Diversity and Social Justice in Technology Design
Reflections on Diversity-Aware Technology

Laura Schelenz
University of Tuebingen, International Center for Ethics in the Sciences and Humanities

Submission date: 12 June 2022; Acceptance date: 28 September 2023; Publication date: 13 July 2023

ABSTRACT
An interdisciplinary endeavour at the intersection of American Studies, Critical Diversity Studies, as well as Science and Technology Studies, this article scrutinises so-called diversity-aware technology. A diversity-aware system is a computer system whose designers a) account for differences between the system’s stakeholders, and/or b) draw on a normative notion of diversity like “inclusion” or “fairness” in its design. Diversity concepts embedded in technology carry contested values and have effects on the technology’s stakeholders. Therefore, it is vital to conduct a critical review of designs leveraging diversity concepts. In an exploration of three cases (diversity-aware datasets, machine learning fairness, and diversity-aware social media), the article sheds light on the shortcomings of mainstream or “individual-level” diversity-aware technology. Such technology leverages individual-level notions of diversity (demographics, personality, culture) to cater to users, thereby obscuring social inequalities among them. Inspired by Black feminism and critical race theory, the article offers a social-justice-oriented conceptualisation of diversity-aware technology. It develops a definition and criteria for critical or “structural-level” diversity-aware technology, where diversity concepts are linked to the visibility and redistribution of power. The article offers inspiration for researchers of technology and designers who work with diversity concepts.

KEYWORDS
Diversity, Diversity-aware technology, Black Feminism, Design, Social Justice

Introduction
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, 2022). 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.

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. 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. 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 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”.

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. It theorises the experiences of Black women as a group, including their intersecting oppression and acts of resistance (Collins, 2000, 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, 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 [539 US 306 (2003)], 23 June 2003). Diversity was 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). 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

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

<table>
<thead>
<tr>
<th>Diversity-Aware Technology</th>
<th>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.</th>
</tr>
</thead>
</table>
| **Individual-level (or mainstream) diversity-aware technology** | **Definition of difference:**  
• demographic differences  
• cultural differences  
• differences in personality and psychological aspects  
• (dis)ability  
• different social practices  

**Normative notion/value of diversity:**  
• tolerance  
• pluralism  
• multiculturalism  
• inclusion  
• fairness  
• equality |
| **Structural-level (or critical) diversity-aware technology** | **Definition of difference:**  
• different experiences of the stakeholders  
• different levels of privilege and oppression  
• structural inequalities  
• power relations  

**Normative notion/value of diversity:**  
• social justice  
• representation  
• redistribution of resources and power |

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 (2000), 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, 2000, 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 (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, 389f). 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 (2000). 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, 2000, 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”. Both Erete et al. (2018) and 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 (2020) 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). In their study “Gender shades”, Buolamwini and Gebru (2018) find that popular facial recognition software fails to recognise and classify Black women 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 (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). 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 the WeNet 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 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, 2000; 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., 2023, 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, 2000, 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 (American Civil Liberties Union [ACLU], 2021; 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 (American Civil Liberties Union [ACLU], 2020). 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 (2019a, 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, 2020), 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.7
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 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. 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, 2000, 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 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, cop. 1995, 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.

ACKNOWLEDGEMENTS

I would like to thank the anonymous reviewers for their feedback on this article. This project has received funding from the European Union under grant agreement No 823783 (WeNet – The Internet of Us). Many thanks to Professor Dr Astrid Franke and Professor Dr Daniel Gatica-Perez for their advice. Many thanks to the participants of the colloquium in American Studies at the University of Tuebingen, Germany, with Professor Dr Astrid Franke for discussing earlier drafts of this article.

NOTES

1. By technology stakeholders, I refer to users but also those who are targeted or affected by the technology without actively making use of it.
2. 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, 2022).
3. 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.
4. 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.
5. Another example is the FairFace dataset by Kärkkäinen and Joo (2019).

6. Of course, in order to qualify as diversity-aware technology, the WeNet platform would also have to fulfill 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.

7. While there are multiple streams of BF (see Collins, 2000, p. 22), this article works with a US-centric version of Black feminist thought due to my expertise in this area.

8. 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.

REFERENCES


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.


