13th ACM Conference on Recommender Systems* (pp. 137–145). ACM. https://doi.org/10.1145/3298689.3347009

Wynter, S. (2015). On How We Mistook the Map for the Territory, and Re-Imprisoned Ourselves in Our Unbearable Wrongness of Being, of Désêtre: Black Studies Toward the Human Project. In L. R. Gordon & J. A. Gordon (Eds.), *Not Only the Master's Tools: African American Studies in Theory and Practice.* Routledge.

Young, I. M. (1990). *Justice and the Politics of Difference*. Princeton University Press.

Yuval-Davis, N. (2006). Intersectionality and Feminist Politics. *European Journal of Women's Studies*, *13*(3), 193–209. https://doi.org/10.1177/1350506806065752

Yuval-Davis, N. (2011). *Gender & Nation*. *Politics and Culture*. Sage.

Zepf, S., El Haouij, N., Lee, J., Ghandeharioun, A., Hernandez, J., & Picard, R. W. (2020). Studying Personalized Just-in-time Auditory Breathing Guides and Potential Safety Implications during Simulated Driving. In T. Kuflik, I. Torre, R. Burke, & C. Gena (Eds.), *Proceedings of the 28th ACM Conference on User Modeling, Adaptation and Personalization* (pp. 275–283). ACM. https://doi.org/10.1145/3340631.3394854

Zhao, Z., Hong, L., Wei, L., Chen, J [Jilin], Nath, A., Andrews, S., Kumthekar, A., Sathiamoorthy, M., Yi, X., & Chi, E. (2019). Recommending What Video to Watch Next: A Multitask Ranking System. In T. Bogers, A. Said, P. Brusilovsky, & D. Tikk (Eds.), *13th ACM Conference on Recommender Systems* (pp. 43–51). ACM. https://doi.org/10.1145/3298689.3346997

Zou, J., & Schiebinger, L. (2018). Ai Can Be Sexist and Racist - It's Time to Make it Fair. *Nature*, *559*, 324–326. https://doi.org/10.1038/d41586-018-05707-8

Zuboff, S. (2019). *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power*. Profile Books.