Journal of Humanistic Mathematics

Volume 12 | Issue 2 July 2022

Beyond Ethics: Considerations for Centering Equity-Minded Data Science

Nathan Alexander
Morehouse College

Carrie Diaz Eaton
Bates College

Anelise H. Shrout
Bates College

Belin Tsinnajinnie
Santa Fe Community College

Krystal Tsosie
Native BioData Consortium, Vanderbilt University

Recommended Citation

Nathan Alexander, Carrie Diaz Eaton, Anelise H. Shrout, Belin Tsinnajinnie & Krystal Tsosie, "Beyond Ethics: Considerations for Centering Equity-Minded Data Science," *Journal of Humanistic Mathematics*, Volume 12 Issue 2 (July 2022), pages 254-300. DOI: 10.5642/jhummath.OCYS6929. Available at: https://scholarship.claremont.edu/jhm/vol12/iss2/14

Beyond Ethics: Considerations for Centering Equity-Minded Data Science

Cover Page Footnote

This work was supported in part by a working group funded by the RIOS Institute and supported by the William and Flora Hewlett Foundation. Correspondence regarding this article should be sent to Nathan N. Alexander, Department of Mathematics, Division of Science, Technology, Engineering, and Mathematics, Morehouse College, 830 Westview Drive SW, Atlanta, GA, 30314. Email: nathan.alexander@morehouse.edu

Beyond Ethics: Considerations for Centering Equity-Minded Data Science

Nathan Alexander

Morehouse College, Atlanta, Georgia, USA
nathan.alexander@morehouse.edu

Carrie Diaz Eaton

Bates College, Lewiston, Maine, USA
cdeaton@bates.edu

Anelise H. Shrout

Bates College, Lewiston, Maine, USA
ashrout@bates.edu

Belin Tsinnajinnie

Santa Fe Community College, Santa Fe, New Mexico, USA
belinmt@gmail.com

Krystal Tsosie

Native BioData Consortium, Vanderbilt University, Nashville, Tennessee, USA
krystal@nativebio.org

Abstract

In this paper, we utilize duoethnography — a research method in which practitioners discursively interrogate the relationships between culture, context, and the mechanisms which shape individual autobiographical experiences — to explore what may be beyond ethics in the context of data science. Although ethical frameworks have the ability to reflect cultural priorities, a singular view of ethics, as we explore, often fails to speak to the multiple and diverse priorities held both within and across institutional spaces. To that end, this paper explores multiple perspectives, epistemologies, and worldviews that can collectively push

researchers towards considerations of a data science education that is equity-minded both in concept and practice. Through a set of dialogues which examine our positionalities, journeys, ethics, local cultures, and accountabilities, this paper explores the contextual realities rooted in the authors' educational settings. These conversations focus on the humanity of our students, the communities from which we come from and serve, as well as the unintentional harms and possibilities associated with the development of data science programs across institutional types. We take a set of five core questions to examine how we made, and continue to make, sense of our diverse cultural perspectives on data science education and equity with/in relation to others' realities. Broadly, this paper seeks to offer reflections on the related but differing functions of ethics and equity in data science education.

Keywords: data ethics, equity, data science, education, social justice

Introduction

As various sectors increase their demands for data science skills, competing discourses regarding the role of academia in data science education are becoming more context-specific [46, 64, 81]. There are guidelines from related disciplines — such as computer science [1], the mathematical and statistical sciences [60, 2], and statistics education [9, 22] — which serve as references for the development of undergraduate data science academic programs [30]. However, less attention has been paid to the related but differing functions of ethics [69, 84, 85] and equity [55, 70, 72] in data science education. Historically, data science has been an industry-led domain driven by the needs and heuristics of business leaders and real-world problems and applications that are often positioned in contrast to academic inquiry (see, for example, [62]). As industries continue to privilege data science practices that center capital acquisition, profit, and rapid innovation [34], there is a need to understand data ethics and its relations to equity-minded, human-centered, and justice-oriented data science practices [56].

As postsecondary educators, especially those in mathematics, statistics, and computer science departments, begin to propose, develop, and update data science curricula and programs, there are new opportunities to pose questions not generally considered by industry-based data scientists. Certain

questions allow both academicians and practitioners to probe some of the widely held assumptions about ethics in data science education and explore the role of context in framing and understanding equity. Specifically, this moment presents space for the field to collaboratively explore, critique, and expand the collective understanding of the cultural, historical, and sociopolitical dimensions of equity in the development of data science and data literacy pedagogies, and in computational analyses more broadly. These reflections have an ability to result in the identification of various intersections that can help speak to the various themes and realities rooted in possible futures for a more equity-minded data science practice. As authors, we understood that the current state and narratives around ethics and equity in data science, while related, are different in terms of practice. As a result, we sought to collectively identify what might exist beyond a singular conception, definition, or framework on ethics in data science.

In this article, we employ duoethnography [16, 50, 66, 67] as a dialogic method [7] to examine heterogeneity in our understanding of ethics and equity. In this process, we mark similarities, variations, contrasts, and possibilities with/in our approaches to data science education. As a result, our concept of equity-mindedness in data science sits at multiple intersections. For example, one intersection allows us to reconsider the meaning of diversity in our contexts, while also providing space for us to examine some common systems at play within and across these various contexts; another intersection takes into consideration our different ways of knowing, many of which are rooted in critical and indigenous methodologies, such as storytelling [5, 29], and others build on some of our understanding of research across the fields of mathematics education and science education, and specifically research that relates to identity-related and sociopolitical issues in both K-12 and undergraduate settings [3, 13, 17, 39, 43, 52, 57, 58, 61, 82, 85].

Through duoethnography, we present one method of exploring and understanding the dimensions of an equity-minded data science based on the contexts of our lived experiences and material realities, our progress towards new realities, and the various pathways to our collective sharing. More broadly, we sought to interrogate ethics and equity as well-configured, normalized, and static terms in data science discourses. From these terms we exercise equity-mindedness as a collaborative social construction to be understood by

our various modes of practice situated across our various institutions, their cultures, how learning occurs across each of these contexts, and the role of individual and collective actions towards criticality and change.

The desire for these reflections came as a result of our noticing that terms such as ethics, equity, and social justice tend to induce individual social reflection as opposed to formations of practice that respond to local systems and result in more immediate and noticeable developments, such as community-centered pedagogies or a collective "deliberation of how [data science] tools might fortify, attenuate, or alter learning opportunities and relationships of power in the classroom" [72]. Despite longstanding calls for improving culturally relevant pedagogies [54] and culturally responsive teaching [38] in science, technology, engineering, and mathematics (STEM) education, data science education has yet to consider some important characteristics of culture and context, as well as other resource-based pedagogies (e.g., culturally sustaining pedagogies [71]; funds of knowledge [63]). Conversely, data science education, especially in postsecondary contexts, has centered on the development of technical skills based on the field's history as an industry-driven domain.

We begin in the next section with a broad exploration of the role of industry and various perspectives on ethics in the field of data science. In a third section, we provide a more extensive discussion of the duoethnographic approach, and we outline five core questions whose responses are presented in a set of evolving dialogues. These dialogues help situate our positionalities as authors, our journeys to the moments of our dialogic engagements, and we consider multiple meanings of ethics and culture within and across our individual contexts; we end with a dialogue on accountability to ourselves and to our individual and collective communities. In a closing section, we explore equity-mindedness further based on the intersections that were observed in a set of summary statements provided after each of the five dialogues. We close with recommendations for the field.

Ethics in Data Science

There are a wide-range of definitions, frameworks, and perspectives on ethics in data science, statistical data analysis, and related areas such as machine learning (ML) and artificial intelligence (AI) [12, 24, 35, 36, 37, 49, 64, 70,

75, 77, 86]. As discourses around ethics, equity, and social justice become increasingly commonplace in data science, specifically, and computation more generally, there is a need to underscore the role of context and identity in making sense of how the field defines and applies these terms in classroom settings. That is, while industry-related applications of ethics may be tied to specific approaches which support data use and privacy, less is known about the pedagogical practices employed for ethics discussions in undergraduate data science education.

Recent work has offered some practical suggestions to integrate ethics into data science courses [11, 18]. For the purpose of the current discussion, however, we posit that in addition to these conversations, additional work needs to be done to understand the related but differing functions of ethics and equity in data science education. To this end, we continue with a brief exploration of frameworks and research on data ethics in undergraduate settings.

Current policy reports on data science education have stressed the need for an ethics code of conduct (see, for example, [64, page 32 and Appendix D]) or directly suggest integrating ethics into data science education [11, 26, 30, 64]. However, there is often less focus on the preparatory work necessary (and what we might term as increasing equity-mindedness) to put these suggestions into practice. The field often assumes that there is a coherent set of practices that are axiomatically ethical in the field, or that there is consistency in their classroom implementation [81]. We remind the reader that ethics, however, are neither fixed nor objective. A lack of curricular implementation in more critical conversations in undergraduate mathematical settings may also be the result of a lack of awareness, comfort, or community [4].

De Veaux *et al.*, in *Curriculum Guidelines for Undergraduate Programs in Data Science* [30], identify ethics as one of six major subject areas for the data science major: "(1) data description and modeling, (2) mathematical foundations, (3) computational thinking, (4) statistical thinking, (5) data modeling, and (6) communication, reproducibility, and ethics" are identified in their curriculum guidelines [page 9]. Here we note that ethics, in this particular framework, has been grouped together with other topics of study, i.e., communication and reproducibility. The authors also discuss the issue of knowledge transfer earlier in the report and in doing so, they group ethics

with reproducibility to make the following note:

> The capabilities of data science introduce new ethical questions. Programs in data science should feature exposure to and ethical training in areas such as citation and data ownership, security and sensitivity of data, consequences and privacy concerns of data analysis, and the professionalism of transparency and reproducibility [page 8].

Elsewhere, the authors list ethics as a related, as opposed to central, course for data science majors. However, little detail is provided for the reader around what might constitute a course in ethics beyond. The reasons for this grouping with other topics of study in some places and as a standalone topic in others can be due to a host of reasons. From a high-level perspective, this might provide an opportunity for different institutions to decide how their approach to engaging ethics fits within a specific curricular framework. From a more granular perspective, however, a lack of detail points to a general use of the term 'ethics' in the field of data science with little to no acknowledgment of the various ways that it might constitute a study of equity issues related to the use of data. In consequence, some of the questions that we respond to as a group offered an opportunity for us to consider a set of approaches to the design and development of data science programs in relation to the intersections of ethics and equity.

The National Academies of Science, Engineering and Medicine's 2018 report, *Data Science for Undergraduates: Options and Opportunities* [64], describes ten data acumen skills, one of which is ethics. These skills are not contextualized in the context of a specific curriculum. However, the proposed Hippocratic Oath for data science in Appendix D begins to outline what might be considered ethical practices for data science in undergraduate settings. We do not know yet to what extent students are introduced to this Hippocratic Oath, Ethical Guidelines for Statical Practice (CPE, 2022), or similar documents in the curriculum. Nor do we know to what extent these conversations are connected to considerations of equity and justice. Some of these guidelines for data science community practice have been grounded in equity and justice (see for example, [26]), but some of them fail to center these connections [79].

For example, a more philosophical approach to the study of ethics in fields such as moral philosophy distinguish between normative ethics — arguments about how the world ought to be — and descriptive ethics — arguments about how different populations currently see the world. Similar calls for "ethics in data science" [8, 76] based on these frames of reference, seem to invoke a set of normative ethics, that is, the need to do data science in a way that reflects a more just world based on a set of moral principles. However, we as a collaborative group ask: "Whose moral principles ought to be at the center of this project?" We find that these answers remain obscure or are detached from many context-specific and material conditions which generate the questions.

To interrogate these ideas, this paper explores the value and meaning of multiple histories and futurities in conceptions of data science. We seek to go beyond ethics as a means to advance towards the material realities of equity-mindedness, as opposed to some social construct primarily used for individual reflection. We argue that equity-mindedness in data science is an ever-evolving set of practices — and more appropriately the collective movement toward a praxis — that combines a set of central questions to consider within and across various contexts. In this way, ethics and equity become deeply rooted in the local context while also being grounded in some larger discourse that seeks to make sense of the many and varied ways that one might make sense of the overlaps and difference in the two terms.

Some related research has called for more humanistic stance in data science education [56]. This work has happened in different ways in the field of mathematics education and, more recently, science education. For example, some of the early iterations of related socio-political consciousness for the field of mathematics education was to move beyond static discourses on race and power in the analysis of data on student mathematics course outcomes. For example, instead of an overemphasis on racial achievement gaps, what some have called "gap-gazing" [42, 44], there was a gradual shift to examining power and systems in the field [58]. Relatedly, by examining the current participation numbers in data science in the name of diversity, equity, and inclusion, and defining equity as a gapless end state, we sought to contemplate how we are constructing equity in more materialized senses. We contend that conceptions of ethics and diversity in data science necessitate diversi-

fying discourses around multiple approaches to reflect the wide expanse of cultural and ideological experiences and perspectives. Moreover, we observe that current and future practitioners of data science bring to the field perspectives in which to consider ways that we may avoid replicating patterns of marginalization in data science and mathematics.

Eve Tuck and K. Wayne Yang, in the forward to their book *Towards What Justice* [83], write that "social Justice is a way to mark a distinction from the origins and habits of almost all disciplines which emerged in the 19th and 20th centuries and are rooted in colonialism and white supremacy" [page 4]. As we explore ethics and equity within the mode of diversification, we acknowledge how such efforts may speak to different, and diverse, dreams of justice [83, page 1]. A goal of this paper then is to integrate stories of our experiences that act as counternarratives to the discourses which dominate quantitative and data-focused education policies [27].

These voices may also lead to different interpretations of social justice, though we find ourselves in intersecting spaces in solidarity with each other's projects [4, 83, 15]. Keeping our positionalities centered as strengths allows us, the authors, to acknowledge both the historical and contemporary realities that shape our experience of data science. It also allows us to consider new futures for data science that extend beyond what some have considered myopic views of ethics and equity. This collective approach, which has been practiced by critical scholars across the field of education allows multiple voices to come together to identify new representations and consider new priorities.

Method

For the current discussion, we adopt an emerging duoethnography methodology (see, e.g., [16]). Duoethnography is a research approach where multiple researchers discursively interrogate the relationships between the cultural contexts and mechanisms which shape their own autobiographical experiences. In this way, duoethnographies have the potential to reveal interactions between narratives and identities that intersect in a particular cultural space [5]. In addition, we draw on Indigenous scholarship centering narrative and storytelling as data [29, 47, 68, 88]; and the roles of sociohistorical contexts and identity in mathematics education research [59, 78].

We are a group of five researchers who are working to build data science education programming. We come to data science with a variety of personal and professional experiences. Our research was constituted by three steps: preliminary discussions, iterative discussions, and synthesis. As part of our preliminary discussions, we met several times to discuss possible overlapping interests. Some of us have collaborated in various capacities before with others, but most are new collaborative relationships. Both our research and the research process itself were emergent. We defined our methods iteratively through discussion and revision, and documented our methods as they changed and developed. In these initial group sessions, we co-composed a set of framing questions that would allow us to explore the relationships between positionality, community, data science and ethics. These framing questions are reflected within the Dialogues section. When we had finalized the questions, we moved on to iterative discussions. In this phase of discussion, each of us responded in writing to a set of questions:

(a) What identities do we hold that inform our perspectives? How do these identities impact our perspectives?

(b) How did we come to program-building in data science?

(c) What does ethics mean for our data science programs? How are we going beyond ethics in our data science education programs?

(d) How do themes of the past, present and future of data science and/or our communities play into our work?

(e) What are some things we are holding true to as we build out these programs and do this work? How are we keeping ourselves and/or our communities accountable?

We each answered each question asynchronously as a first pass, using Google Docs as a collaborative editor. We then used collaborative meeting time to respond to each other and to verbally discuss our overlapping interests or emerging themes. We also documented these discussions in the document. Our final research stage digested and synthesized key themes and outcomes. Each author independently encoded one or more sections with a set of thematic tags that they developed. These tags necessarily reflected our own

subjectivities and positionalities. We then compared our thematic tags and where we applied them within each section. In the process, we reflected on the ways in which our positionalities and subjectivities shaped our readings of the answers to each section. Taken together, these tags that were produced individually but collectively honed and discussed became the basis for a short synthesis after each dialogic prompt, as well as a more comprehensive synthesis at the end of the paper. Finally, we re-examined the full text of this document with these key emergent ideas in mind, reshaping the paper as necessary to help guide the reader in this paper.

A note on language: As part of this process, we also discussed the implications of language choices and the use of specific terms. We come to this paper from different positionalities and with different linguistic practices. Part of our work was to recognize those differences, and agree, for the purposes of this work, on shared practices. We decided to render instances of white as a racial identity in lower case, because we do not want to give whiteness any additional power or reify it as a category. This decision does not reflect consensus among scholars or journalists, but does reflect our understanding of the needs of this paper and audience [6, 23, 65]. We use this as an example of lack of agreement on the language even within our communities, and our individual decision to capitalize may vary based on the context outside of this paper.

In our answers to these questions, we individually and collectively argue that ethical frameworks reflect cultural priorities. In predominantly white fields, and at predominately white institutions, any effort to craft ethical frameworks will necessarily reflect racialized understandings of what is just and ethical. In order to counternarrate this hegemony, our methodology centers multiple perspectives, epistemologies and worldviews that can collectively push us towards better data science education. We consider the humanity of the people we are educating, the communities we are serving, and the data science harm or good we are enacting. It is critically important then to understand our own cultural perspectives on data science and data science education in relation to others as a means of decentering whiteness and challenging the objectivity of removing these perspectives.

We are excited to share this journey with you. We recognize that this paper might look different from those that you are used to reading. We invite you to

reflect on the questions that we responded to, consider your positionalities as they intersect or differ from ours, and consider your responses in conversation with ours. The intersections sections at the end of each dialog are intended to highlight the themes that we collectively recognized in our own answers. In particular, we wish to draw attention to the following themes:

- That ethics in mathematics and data science need not be one fixed idea or approach, but rather reflective of the multiple approaches, epistemologies and worldviews that are already held by practitioners of data science.
- That we each have different dreams, but we share the challenges of assimilationist thinking that plagues policy and practices no matter where we are.
- That our communities inform who we are doing this with and for, and that our communities are both important responsibilities as well as sources of support.

Equity-Mindedness through Collaborative Dialogue

In this section, we offer readers a view into our collaborative process and practices that were grounded by duo-ethnographic methods. Through a series of five dialogues, we examine the diverse and intersecting, as well as competing, priorities across institutional settings in relation to building out what we have framed as equity-mindedness in the material sense. Namely, we seek to explore the ways that our diverse experiences extend, again, beyond ethics and towards more complex conceptions of our educational practices. To that end, we encourage readers to not only read along but also engage in the dialogues by writing through their own histories, perspectives, and settings.

Dialogue 1: Positionality

What identities do we hold that informs our perspectives? How do these identities impact our perspectives?

CDE: I carry with me many identities which are marginalized in society and STEM. At different times, different identities seem to be driving my life. For example, when I was pregnant with my oldest child, it was my identity as a mother that drove many of my actions and decisions. My son taught me about appreciating neurodiversity in the classroom instead of penalizing it, even before I really understood my own neurodivergence. Like my neurodivergence, many of my identities have been easier to hide than some, and I toe the line between marginalized and privileged as I navigate acceptance of both. For example, I identify as queer, but enjoy both the privilege of a heterosexual marriage and the marginalization of being identified as a woman. My father is from Peru and my mother is from New England — thus I am Latinx, born in the US, but by genetic chance am often read as white, despite my dark hair and eyes. This compounded my struggle to be recognized as a part of the Latinx community I loved — because my family had been encouraged to assimilate, because I did not grow up speaking Spanish fluently, because I looked too much like my mom. Dueling identities between self-perception and societal perception drove me to act as a boundary spanner. I often played translator, peacemaker, and educator between my parents and between members of my diverse communities. Later, I adopted similar boundary spanning disciplinary identities — mathematician, biologist, STEM Education researcher, data scientist — and eventually reclaiming social justice advocate as part of my academic identity.

NNA: My identities and personal development have been based on the various phenomena in my local surroundings and social contexts. For example, as an undergraduate student attending a predominantly white institution (PWI), I was very engaged in organizations that supported my own sense making in relation to my racial identity. However, this development came as a result of me consistently framing my own identity using a "minority-majority" paradigm. Now, working at an historically black college, I find that many other identities come into the forefront for my students as I develop both as learner and their teacher. One of those is a set of specific histories, privileges, and issues in our positionality within the broader world as Black male-identified persons. When I am abroad, as some others may experience, I am often more immediately aware of and critical of my position as a citizen of the United States. As a result, I consider this shifting and

flexible nature of our identity formation to be context-based as discussed by education scholars, and how these contexts often situate our experiences and perspectives in terms of the various spaces that we find ourselves in as we think about the world.

BMT: Shí éí Ilocano nishłį́ dóó Ta'chii nii bashishchiin. Ilocano dashicheii dóó Tsi'naajinii dashinalí. Ȧkót'éego diné nishłį́. (English translation: I am Ilocano/Filipinx, born for the Red Running into the Water clan. My maternal grandfather is Ilocano and my paternal grandfather's clan is Black Streak in the Forest.) My mother's parents grew up in the Philippines, then fled the violence of war to Hawai'i. My mother was born and raised in Hawai'i, but my sisters and I were raised in the community that my father was raised in, Na' Neelzhiin. My parents were both educators, committing to serve their communities through teaching in schools. This form of service to our community through formalized education greatly influenced my decision to transition to recenter my studies from mathematics to mathematics education.

I think about and am asked a lot about my Diné, Filipinx, Indigenous identities, transitioning from mathematics to mathematics education, search for a career that is close to home, working towards service to communities via TCUs and Community Colleges. Being a parent, committed to family has a different kind of stakes with respect to my work as an educator that I am still grappling with. Also I started training in decontextualized mathematics (I had been moving towards focusing on research in computational group theory). In switching to mathematics education, focusing on equity there was a shift in the ways I needed to think about mathematics in social, cultural, and political contexts. Being an Indigenous person in mathematics education often means that I am asked to provide perspective on culturally relevant or contextualized mathematics for Indigenous communities. Well, this then requires that I grapple with tackling what counts as mathematics and simultaneously grappling with questions of Indigeneity in the 21st century and the nature of my own connections to Native communities.

> CDE: Belin - I do not think I knew before your writing that you were also Filipinx. But there is something about the way we present to others and are presented by others that often is less dimensional than reality. My dad assimilated into the US, marry-

ing a white woman and raising his kids without speaking Spanish and thousands of miles from his family, so all of these factors led to the erasing of this side of my identity. I was left with only symbols of who I was — a copper plate with the Incan Sun God from Peru that I would take to show and tell. For years, I would bring my dad to any "Latinx community" event that I wanted to go to so that I could assert legitimacy. I think I only started healing much later in life after visiting Peru myself, experiencing the place and the artifacts of their science and engineering they left behind, and realizing that I embody part of their legacy as a natural philosopher and mathematician.

BMT: Carrie, I am constantly grappling with the sense that I am not always presenting my Filipinx background. I often don't know or feel that I can claim it, having lived in Indigenous communities on the continent that many now call North America. I think like you after visiting Peru, I felt more in tune with my Filipinx heritage after each of my maternal grandparents passed away and visiting my mother's family in Hawai'i. I was guided through the mourning and healing processes through my Filipinx relatives. Currently, I feel "legitimacy" in my Filipinx identities through the foods that I've learned to prepare through my mother and my sisters. Also, I am more aware about the complexities of the Indigeneity of Filipinx peoples. I also think about how even when I say my clans in identifying myself in the Diné way, I am resisting blood quantum notions of identities, knowing that being Diné in this way means that I am also fully embracing being Filipinx.

AHS: I am a white woman who has worked mostly in predominantly white institutions, and in spaces designed to make me, and people like me, comfortable. I was trained as an historian of the nineteenth-century Atlantic world, which has led me to think about the ways in which historical processes — most notably racism and colonialism — shaped the world in the past and continue to shape the world we live in today. The nineteenth century was a time of increased data production. Colonizer and settler governments used data to enact violence on indigenous communities, enslavers collected data about the people who they held in bondage and governments around the

world used data to demarcate "desirable" and "undesirable" citizens and subjects. My research concerns the data that was produced in these conditions, as well as methodologies that acknowledge and disrupt the violence inherent in historical data. In the time that I have been a working professional in academia, I have been located in programs and departments that are primarily identified with the humanities and social sciences. More recently, I work in an interdisciplinary program with strong connections to STEM and Data Science. Women are not disproportionately historically marginalized in history departments (that is, any more than they are in the academy as a whole). In fact, all of the faculty appointed to the Atlantic history program at NYU (where I did my PhD) identified as women. Moving into STEM adjacent spaces has meant that while my gender identity has not changed, the way that it, and I, am perceived has been.

KST: Shí éí Kinłichii'nii nishłį́ dóó Naakai diné'e bashishchiin. Tòdichii'nii dashicheii dóó Tłizi'łaní dashinalí. Ȧkót'éego diné asdzáán nishłį́. (English translation: I am Red House clan, born for the Mexican Peoples clan. My maternal grandfather's clan is Bitterwater and my paternal grandfather's clan is Many Goats. In this way, I am a Navajo woman.) First and foremost, I identify as being an Indigenous citizen of the Navajo Nation. In our tradition, it is important to identify our ancestors even before our own names to acknowledge those who came before us and to acknowledge our relationalities and kinship to each other. It is this Indigenous dimensionality that grounds my sense of belonging and relative isolation in the space of data science, which has a severe underrepresentation of Native American women (Corbett & Hill, 2015) Considering that most of my waking moments for the past couple of decades have been spent in higher education, it makes sense that I next identify as a population geneticist and biostatistician with intersections in ethics and public health. The added stratification of education attainment as an Indigenous woman with graduate STEM degrees means that not only do I have a narrow circle of peers but that I also have a huge sense of responsibility in advocating for data equity for my people.

Dialogue 1 (Positionality) summary.

This prompt for this discussion suggested that the identities we hold inform our perspectives. Though we found diversities across our racial, gender, and

professional identities and experiences, we were drawn to the connections that we made through our dialogues. What our discussion reflects are the complexities and dynamics of the ways our identities interplay with our goals, our decisions, and the spaces that we occupy. We reflected on the ways that some of our identities, such as our racial and gender identities have been historically marginalized and how we are aware of that ongoing marginalization in the spaces that we occupy today. We refer to how these identities intertwine with roles that we have taken on, such as our professional identities or our identities as parents. The ways that we embrace and are informed by our identities are not static. We share how we may question and/or assert our sense of belonging in various spaces. Our identities play a significant role in our sense of responsibility in advocating for our communities in spaces that have worked to actively marginalize our communities. The ways in which we describe the ways that racial and gender identities intertwine with ongoing professional goals. There is a theme of the constant negotiation within ourselves in our awareness of the ongoing marginalization of communities that are dear to us and also the sense of privilege of being in the positions that we are in. The ways that we embrace our identities are dynamic, shifting, and evolving with our agencies, goals, and sense of responsibilities as we navigate our professional, academic, and personal journeys (e.g., Grande, 2015; Hull & Greeno, 2006; Nasir, 2002). As with our past, present, and futures, our approaches to engaging in data science education are intertwined with our positionalities. As with all the work that we engage in, we bring the ways that our communities are marginalized, the privileges that we are afforded, and our relationships to our past, present, and future to our work and conversations in data science.

DIALOGUE 2: Journey

How did we come to program-building in data science?

NNA: Many of my interests while learning mathematics in school tended toward real-world applications; I always enjoyed the challenge of word problems. I was also interested in statistical applications as an undergraduate, and I spent a lot of time exploring fractals in nature, literature, and dance, as well as coding and analyzing data. As data science has become a more popular term in academia, I began to think about how my background and

interests in these real-world applications might contribute to a more critical analysis of quantification in the contexts of African American life. In making connections with colleagues who are doing similar work, I believe that we began to see some of our different but overlapping dreams of ensuring that the negative aspects of what we have experienced in the learning of mathematics would not be carried over into this expanding interest in education research focused on quantification, statistical analysis, and data science education. On another end, I felt that some of my colleagues from graduate school seem to have committed to the publish-or-perish mindset often present in academic circles, whereas I wanted to better understand the dimensions of more critical work as it might play out by teaching students directly in a classroom setting. Additionally, program building has provided a way for me to integrate these two in ways that speak to our overlapping dreams.

> AHS: I want to dwell for a bit in "dreams" — because I think so often data science programs are expected to be data-driven which can mean privileging certain kinds of evidence (i.e. what fields are most lucrative for particular students with particular positionalities) and ignoring others, like the evidence that we bring, differently about what it is possible or desirable to dream for.
>
> NNA: This is an important point, especially when defining what programs constitute as "data," especially in the development of courses and learning outcomes. Recently, I have been considering what Irizarry (2020) notes about data science being a collection of skills and modes of doing, which involves many different individuals who may have non-overlapping skills. In this way, defining "data" has the ability to take on new meanings. So the idea that someone is or is not a "data scientist" is disrupted. For me, this removes one of the often unspoken barriers to many people's dreams of identifying as a "mathematician"; the current price is a PhD in mathematics ... but everyone does mathematics and each one of us can define what mathematics is, or what constitutes mathematics, for our context or thinking. These are the dreams I speak of: programs that are not only inclusive, equitable, or ethical in name but also in action and identifiable practices. Despite these dreams, data is still often viewed in the flat context of

the qualitative versus the quantitative; and we still see that the quantitative dimension is often privileged over others in terms of science. So, to your point again, we want to be sure that privilege like this does not carry over when dreaming.

CDE: I haven't called it as much dreams, but that leans into Tuck and Yang, which we have been reading together during this journey. I had been thinking about our imaginations as a form of futurism. Dreams feel intangible, though full of meaning. Imagining our futures feels like something we may have the possibility to live. But perhaps I should think of dreaming more when I get stuck in the limits of reality.

NNA: Appreciate this thinking CDE.

BMT: At my institution, my input as a mathematics educator has been sought for in developing grant proposals. I've been interested in providing input because I've been wanting to rethink our mathematics offerings at the community college that decentralized the calculus track. One goal for me is to also rethink options for those who do not test into current college level offerings. What role can a focus in data science education play in reshaping mathematics in community college contexts?

AHS: I come to data science via the digital humanities. In the nearly ten years since I completed my PhD I have worked as a visiting assistant professor of history, a postdoc in digital studies, an assistant professor of digital history and (now) an assistant professor of digital and computational studies. I came to Bates College (and in this latter role) to help build a new interdisciplinary program that sought to "infuse computing throughout the curriculum" without replicating the racialized, gendered classist and otherwise exclusionary practices often seen in traditional computer science programs. This work is difficult because so many of the established pathways are built on colonizer and/or white supremicist educational structures. In DCS at Bates, we hope that data science has the potential to bring together computational and quantitative literacies, critical theories required to see and understand the ways in which power operates in data collection and analysis, and the human context that make data analysis meaningful.

CDE: After spending eight years as a mathematician in a biology department,

working to develop the mathematical and computational skills of biology and environmental science students, I moved to Bates College at the same time as AHS to help build the Digital and Computational Studies (DCS) program. The program strives to connect digital and computational critiques and approaches to multiple departments with intentionality towards equity and social justice. It was the commitment to both that attracted me - this was not going to be just another computer science department, but something more culturally relevant to me, personally. DCS instituted a minor last year, not specifically in data science, but it prepares students to flexibly pursue careers in tech broadly speaking. This minor structuralizes our intentions to send responsible graduates into the industry by requiring a course in critical digital studies.

> NNA: In much of my (our?) dreaming for today's mathematical education is the need to consider how more humanizing practices can follow in line with some of the processes you have outlined. These can be instituted with a specific goal of challenging what we think we know. We know that a silver bullet does not exist but what are some of the non-negotiable points in our thinking about how to support students' mathematical development and sociopolitical consciousness, especially during their transition to college. One example is to develop a first-year experience (FYE) course where students get to understand some of the critical theoretical approaches to mathematical thinking and practice. This introduction can be framed around a set of case studies using various forms of data. Additional work can be found through online repositories and on social media (e.g., GitHub and, more specifically, the #DuBoisChallenge). There are many options when considering the varied methods used to approach quantification and literacy development using a historical or justice-based lens. The more we explore departmental practices and norms against what we say we believe but have failed to put into motions can provide us with knowledge about the varied ways of uncovering what we say and what we do. Reading research alone can not get us to some equitable place, and our pedagogical priorities should be a way for us to extend these questions to improve our classroom practices.

CDE: Ahh - "departmental priorities!" I think they really constrain what we can do here. Who is determining those priorities, who are we centering there?

NNA: Yes, in my case I often feel vulnerable as a pre-tenure faculty trying to learn more about where things have been as I think about what is truly possible; and at times it feels inauthentic in some ways, or as an attempt to follow the norms that have been set versus do what we know is right under, often under the guise of something elseâĂęthis is all political.

CDE: YEESS — authenticity and the struggle to hold onto yourself. I wrote a poem about this a few years ago [32]. About a fear I hold deep: losing myself to the demands of assimilation. About being kept away and sheltered because others thought it wasn't my time. And about losing myself when the time finally arrives.

KST: I am a former tribal college researcher and educator. In addition to conducting community-engaged genomics research, I spent two years teaching mathematics and statistics to Indigenous students at a rural tribal college. Tribal colleges and universities (TCUs) are unique minority-serving institutions because each of the 37 TCUs nationwide operates separately within bounds of tribal sovereignties. As smaller institutions, a majority of TCUs are 2-year institutions and relatively few have bachelor-level degree programs in STEM fields, nonetheless in emerging fields like data science. I had the privilege to not only teach Indigenous-centered curricula in mathematics , but also had the opportunity to really think on the challenges and possibilities of building Indigenous-led data science modules for tribal college students from heterogeneous disciplines and institutional types. The need for online training opportunities was prevalent by nature of the rurality of TCUs, but it became even more apparent when distance-learning options were forced by the onset of the COVID-19 pandemic, which disproportionately affected US Indigenous communities. (As an aside, there truly is a "digital divide" in many Indigenous communities, and those inequities to accessing the Internet were just exacerbated when tribal schools were forced to meet online. Many tribal households lack standard broadband at a rate four times as the general population, and 18 percent of tribal reservation residents lack Internet access at home [45]. It is sometimes difficult to push for data science education in

tribal colleges when the technology disparity is so apparent.) While I am no longer a TCU academic and am transitioning to a role at a more mainstream university, I still cooperate with TCU partners to ensure that data science education opportunities are Indigenous-driven and Indigenous-centered. One recent success is our IndigiData Indigenous Data Science Summer Workshop, which met for the first time in Summer 2021.

> BMT: I am also someone who has held leadership positions as a tribal college educator. I spent eight years teaching mathematics to a broad range of Indigneous students at a tribal college providing four year fine arts degrees serving Indigenous communities primarily from North America. Developing and coordinating mathematics and quantitative reasoning curricula in this context meant not only thinking about exploring ways that mathematics and quantitative reasoning could be made meaningful and accessible to students coming from a broad range of diverse cultures and experiences, but also exploring ways to make mathematics and quantitative reasoning meaningful and accessible to students in creative focused arts programs. This meant rethinking what counts as mathematics so that we build on notions of mathematics that are not meant to prepare students for calculus based courses in STEM, but for mathematics that supports and draws on the creative strengths of the students. One of the challenges that I faced as a TCU was navigating the tensions from the need to be responsive and adaptive to the creative and cultural strengths of our students and the need to respond to institutional pushes to be data driven in the assessment of our student learning. By being at an institution that did not offer calculus-based STEM programs, comparing student success in mathematics and the assessment of our mathematics learning outcomes with that of other non-TCU or non-STEM oriented programs gave a different type of tension, a need to validate the types of mathematics and brilliance that our students were engaged in. This often limited my capacity to feel in connection to not only typical 2-year and 4-year institutions, but also to other TCUs that did prepare students for calculus-based STEM tracks.

I am currently at an Hispanic Serving Institution community college. Still, much of my energy and focus has focused on the need to bring issues of justice, equity, diversity, and inclusion as both a Minority Serving Institution (MSI) and in response to the nationwide resurgence of the Black Lives Matter movement that pushed for more discourse and action from higher education institutions. There is an ongoing challenge drawing links between social justice discourses and learning outcomes assessment of college algebra at a two-year college. Many institutions have responded to calls for social justice, diversity, and inclusion by making statements in response to such events as the murder of George Floyd, anti-Asian sentiments, violence at the borders, and trends towards land acknowledgement statements. Though these statements can be interpreted as performative without action, I am still influenced by my work at a TCU and the missions of TCUs to be Indigenous-driven and Indigenous-centered . In other words, even at a non-TCU I am still feeling the same forces of assimilation in the sense that we still feel pressures to be accountable to structures that do not center Indigenous communities, that do not center our local communities, and are grappling to see how the Black Lives Matter movement also matters to our community college.

AHS: MSIs also provide interesting alternatives to data science education. The alternative that I am most familiar with comes out of a consortium of MSI's devoted to digital ethnic studies, and in particular out of work being done at California State University Fullerton. The Cal State Universities operate in tandem with California Community Colleges and the University of California system. I worked at Cal State Fullerton for two years, in a job focused on digital history. While I was there, and after I've left, that program has increasingly focused on data literacy and analysis. This effort has been driven by Jamila Moore-Pewu, who is building something that is not a "traditional" data science program, but which addresses some of the things we've talked about in this paper. It is not housed in a STEM discipline, and is not explicitly related to industry. Rather, it is the product of an

internal effort to help students build skills that will be helpful to them after college. A recent three million dollar grant from the Andrew R. Mellon foundation, to CSUF and other public MSIs, is funding efforts to "develop programming in digital ethnic studies, which uses data visualization to highlight the experiences of Black, Indigenous and People of Color." Something that might be called data science in other spaces or disciplines looks very different at CSUF. Rather than beginning with Tufte or Bayes, these classes begin with Ida B. Wells-Barnett — arguably the first public data scientist. In designing a program that centers Black, Indigenous and other communities of color, the program that Moore-Pewu is developing is not likely to look like other data science programs. That might be a good thing.

CDE: Bates College, as AHS mentions, is an institution of enormous privilege. This type of institution is not the kind in which I experienced my own education — my education and my siblings' education were all in completely public institutions. Despite applying to a variety of positions, earlier in my career, I was only offered full-time jobs at predominately white private colleges. That said, our institution is still a home to approximately 20% students who identify as domestic Latinx, Black, and/or Indigenous, 13% first-generation, and 10% Pell-eligible. I have a responsibility to make their experience as joyful as possible. In addition, I feel I have a responsibility to leverage resources in directions that support other institutions and settings. This is not always perceived well at small colleges to have a significant level of service aimed at national interventions, but it is important to me that I am using the privilege I have to do something important for my community as a whole. I also think that places like Bates should learn from what places like Morehouse and CSUF are doing - identities in positive ways that shape agency for students.

Dialogue 2 (Journey) summary.

As one might expect, we found that our journeys to building out new ways of thinking about data science in relation to ethics came from a set of diverse dreams, sometimes asking "toward what justice?" to build on the reference

presented by Tuck and Yang in their edited volume *Toward What Justice? Diverse Dreams of Justice in Education* [83]. This volume presents narratives and research-based perspectives on both the promise and the tensions of different communities' pathways toward understanding and realizing justice. Across these complex dreams — because dreams are, all at once, philosophical, historical, cultural, social, economic, material, and metaphysical in our embodiment of some realized futures — a set of specific themes surfaced in our mutual understanding and conversations. Namely, we notice that making a difference in a primarily academic context is a matter of attending to the historical structures that have taken hold of one's ability to not only dream but to implement said dreams. So a question of leadership comes to the fore; and then a struggle for authenticity as one becomes, for lack of better words, more "instituted."

We believe most readers can see where this story often leads, so our discussions lead toward the specific ways that we resist orientations of assimilation into the white supremacist structure, while also acknowledging the complex ways that our positions can be used to generate new and more equitable systems. In this way, one main challenge in our discussions focused on understanding how power manifests in our local context. Elsewhere, we noticed an important thread of critiques in relation to the content and normal practices that have been historically instituted. As can be imagined, the manifestation of power functions in the development of certain kinds of content and the privilege of certain practices. Diverging from these practices, we discussed, becomes another threat to being true to ourselves, while taking into consideration our need for survival, for work and pay. To deal with the many and complex components that we identified, we talked much about our lived experiences, that of our ancestors, and how we have and continue to use this knowledge to make sense of and act upon shifts in systems, and future ways of being as we build practices with our students.

DIALOGUE 3: Ethics

What does ethics mean for our data science programs? How are we going beyond ethics in our data science education programs?

KST: We must be careful that we as instructors and practitioners of data science are cognizant of any implicit biases that we impart onto our stu-

dents and mentees. For instance, much of "data science ethics" is taught from a very Western Eurocentric perspective which views justice and equity guiding principles from the perspective of data users, which are likely to be researchers and not the individuals and communities who initially provided the data. Recent NIH grants evaluate ethics based on FAIR principles of data governance (findability, accessibility, interoperability, and reusability) [87], but these center open data sharing for the benefit of researchers without acknowledging the direct risks and impacts onto underrepresented and minoritized populations who are often the subject of calls for increased inclusion for diverse datasets. We must remember that data has value and data is never devoid of some form of risk, and it is a fallacy to assume that anonymizing datasets, particularly of biological data which bears group re-identification risks, removes risks from individuals and communities. Data science education programs that are led by members of our representative communities best impart knowledge of these differences in ethics to the next generation of leaders. Scholars of color often feel the disproportionate burden of educating others within their institutions, and these additional labors are often uncompensated or added expectations differentially felt compared to white colleagues [74]. However, there is a reaffirming feeling when we feel like we are contributing to the training of our own peoples.

> AHS: This made me think of the CARE principles (Collective Benefit, Authority to Control, Responsibility and Ethics) [20, 21].
>
> NNA: I'd like to hear more about the CARE principles and how, in considering how scholars of color, in relation to KST's points, might begin to frame and push against this troubling focus on ethics which supports the maintenance of Western Eurocentric ways of being. One way that we are attempting to move beyond ethics is to center our collective histories and the work of Black scholars and educators within and beyond the borders of the U.S. Still, this work falls short in some ways because there are so many intersectional points of consideration that need to be situated in developing new ways to contribute.
>
> AHS: I came to CARE from the GIDA group - https://www.gida-global.org/care. They've wanted to build on the framework of FAIR because "Existing principles within the open data

> movement (e.g. FAIR: findable, accessible, interoperable, reusable) primarily focus on characteristics of data that will facilitate increased data sharing among entities while ignoring power differentials and historical contexts."

AHS: Ethics is not a neutral concept. The philosophical subfield of ethics is traditionally seen to stem from Aristotle, and was incorporated into theories of "natural law" that were developed by European and Euro-American philosophers during the enlightenment. These same men (and they were mostly men) also developed moral theories that condemned unfreedom (for example, restrictions on the speech of activities of propertied European men) while still making space for racialized and heritable slavery. All of this is to say that ethics carries a lot of racist and colonialist baggage. I get nervous about the uncritical application of "ethics" to data science, without acknowledging that what is considered fair, just, or right is highly culturally informed, and without making space for formulations of justice that are not descended from Aristotlian or enlightenment-era ethics. We should have data science that works towards justice. We should have data science that is equitably accessible. We should not have data science that does harm. But we should also not uncritically adopt ethics as a framework without being very clear what we mean by it.

> KST: Yes, I agree. I also state later that justice needs to be defined by communities themselves. I want to add that "equity" is not the same as "equality." "Data democratization" is a current and popular buzzphrase used by data scientists, but I think the phrase warrants some critical unpacking lest it be misused. First of all, "democracy" can be a Western, individualistic system, and the implied American system of democracy should not be prioritized over Indigenous democracies, such as the Haudenosaunee Confederacy, which may have different sets of ethical principles. Secondly, any system in which "majority" rules will, ultimately and fundamentally, disenfranchise small, underrepresented communities.

CDE: One thing that DCS is doing is decoupling ethics from racism, sexism, ableism and the other -isms that are the foundation of and plague of STEM.

We also treat these conversations as necessary to have throughout multiple courses, not just in a siloed course. As a math educator, one of the things we used to tell programs about quantitative literacy is that you can't expect one math class to make students quantitatively literate — it needs to be also woven through the curriculum. So why should we expect conversations about critical race studies to be any different?

Furthermore, why is the end goal ethics? Data justice and algorithmic justice movements such as Data 4 Black Lives and the Algorithmic Justice League call us to go beyond simply data ethics [28]. We are now seeing movements beyond open data or open science calling for a movement to information justice [48]; instead of merely inclusive teaching and open education, we might strive towards educational justice [31]. Here open becomes a tool for social justice communities, not the goal of social justice. It then allows us to understand "Open for whom?": when open is a tool that should be wielded and when open adds to further oppression and marginalization.

> AHS: I like the emphasis on justice — and I don't want to be *that* humanist, asking "but what do we mean by ..." into infinity, but I think (referencing what KST wrote above) that being as specific as possible about who is defining justice, ethics, democracy is important. I gather that ethics in CS (for example) had the potential to be important and novel, but that "ethics" is in the process of being defined into nothing.

NNA: In my current context, ethics has not been a central part of our discussions around data science. However, we often think about the role of data in helping to eradicate inequities. So, as we think about what equity in data science means in our local contexts, some of our questions could surround, for example, historical examples of how data has been used to expose inequity and injustice, and how those inequalities have, for a large part, continued as a result of a lack of ethics from stories that come from the data. From a more intersectional lens, we could consider gender and class, language and global considerations that relate to the Black diaspora; even more specifically, the historical role of data in maintaining systemic inequities rooted in anti-blackness. Outside of our specific context, I also think about data as a way to tell a story and provide opportunities to expand our thinking about

what data is, how data can be leveraged for justice, and what futures can be conceptualized based on our knowledge of and use of data.

BMT: A concern of mine would be developing a data science program that only seeks to prepare students for a growing job market. That is, my concern would be getting into the trap of figuring out the nuts and bolts of a program and not prioritizing critical perspectives about what data science is, who it is for, and questions of ownership. In developing a data science program with the missions of TCUs and community colleges in mind, how can we prioritize community involvement? Maybe not just involvement, but also thinking about ownership of and access to data itself.

> AHS: I've also been thinking about Mary Poovey's work (she also has some good histories of the development of statistics as a tool for state control) on white women in medicine, and how, as soon as white women began to take on professional medical roles, the professional identity that they inhabited, nursing, was deprofessionalized. I think that as we talk about job preparation, we also need to think about the ways in which structures are going to gate keep the boundaries of data science, so that it continues to be capital-enchancing for some (white) students. What happens when it becomes capital-enhancing for marginalized students?
>
> CDE: Probably it will become less attractive/pay less. So how do we re-engineer that paradigm in anticipation of that moment that is currently a marker for educational justice? Real justice will be a moving target. But also, I will add to what Belin said — that there is a capitalistic reason for workforce development. So when we take a step back, we ask why the white power that holds capitalistic power now cares about us. And then we are reminded that this is a form of racial commodification — that our bodies are now useful in a tech race that the US is losing [10].

Dialogue 3 (Ethics) summary.

What we learned is that even tying "ethical" practices to guide conversations related to data science and education potentially evokes implicit biases related to the values stemming from which systems of ethics we prioritize and

de-center. Not paying due attention to these biases unfortunately can substantiate the dominance of one system of ethics over another, which can reinforce and serve as a microcosm of similar inequities and power imbalances that undergird data science and education currently. Some key questions that arose from our discussions include "who are we centering" and whether we are contributing to the marginalization of underrepresented voices and peoples when we frame data as being "neutral" and data science as perpetuating "objective research" [33]. These latter points, in particular, shifted our conversation from globally unpacking data science ethics to an operational discussion about justice, namely "for whom does data and data science benefit." Since some of us identify with coming from historically disenfranchised groups or have experience in working with MSIs, a common theme from our responses highlighted the importance of positioning scholars of color and their communities as being agents of determining justice- and equity-centered means for moving the field of data science and education forward. For those who are not from these communities, we discussed briefly the importance of what allyship entails, and the importance of listening before trying to act on behalf of communities. Ultimately, what came out of our conversation was a move towards not just decolonial but anti-colonial perspectives of data and informational justice which prioritizes communities as agents for stewarding their data, not researchers and outsiders. Moving towards justice requires an internal reflection within the fields of data science to understand Eurocentricity (or whiteness axiomatically), and to reframe ethics as more than just opposition to"isms" such as sexism, ableism, and racism. Data science should not be content to merely avoid sexism, ableism and racism. Seeking justice for data science means actively and materially working against the power structures that support these "isms." Hopefully these critiques will lend itself towards a positive outlook and a move towards actual data equity for all in the future.

DIALOGUE 4: Culture

How do themes of the past, present and future of data science and/or our communities play into our work?

CDE: I teach a course called the "Past, Present, and Possible Dystopian Future of Computing." At first, I was thinking about it as a standard

stand alone "computing ethics" course, but its focus became more focused on marginalization in STEM. It's a film-based course, and it satisfies the writing intensive requirement. Both of these design parameters make it ideal for holding conversations about racism, sexism, and bigotry in STEM. We watch film together and discuss at intervals, so there is a common viewing experience at the same pace. As a writing intensive class, students use writing to further their understanding and foster reflection. I use writing both as a tool to learn as well as to help students craft writing in their discipline. The choices of movies range from the familiar blockbuster like Hidden Figures to the lesser known Sleep Dealer and range in how computing technology is developed and used. The lesser known film Sleep Dealer strongly resonates with my community identity, as it explores border labor with its dystopian future in which Mexicans "plug into" the border wall to perform labor for the US.

I also teach *Calling Bull* (`(https://qubeshub.org/community/groups/callingbull/)`, a course that focuses on data and scientific literacy as it plays out in contemporary news, social media, and among their family and friends. But it is also a digital advocacy course which asks students to imagine and advocate for the future they want to see. The past is where we understand that marginalization as a pattern throughout history once we understand the structures, we use it as a tool to examine the contemporary. Only when we understand how it persists can we reimagine a new future and bring that future to fruition. The students I teach become tech industry leaders partly by virtue of their own work, but also because we are an elite private school with an active job placement network - so we have a responsibility to send leaders into the future that can imagine the consequences of their technological advancements and design against it if necessary.

NNA: I have been thinking a lot about the practical aspects of data science education in undergraduate settings.. As I apply some of my thinking regarding equity, justice and futurity, and specifically Afrofuturism, my starting point is often with abolitionists such as Ida B. Wells-Barnett, and her work using data on lynching, and W.E.B. DuBois and his work examining African American communities. Their use of data to disrupt deficit narratives and expand new ways of thinking have positioned possibilities for students to explore historical data and consider what will be needed for the future to

help continue these legacies. In this way, data science education becomes a part of a longer lineage of cultural practices as opposed to a static set of literacies across the various domains that inform data science. I imagine that this begins to set the foundation for the future possibilities of the type of educational paradigms that we want to move toward, as we figure out ways, methods, and practices that allow us to shift away from the current status quo for college classrooms (i.e., banking model of education, as noted by Freire).

BMT: In my work, I've been resistant to the "overquantification" of things [42]. For example, I've been influenced by Rochelle Gutierréz's work [42, 44] that has been critical of "gap-gazing" as a central notion in discussing equity in mathematics education. I'm also mindful of the ongoing conversations in Indigenous communities regarding the quantifications of Indigenous identity and membership through blood quantum and the continued harmful impacts that has had on our communities. In my teaching, I've been trending towards "ungrading" [14] in my classes in resistance to reducing the complexities of learning and assessment to points, percentages, and letter grades. So coming into conversations on data science, I come from places where I am constantly thinking about whether it's appropriate for institutions/people in power to have conversations that use data without centralizing the perspectives of the communities that they are talking about. So the future of data science can be transformative if community voices and ownership is prioritized. The future of data science education can also be damaging if the central focus downplays the needs of communities from justice-oriented frameworks.

AHS: I cannot approach data science without the context of the past harms that data has facilitated. It is impossible, I think, to use this data, or any data produced to do harm (and a lot of data is produced to do harm or enact power) without grappling, both theoretically and methodologically with the ways in which that harm and power infects our data.

KST: My own operability in data science is informed by my experiences as an Indigenous academic and advocate in population genetics. If we think back even two decades ago, there was a huge push for "increased diversity and inclusion of underrepresented people" in genomic datasets, but this was under the explicit aim of recruitment. Power dynamics were unbalanced, with a very hierarchical, top-down relationship between researcher and communities

with the underlying drive of collecting as much data from Indigenous groups before they "vanished" [80]. Indigenous groups then were concerned about the open accessibility of their biological data and concerned about the potential for biocommercial exploitation by companies using their Indigenous DNA for patented technologies that primarily benefit companies' stock prices. History is cyclical, and now 20 years later, we still see many of the same concerns persist today in the form of 1) concerns about open data, 2) data privacy, group re-identifiability and discrimination, and 3) usurpment of Indigenous data sovereignties for the development of commercial intellectual domains.

Dialogue 4 (Culture) summary.

When we thought about the past, present and future hopes for data science in and with our own communities, many of us focused on the past, and on harm. We understood this harm and usability both structurally, and in terms of the specific ways in which our own communities are impacted. These specificites, too, reveal the structures. This harm is felt most deeply in communities historically and contemporaneously marginalized in STEM. It includes overquantification, which takes the form of white and settler researchers weaponizing data and numbers to police Indigenous communities, and to fetishize assessments that "identify disparities between middle-class [w]hite students and students who are Black, Latina/Latino, First Nations, English language learners or working class." [42, page 357]. This harm also includes white and settler researchers' exploitation, commercialization and monetization of data about Indigenous communities without the consent of those communities. As we described these harms, and called for attention to the ways in which data science has exacerbated them in the past, we also noted that data science's past could be usable, hopeful, and maybe even liberatory. We described how understanding this can help to re-write the origin story of a more critical data science, highlighting, for example, the foundational work of Ida B. Wells-Barnett and W. E. B. DuBois. This potential for historical reframing dovetailed with our dreams for the future of the field. When we posit that the roots of modern data science began with others, such as Wells-Barnett and Du Bois, rather than with eugenicists and traditional statisticians, we argued, it is possible to "disrupt deficit narratives," "expand new ways of thinking ," "reimagine a new future and bring that future work to fruition." These possibilities reflect not a single ethical framework, but

rather a set of hopes and dreams that collectively help to make data science "how it ought to be," rather than how it has been, or is.

DIALOGUE 5: Accountability

What are some things we are holding true to as we build out these programs and do this work? How are we keeping ourselves and/or our communities accountable?

CDE: In everything I do, I try to develop a community of support. These are communities of people working together towards shared visions and honest conversations. It follows the paradigms of both communities of transformation [51] and collective action theories as we make change together. An important feature is how the spaces we create become spaces of challenge to motivate ourselves to reach for something more. My peers in these communities of transformation hold me accountable in this way. A second thing I am thinking about a lot is how we SAY we want justice, but we don't commit to it in writing or with our resources. So how do we get people to formalize their commitments into policies and action? In the classroom and in professional development, one way I try to get at this is through helping people identify and build their agency to act. We've also worked on developing data-informed insights about our classrooms. Are our Black, Latinx and Indigenous students thriving in our classes and continuing to join us in future experiences? Is our community developing cultural competencies in computing? The data show that we still have work to do.

BMT: I come from mathematical experiences that are not rooted in data science, yet because numbers and calculations are involved, I am asked to provide expertise. I've noticed a tendency for professional mathematicians to take up space in discourses outside of their expertise just because computations are involved. I'm thinking about the need for mathematicians in mathematics departments getting involved in developing data science programs to step back and learn from the data science educators who have been doing the work and thinking about establishing respectful relationships with communities.

NNA: I find myself struggling with developing truly interdisciplinary courses that challenge what is considered mathematics. In practice this means con-

sidering when and where to emphasize real-world applications in relation to their abstractions in mathematical thinking. These changes also require accountability with course learning outcomes, and how some courses may be more easily developed and institutionalized in a rapidly developing data science ecosystem.

AHS: The unofficial mantra of DCS might be "who are we centering here?" It is easy, at a predominately white institution with an (as of July 2021) entirely white senior administration, to center the articulated needs of students who come into our classrooms with considerable social privilege, and to center the articulated needs of our college administration and donors. These groups hold us accountable to their needs whether we seek out their opinions or not. Funnily (not funny) these needs seem to push us towards educational structures that we know — from experience and from the literature — are exclusionary. As we build the program, we're working to remember that the loudest (or best-funded) voices in the room are not the only, or the most important voices.

> CDE: These needs are like a strong wind blowing us in the wrong direction — or currents that threaten to run us over a waterfall. That's how I am picturing this.

KST: There is, for whatever reason, a stigmatization against scholars of color also serving as advocates for their own communities. For instance, a top-tier peer-reviewed biomedical journal declined a commentary from Indigenous scholars because of a misguided notion that we could not be "objective" about scientific concerns that directly relate to our own people. Instead, they pushed for a perspective that, again, predominantly centered white scholars on matters that concerns our communities. This fallacy that data science is "objective" also reinforces Indigenous peoples as "objects" of research [53]. The notion that data science should be "dispassionate" needs to be abandoned. Data has value and data is inextricably linked with power, therefore data decisions must be driven by communities. The last couple of years of racial unrest (that has always existed but only recently became prominently featured in the news) has demonstrated that data science needs to be equitable and justice-centered. Those principles need to be defined by communities themselves. We as scholars of color are accountable to our communities. But to whom are white scholars accountable?

White scholars outnumber scholars of color, so we do appreciate and rely on white allyship to amplify our voices. Unfortunately, white amplification can also dangerously approach white co-optation of ideas that originate from scholars and communities of color. Sometimes, advocacy entails sitting down, letting us speak for ourselves, and truly listening to what we have to say.

> AHS: I keep coming back to the fact that objectivity is as much of a construct as is ethics. There is no God-View [33] that does not also imply relative privilege. These aren't fully fleshed thoughts — but what does it mean to make a data science program that *isn't* predicated on objectivity? We have a colleague at Bates who has started to begin her stats classes with a discussion of eugenics — and it changes the whole conversation about statistical objectivity. Referencing NNA above — such a program would, I think, need to focus on what data is used for (for good and evil) rather than (referencing BMT) checking "professionalizing" boxes (yet another construct).
>
> CDE: Sociologist Victor Ray edits a blog called Conditionally Accepted and has a great article called "The Unbearable Whiteness of Mesearch" [73] that talks about how scholarship of our communities, and the application and collaboration of that scholarly work in and with our communities is brushed off as simply service. I worry that we too easily see research which prioritizes our communities and tells our stories as "Mesearch." The academy is too willing to accept that the teaching which prioritizes our communities and tells our stories should be simply optional. To view social justice as optional is to insist your privilege afforded by white supremacy should be centered. How should you be held accountable for that open proclamation?

Dialogue 5 (Accountability) summary

Together as we think about accountability, we realize that our answers may depend on positionality. Systems are constructed in which we are and have been accountable to white supremacy. As scholars from various communities historically and contemporaneously marginalized in STEM and now in data science, we ask - when will data science be accountable to our communities?

One pathway is by empowering those in our communities to become data scientists, because of the deep understanding brought and because of the commitment to accountability. However, at the same time we say we need to take more space in data science, we simultaneously critique ourselves by saying that data science is not the only way to make a difference in our communities and no future is entitled to the continued commodification of our bodies. If we center race and ethnicity in our discussion of marginalization, then white people need to get out of the way and support the ways in which we are constructing our spaces orienting towards justice. The words of Joy Buolamwini [19], "Who codes matters, How we code matters, and Why we code matters." help us remember — who are we centering in this conversation about data science ethics and futures? Whose knowledge is centered? If we really centered marginalized communities, if were accountable to these communities, if we asked ourselves "Who is doing data science, How we are doing data science, and Why we are doing data science," and if we are centered on human rights and the human experience, then, would we even need to invoke ethics as a framework for building our future? Maybe we are asking more questions than providing answers, but the room for possibilities adds to the excitement of building our future together.

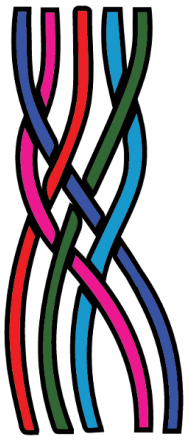

Figure 1: A regular braid of five strands, modified by AHS from "A regular braid of five strands" (2013) created by Stifehler, available at https://commons.wikimedia.org/wiki/File:5_Strand_Braiding_Technique.png using Creative Commons License BY-SA 3.0.

Intersections

Instead of a traditional conclusion, we offer reflections on intersections found in our dialogues. Our positionalities, our journeys to this place, our contemplations on ethics and social justice, our dreams for the future, and the accountability we bear to our communities are all intertwined. Each of these five strands are simultaneously present in our reflections. Just as our five lives and stories are intertwining together in this moment, in this paper, to weave in a larger tapestry to show you how we see data science education. We hope that these dialogues leave with you the message that ethics or better yet, encouraging respect for human dignity, in data science is not about one fixed idea or approach, but should embrace the multiple approaches, epistemologies and worldviews that we are portend to serve. Instead of leading this conversation predicated on the thoughts reflected by those from a dominant social construct, we should center the practitioners of data science from groups that have been historically and contemporaneously marginalized.

As we discussed the strands of our reflections, we thought about why the place of our education work matters to us and shapes so much of our journeys instead of our positionalities. In the course of our discussions, we realized one response is that our students are taking the journey with us, not just our fellow practitioners, educators, and administrators. Our students are also an unheard voice in these discussions, and we find connections with students who remind us of ourselves, trying to navigate this system of education. In our reflections of the past and present, we grounded ourselves in harm. Why did we do this, we asked ourselves? Because this is the experience we try to shield our students from — by playing defense and offense at the front lines of classrooms, administrative, institutional, and national conversations. Our future, our students' future, our future students' experience, our communities will all depend on these battles.

But this work is difficult and draining. We ask ourselves if we should leave, but we bear this sense of responsibility and accountability to advocate for our communities in spaces. Some of us have left. We've moved institutions to try to find our place of belonging where we can contribute our strengths and hope for shared dreams. We've moved to institutions with "students like us" or to institutions that told us we could build something new in the

age of data science. But these dreams we dream feel like only dreams when we are faced with the harsh reality of the landscape which still expects us to assimilate. Like being asked to weave a tapestry with muted colors and coarse yarn. The difference is — are our institutions, our colleagues, our students giving us the space, the time, the support, the encouragement to keep us going or are they too tied to the strings of white supremacy that they do not know how to dream the dreams we see?

It is paradoxical that the liberatory promise of education is tied to the way that settler colonialism manifested in the educational state in the US as a tool of assimilation. These issues exist across different contexts, including minority serving institutions. As we dive into the root of our current struggle reimagining data science, every problem that we've named has been about fighting the way data science is continuing to be constructed within a white supremacist framework. We have to go beyond ethics. These issues are deeply rooted in our society, in STEM education, and also deeply rooted within ourselves. It has been critically important to give ourselves the time and space to self-reflect, instead of placing the blame solely on others. Sadly, what seems to bring us together is this common battle against assimilation and within ourselves — the confrontation of the tools of assimilation we have wielded to get here and the desire to do better. But in this space, where we intersect, is this moment of solidarity, harmony, and respect, that encourages us to keep dreaming and fighting for our dreams.

References

[1] Association of Computing Machines (ACM) and Institute Electrical & Electronic Engineers (IEEE), *Computer Science Curricula 2013: Curriculum Guidelines for Undergraduate Degree Programs in Computer Science*, ACM, New York, 2013.

[2] American Statistical Associaition (ASA), *Curriculum Guidelines for Undergraduate Programs in Statistical Science*, ASA, Alexandria, 2014.

[3] Adiredja, A. P., & Andrews-Larson, C., "Taking the sociopolitical turn in postsecondary mathematics education research," *International Journal of Research in Undergraduate Mathematics Education*, Volume **3** Issue 3 (2017), pages 444–465.

[4] Alexander, N. N., Teymuroglu, Z., & Yerger, C. R., "Critical conversations on social justice in undergraduate mathematics," *Problems Resources and Issues in Mathematics Undergraduate Studies*, Volume **29** Issues 3–4 (2019), pages 396–419.

[5] Archibald, J., Lee-Morgan, J., & De Santolo, J. (editors), *Decolonizing Research: Indigenous Storywork as Methodology*, Bloomsbury, London, 2019.

[6] Associated Press (AP), "Explaining AP Style on Black and white," *AP News*, July 2020. Available at https://apnews.com/article/archive-race-and-ethnicity-9105661462, last accessed on July 30, 2022.

[7] Bakhtin, M.M., *The dialogic imagination: Four essays*. University of Texas Press, Austin, 1981.

[8] Barocas, S., & Boyd, D., "Engaging the ethics of data science in practice," *Communications of the ACM*, Volume **60** Issue 11 (2017), pages 23–25.

[9] Bargagliotti, A., Franklin, C., Arnold, P., Gould, R., Johnson, S., Perez, L., & Spangler, D., *Pre-K-12 Guidelines for Assessment and Instruction in Statistics Education (GAISE) report II*, American Statistical Association and National Council of Teachers of Mathematics, 2020.

[10] Basile, V., & Lopez, E., "And still I see no changes: Enduring views of students of color in science and mathematics education policy reports," *Science Education*, Volume **99** Issue 3 (2015), pages 519–548.

[11] Baumer, B. S., Garcia, R. L., Kim, A. Y., Kinnaird, K. M., & Ott, M, Q., "Integrating data science ethics into an undergraduate major: A case study, *Journal of Statistics and Data Science Education*, Volume **30** Issue 1 (2022), pages 15–28.

[12] Benjamin, R., *Race After Technology: Abolitionist Tools for the New Jim Code*, Polity, Cambridge, 2019.

[13] Blikstein, P., "Travels in Troy with Freire: Technology as an agent of emancipation," pages 205–244 in *Social justice education for teachers: Paulo Freire and the possible dream*, edited by P. Noguera and C. A. Torres (Sense, Rotterdam, 2008).

[14] Blum, S. D., & Kohn, A., *Ungrading: Why rating students undermines learning (and what to do instead)*, West Virginia University Press, 2020.

[15] brown, a. m., *Emergent Strategy: Shaping Change, Changing Worlds*, AK Press, New York, 2017.

[16] Breault, R. A., "Emerging issues in duoethnography," *International Journal of Qualitative Studies in Education*, Volume **29** Issue 6 (2016), pages 777–794.

[17] Buckmire, R., "'Who Does the Math?': On the Diversity and Demographics of the Mathematics Community in the USA," pages 1–12 in *Improving Applied Mathematics Education* edited by R. Buckmire and J. Libertini (Springer, New York, 2021).

[18] Bullock, B. B., Fernando L. Nascimento, F. L, & Stacy A. Doore, S. A., "Computing Ethics Narratives: Teaching Computing Ethics and the Impact of Predictive Algorithms," pages 1020–1026 in *Proceedings of the 52nd ACM Technical Symposium on Computer Science Education (SIGCSE '21* (Association for Computing Machinery, New York, 2021).

[19] Buolamwini, J., *How I'm fighting bias in algorithms*, TED Talks: TEDx Beacon Street, November 2016. Recording available at https://www.ted.com/talks/joy_buolamwini_how_i_m_fighting_bias_in_algorithms, last accessed on July 30, 2022.

[20] Carroll, S., Garba, I., Figueroa-Rodríguez, O., Holbrook, J., *et al.*, "The CARE Principles for Indigenous Data Governance," *Data Science Journal*, Volume **19** (2020), Article 43. doi:10.5334/dsj-2020-043

[21] Carroll, S.R., Herczog, E., Hudson, M., Russell, K., & Stall, K., "Operationalizing the CARE and FAIR Principles for Indigenous data futures," *Scientific Data*, Volume **8** (2021), Article 108. doi:10.1038/s41597-021-00892-0

[22] Carver, R. *et al.*, *Guidelines for Assessment and Instruction in Statistics Education (GAISE) college report 2016*, American Statistical Association, 2016.

[23] Chicago Manual Style, "Black and White: A Matter of Capitalization," *CMOS Shop Talk*, June 22, 2020. Available at https://cmosshoptalk.

com/2020/06/22/black-and-white-a-matter-of-capitalization/, last accessed on July 30, 2022.

[24] Committee on Professional Ethics, *Ethical guidelines for statistical practice*, Technical report, American Statistical Association, 2018. Available at http://www.amstat.org/asa/files/pdfs/EthicalGuidelines.pdf, last accessed on July 30, 2022.

[25] Corbett, C., & Hill, C., *Solving the Equation: The Variables for Women's Success in Engineering and Computing*, American Association of University Women, Washington, DC, 2015.

[26] Crall, A., Diaz Eaton, C., Gram, W., Gupta, S., Lauer, A., Shanley, L., Stone, S. A., Swan, B., Sudduth, E., Whitney, K. S., Yong, K., *2019 EDSIN Conference Proceedings*, EDSIN: Environmental Data Science Inclusion Network, QUBES Educational Resources, 2021. doi:10.25334/9C03-1518

[27] Delgado, R., "Storytelling for oppositionists and others: A plea for narrative," *Michigan Law Review*, Volume **87** (1989), pages 2411–2441.

[28] Dencik, L., Hintz, A., Redden, J., & Treré, E., "Exploring data justice: Conceptions, applications and directions," *Information, Communication & Society*, Volume **22** Issue 7 (2019), pages 873–881. doi:10.1080/1369118X.2019.1606268

[29] Denzin, N. K., Lincoln, Y. S., & Smith, L. T. (editors), *Handbook of critical and indigenous methodologies*, Sage, Thousand Oaks, 2008.

[30] De Veaux, R., R. D., Agarwal, M., Averett, M., Baumer, B. S., Bray, A., Bressoud, T. C. *et al.*, "Curriculum Guidelines for Undergraduate Programs in Data Science," *Annual Review of Statistics and Its Application*, Volume **4** (2017), pages 15–30.

[31] Diaz Eaton C., Bonner K., Cangialosi K., Dewsbury B., Diamond-Stanic M., Douma J., Smith M. K., Taylor R., Wilfong K., & Wojdak, J.., "Sustainability Challenges for Open Resources to promote an Equitable Undergraduate Biology Education," to appear in *CBE-Life Science Education*.

[32] Diaz Eaton, C., "No, Not yet." blog post at *Math Mamas*: American Mathematical Society Blog, June 6, 2019. Available at https://blogs.ams.org/mathmamas/2019/06/17/no-not-yet/, last accessed on July 30, 2022.

[33] D'ignazio, C., & Klein, L. F., *Data Feminism*, MIT Press, 2020.

[34] Dixon-Román, E., & Parisi, L., "Data capitalism and the counter futures of ethics in artificial intelligence," *Communication and the Public*, Volume **5** Issue 3-4 (2020), pages 116–121.

[35] Elliot, A, C., Stokes, S. L., & Cao, J., "Teaching ethics in a statistics curriculum with a cross-cultural emphasis," *The American Statistician*, Volume **72** Issue 4 (2018), pages 359–367.

[36] Fiesler, C., Garrett, N., & Beard, N., "What do we teach when we teach tech ethics? A syllabi analysis," pages 289–295 in *Proceedings of the 51st ACM Technical Symposium on Computing Science Education* (2020). doi:10.1145/3328778.3366825

[37] Floridi, L., & Taddeo, M., "What is data ethics?' *Philosophical Transactions of the Royal Society A*, Volume **374** (2016), 20160360. doi:10.1098/rsta.2016.0360

[38] Gay, G., *Culturally responsive teaching: Theory, research, and practice*, Teachers College Press, New York, 2010.

[39] Gholson, M. L., & Wilkes, C. E., "(Mis)Taken Identities: Reclaiming Identities of the 'Collective Black' in Mathematics Education Research Through an Exercise in Black Specificity," *Review of Research in Education*, Volume **41** Issue 1 (2017), pages 228–252.

[40] Goffney, I., Gutiérez, R., & Boston, M., *Rehumanizing Mathematics for Black, Indigenous, and Latinx Students*. National Council of Teachers of Mathematics, 2018.

[41] Grande, S., *Red pedagogy: Native American social and political thought*, Rowman & Littlefield, 2015.

[42] Gutiérrez, R., "A gap-gazing fetish in mathematics education? Problematizing research on the achievement gap," *Journal for Research in Mathematics Education*, Volume **39** (2008), pages 357–364.

[43] Gutiérrez, R., "Introduction: The need to rehumanize mathematics," pages 1–10 in *Rehumanizing mathematics for Black, Indigenous, and Latinx students*, edited by I. Goffney and R. Gutiérrez (National Council of Teachers of Mathematics (NCTM), 2018).

[44] Gutiérrez, R., & Dixon-Román, E., "Beyond gap gazing: How can thinking about education comprehensively help us (re) envision mathematics education?" pages 21–34 in *Mapping equity and quality in mathematics education* (Springer, Dordrecht, 2010).

[45] Howard, B,, & Morris, T., "Tribal technology assessment: The state of internet service on tribal lands," in *TPRC47: The 47th Research Conference on Communication, Information and Internet Policy*, 2019. doi:10.2139/ssrn.3427547

[46] Irizarry, R., "The role of academia in data science education," *Harvard Data Science Review*, Volume **2** Issue 1 (January 31, 2020).

[47] Iseke, J., "Indigenous storytelling as research," *International Review of Qualitative Research*, Volume **6** Issue 4 (2013), pages 559–577.

[48] Johnson, J. A., "Open data, big data, and just data," pages 23–49 in *Toward Information Justice: Public Administration and Information Technology* (Volume 33, Springer, Cham, 2018).

[49] Kantayya, S., & Buolamwini, J., *Coded bias*, 7th Empire Media, 2020. http://www.codedbias.com/, last accessed on July 30, 2022.

[50] Kennedy, B., & Moore, H., "Grounded Duoethnography: A Dialogic Method for the Exploration of Intuition Through Divergence and Convergence," *Forum Qualitative Sozialforschung / Forum: Qualitative Social Research*, Volume **22** Issue 2 (2021).

[51] Kezar, A., Gehrke, S., & Bernstein-Sierra, S., "Communities of transformation: Creating changes to deeply entrenched issues," *The Journal of Higher Education*, Volume **89** Issue 6 (2018), pages 832–864.

[52] Kokka, K., "Healing-informed social justice mathematics: Promoting students' sociopolitical consciousness and well-being in mathematics class," *Urban Education*, Volume **54** Issue 9 (2019), pages 1179–1209.

[53] Kukutai, T. & Walter, M., "Indigenous statistics," *Handbook of Research Methods in Health Social Sciences* edited by P. Liamputtong (Springer, Singapore, 2017). doi:10.1007/978-981-10-2779-6_40-1

[54] Ladson-Billings, G., "Toward a theory of culturally relevant pedagogy," *American Educational Research Journal*, Volume **32** Issue 3 (1995), pages 465–491.

[55] LaMar, T., & Boaler, J., "The importance and emergence of K-12 data science," *Phi Delta Kappan*, Volume **103** Issue 1 (2021), pages 49–53.

[56] Lee, V. R., Wilkerson, M. H., & Lanouette, K., "A Call for a Humanistic Stance Toward K–12 Data Science Education," *Educational Researcher*, Volume **50** Issue 9 (2021), pages 664–672.

[57] Madkins, T. C., & McKinney de Royston, M., "Illuminating political clarity in culturally relevant science instruction," *Science Education*, Volume **103** Issue 6 (2019), pages 1319–1346.

[58] Martin, D. B., "Race, racial projects, and mathematics education," *Journal for Research in Mathematics Education*, Volume **44** Issue 1 (2013), pages 316–333.

[59] Martin, D. B., *Mathematics success and failure among African-American youth: The roles of sociohistorical context, community forces, school influence, and individual agency*, Routledge, 2000.

[60] Mathematical Association of America (MAA), *2015 CUPM Curriculum Guide to Majors in the Mathematical Sciences*, MAA, Washington, DC, 2015. Available at https://www.maa.org/sites/default/files/pdf/CUPM/pdf/CUPMguide_print.pdf, last accessed on July 30, 2022.

[61] Mensah, F. M., "Retrospective accounts in the formation of an agenda for diversity, equity, & social justice for science education," pages 317–335 in *Moving the equity agenda forward* (Springer, Dordrecht, 2013).

[62] McKinsey Global Institute, *Big Data: The Next Frontier for Innovation, Competition, and Productivity*, McKinsey & Co., New York, 2011.

[63] Moll, L. C., & Gonzalez, N., "Lessons from research with language-minority children," *Journal of Reading Behavior*, Volume **26** Issue 4 (1994), pages 439–456.

[64] National Academies of Sciences, Engineering, and Medicine (NASEM), *Data Science for Undergraduates: Opportunities and Options*, National Academies Press, Washington, DC, 2018.

[65] National Association of Black Journalists, *NABJ Statement on Capitalizing Black and Other Racial Identifiers: June 2020*, NABJ Style Guide, web resource at NABJ Website, at https://www.nabj.org/page/styleguide, last accessed on July 31, 2022.

[66] Norris, J., "Duoethnography," pages 233–236 in *The SAGE encyclopedia of qualitative research methods*, Volume 1, edited by L. M. Given (Sage, Los Angeles, 2008).

[67] Norris, J., Sawyer, R. D., & Lund, D. E. (editors), *Duoethnography: Dialogic methods for social, health, and educational research*, Left Coast Books, Walnut Creek, 2012.

[68] Ober, R., "Kapati time: Storytelling as a data collection method in Indigenous research," *Learning Communities: International Journal of Learning in Social Contexts* [Special Issue: Decolonising Research Practices], Volume **22** (2017), pages 8–15. doi:10.18793/LCJ2017.22.02

[69] Oliver, J. C., & McNeil, T., "Undergraduate data science degrees emphasize computer science and statistics but fall short in ethics training and domain-specific context," *PeerJ Computer Science*, Volume **7** (2021), e441.

[70] O'Neil, C., *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy*, Crown, New York, 2016.

[71] Paris, D., "Culturally sustaining pedagogy: A needed change in stance, terminology, and practice," *Educational Researcher*, Volume **41** Issue 3 (2012), pages 93–97.

[72] Philip, T. M., & Olivares-Pasillas, M., "Learning Technologies and Educational Equity: Charting Alternatives to the Troubling Pattern of Big Promises with Dismal Results," *Teachers College Record*, (2016), ID Number: 21616.

[73] Ray, V., "The unbearable whiteness of mesearch," *Inside Higher Ed*, Volume **21** (October 2016). Available at https://www.insidehigher

ed.com/advice/2016/10/21/me-studies-are-not-just-conducted-people-color-essay, last accessed on July 31, 2022.

[74] Riggs, B., "Training matters! Narrative from a Black scientist," *Molecular Biology of the Cell*, Volume **32** Issue 3 (2021), pages 223–225.

[75] Rose, M. M., "Integrating Ethics and Diversity into Economics and Business Statistics," *Journal of Education for Business*, Volume **72** Issue 1 (1996), pages 13–17.

[76] Saltz, J. S., & Dewar, N., "Data science ethical considerations: a systematic literature review and proposed project framework," *Ethics and Information Technology*, Volume **21** Issue 3 (2019), pages 197–208.

[77] Schlenker, L., "The ethics of data science," *Towards Data Science*, May 9, 2019. Available at http://www.towarddatascience.com/the-ethics-of-data-science-e3b1828affa2, last accessed on July 30, 2022.

[78] Sfard, A, & Prusak, A., "Telling identities: In search of an analytic tool for investigating learning as a culturally shaped activity," *Educational Researcher*, Volume **34** Issue 4 (2005), pages 14–22.

[79] Shrout, A. H., Diaz Eaton, C., & Higdon, J., "Who Are We Centering in This?: Analyzing language around equity and justice in STEMM education policy," *SocArXiv*, June 4, 2022. doi:10.31235/osf.io/5837t

[80] TallBear, K., "Narratives of race and indigeneity in the Genographic Project," *Journal of Law, Medicine & Ethics*, Volume **35** Issue 3 (2007), pages 412–424.

[81] Tang, R., & Sae-Lim, W., "Data science programs in US higher education: An exploratory content analysis of program description, curriculum structure, and course focus," *Education for Information*, Volume **32** Issue 3 (2016), pages 269–290.

[82] Tolbert, S., & Bazzul, J., "Toward the sociopolitical in science education," *Cultural Studies of Science Education*, Volume **12** Issue 2 (2017), pages 321–330.

[83] Tuck, E., & Yang, K. W.. *Toward What Justice?: Describing Diverse Dreams of Justice in Education*, Routledge, Abingdon, 2018.

[84] Utts, J., "Enhancing Data Science Ethics Through Statistical Education and Practice," *International Statistical Review*, Volume **89** (2021), pages 1–17.

[85] Vakil, S., "Ethics, identity, and political vision: Toward a justice-centered approach to equity in computer science education," *Harvard Educational Review*, Volume **88** Issue 1 (2018), pages 26–52.

[86] Vardeman, S. B. & Morris, M. D., "Statistics and Ethics: Some Advice for Young Statisticians," *The American Statistician*, Volume **57** Issue 1 (2003), pages 21–26.

[87] Wilkinson, M. D., Dumontier, M., Aalbersberg, I. J., Appleton, G., Axton, M., Baak, A *et al.*, "The FAIR Guiding Principles for scientific data management and stewardship," *Scientific Data* Volume **3** (2016), Article 160018. doi:/10.1038/sdata.2016.18

[88] Witherspoon, G., *Language and art in the Navajo universe*, University of Michigan Press, 1977.