2021

Exploring Approaches to Data Literacy Through a Critical Race Theory Perspective

Britney Johnson
Georgia Institute of Technology

Benjamin R. Shapiro
Georgia State University

Betsy Disalvo
Georgia Institute of Technology

Annabel Rothschild
Georgia Institute of Technology - Main Campus

Carl Disalvo
Georgia Institute of Technology

Exploring Approaches to Data Literacy Through a Critical Race Theory Perspective

BRITNEY JOHNSON

Georgia Institute of Technology, bjohnson324@gatech.edu

BEN RYDAL SHAPIRO

Georgia State University, bshapiro@gsu.edu

BETSY DISALVO

Georgia Institute of Technology, bdisalvo@cc.gatech.edu

ANNABEL ROTHSCHILD

Georgia Institute of Technology, arothschild@gatech.edu

CARL DISALVO

Georgia Institute of Technology, cdisalvo@gatech.edu

In this paper, we describe and analyze a workshop developed for a work training program called DataWorks. In this workshop, data workers chose a topic of their interest, sourced and processed data on that topic, and used that data to create presentations. Drawing from discourses of data literacy; epistemic agency and lived experience; and critical race theory, we analyze the workshops' activities and outcomes. Through this analysis, three themes emerge: the tensions between epistemic agency and the context of work, encountering the ordinariness of racism through data work, and understanding the personal as communal and intersectional. Finally, critical race theory also prompts us to consider the very notions of data literacy that undergird our workshop activities. From this analysis, we offer a series of suggestions for approaching designing data literacy activities, taking into account critical race theory.

CCS CONCEPTS • Human-centered computing ~ Human computer interaction (HCI) ~ HCI theory, concepts and models • Social and professional topics ~ Professional Topics ~ Computing education

Additional Keywords and Phrases: Critical Race Theory, Data Literacy, Qualitative Methods, Participatory Design, Education/Learning, Workplaces

ACM Reference Format:

Britney Johnson, Ben Rydal Shapiro, Betsy DiSalvo, Annabel Rothschild, and Carl DiSalvo. 2021. Exploring Approaches to Data Literacy Through a Critical Race Theory Perspective. In Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems (CHI '21), May 08–13, 2021, Yokohama, Japan. ACM, New York, NY, USA, Article 706, 1-15. DOI: https://doi.org/10.1145/3411764.3445141.

1 INTRODUCTION

The field of human-computer interaction (HCI) is beginning to recognize that race and racism need to be considered in the design, development, and deployment of technology. For example, some research shows that pre-existing biases about race, politics, and gender, are often reflected in technology, making it less usable and more harmful to underrepresented minorities [24]. Other studies call for HCI researchers to become attuned to issues of race by inviting the participation of historically underrepresented minorities in their research activities and then actively combatting racial disparities that are subsequently identified [43]. Still others have demonstrated the importance of critical race theory (CRT) as a methodology in HCI, to help achieve fairness in algorithmic systems and data. CRT interrogates systems and data from the perspective of who is served through racial classifications and who else has the power to create them [25].

Ogbonnaya-Ogburu et al. highlight the pressing need to engage with race in HCI, particularly through the application of CRT as a lens to conduct research [43]. This paper draws from and extends this body of research in a variety of ways. We leverage CRT as a lens to conduct research to support equitable computing and data practices in the context of everyday work. Namely, we describe and discuss a series of workshops conducted with novice data workers as part of DataWorks: a work training program for developing entry-level skills in data science that aims to broaden participation in the everyday work of computing. Through these workshops, data workers explored issues of their choosing such as gentrification, gun violence, and mental health by locating, analyzing, and formatting data related to those issues.

These workshops were designed for a dual purpose. As part of the work context, they were meant to enable the data workers to develop and practice data skills. The workshops also provided an opportunity for researchers to understand how the data workers' lived experiences, including with race and racism, figured into novice data work. While we **did not** use CRT in the design of the workshops, CRT was the framing perspective we used in our analysis and interpretation of the workshops. This disjuncture has proved useful, as it highlighted problematic assumptions in the very framing of our research, prompting us to question notions of data literacy and workforce development.

Our analysis of these workshops illuminates a set of assumptions and dilemmas at the junction of lived experience, race, and data literacy. We suggest that responses to these assumptions and dilemmas will enable us and others to move towards understanding and fostering equitable computing and data practices in the context of everyday work. The assumptions and dilemmas identified are certainly limited, because they emerged from a preliminary project. Nonetheless, they make a meaningful contribution to the growing body of work at the intersection of HCI and CRT.

We begin by briefly discussing efforts to broaden participation in computing and describe the DataWorks program. Next, we review related work on data literacy, lived experience and epistemic agency, and critical race theory. A description of our methods follows, including how we leveraged CRT in our analysis of how we designed the workshops, which serve as the empirical basis of this paper. We then illustrate key themes from our analysis of these workshops, as well as from critical reflection on the research process and the DataWorks program. We conclude by highlighting key implications and limitations of our work that draw from and extend existing efforts to integrate HCI and CRT.

2 MOTIVATION & BACKGROUND

2.1 Broadening Participation in Computing

For decades, United States researchers and practitioners have sought to broaden participation in computing [46, 21]. One of their important goals has been to increase the number of Black Americans in computing fields by improving the pipeline of potential computer science students. However, while these efforts have been in place since the early 2000’s, there has not been a measurable impact on the percentage of Black Americans majoring in computing fields. Consistently over the past two decades, only 3.6% of undergraduate degrees in computing have been awarded to Black Americans [7]. There have been isolated efforts to sustain computing engagement with minoritized youth through opportunities such as jobs [14], robotics clubs [67], and paid participation in advance placement (AP) coursework [18]. But these unique, high-touch programs are expensive and difficult to replicate. More frequently, we see programs that seek to change youths' opinions on what computing is and what it can do as a way to build their interest in computing [41, 46], make computing personally relevant [17], and tie computing to their culture [33]. While such programs are necessary to build an interest in computing, there is little evidence of their long-term impact when they are not coupled with other ongoing high-touch efforts. This literature informs our work in this paper because it suggests that looking at the lived experiences of young people from historically minoritized communities—how they engage with data, how they perceive their communities in data, and how they see themselves in the technology workplace—can tell us about concrete and pragmatic ways to engage them in computing.

2.2 What is DataWorks?

Recently, many models have been proposed to develop training programs that “skill up” workers to work with data and contribute to a data-driven economy. Many of these models take the form of virtual data science courses or bootcamps, such as those provided by General Assembly (see: https://generalassemb.ly/). These courses or bootcamps are typically not free, nor are they designed for people from historically minoritized and underrepresented backgrounds. Other models, such as Data Science for All/Empowerment, are explicitly designed for such people, but continue to take the form of virtual coursework taught by university faculty on weekends or evenings [8]. More recently, organizations such as the National Science Foundation (NSF) have proposed innovative models incorporating data science fundamentals at Historically Black Colleges and Universities (HBCUs) [62]. However, these models remain early in development and continue to use formal coursework as a means to engage underserved communities in data work.

DataWorks draws insights from these programs to develop a new type of model for providing data services to companies and non-profit organizations through a workplace training program housed within universities. Namely, DataWorks employs people from historically minoritized and underrepresented communities to learn basic data work such as cleaning, formatting, and labeling datasets submitted by companies and non-profit organizations. Moreover, DataWorks also draws insight from informal learning programs, such as Glitch Game

Testers [15], that leverage principles of legitimate peripheral participation [36] to develop a community of practice centered around work with data, and structure expert-novice relations as data workers enter and progress through the DataWorks program.

Our motivation for creating DataWorks is two-fold. The first motivation for creating DataWorks is to develop a model for training members of historically minoritized and underrepresented communities in entry-level data science skills as a pathway to long-term, full-time employment. In this context, we use the term “minoritized” after the work of Roderic Crooks, who developed it to signal that these communities have been subjugated through actions and systems of oppression [10]. The second motivation for creating DataWorks is to develop a workplace training model that supported sustained research into the contextual qualities of data science in an authentic work environment. Through these efforts, we aim to develop new approaches to data science that acknowledge and respect diverse subjectivities, and do not just reproduce the white, male, heteronormative hegemony of data science.

Our work in this paper focuses on the first DataWorks program launched at Georgia Institute of Technology, a large public technology university located in the U.S. South. This program was formally launched in January 2020 with the goal of being self-sustaining and scalable by the end of 2021. Thus, this paper reports on research about the first four data workers hired in this program. These data workers were hired as temporary employees and paid a competitive hourly wage ($15 per hour) for 20 hours of work a week. The cohort included three women and one man, all Black and aged between 19-26 years old.

3 RELATED WORK

3.1 Data Literacy

Our goal for the four individuals employed in the DataWorks program was for them to see data as the object of a mechanical process, or a series of *data moves* [66], while simultaneously preserving the dataset's original context. Understanding data from mechanical and sociotechnical perspectives required developing their basic data literacy, along with critical data literacy.

Owing to its transdisciplinary nature, data literacy has varying definitions. In the information science field, data literacy is understood as the process guiding information retrieval, while in statistics the same term is applied to analytic skills associated with reading and interpreting information [56]. Koltay also discusses how literacy, once a straightforward foil to illiteracy, is now difficult to define in today’s expansive digital landscape [34]. Clegg et al. clarified data literacy within the CHI community, specifying two major groupings of definitions: first, as a subset of information literacy, emphasizing the discovery and use of relevant data; and second, as fluency in the core concepts and methods of data science [6]. Work focused on developing data literacy skills is frequently stratified by this division, as well as two specific student maturity levels: K-12 and professional students. There is extensive research on and implementation of instruction for K-12 students in proper consumption of insights and visualizations generated from data, following the informed citizen theory [68, 20]. For professional students, there has been research on understanding data literacy as an outgrowth of technical data science skills [22, 31] in addition to the many for-profit educational efforts (e.g., General Assembly and other “coding bootcamps”). Professional data science practitioners are commonly instructed in the finer technical details of the data lifecycle, or the generation of new data-driven insights, of which data literacy is seen as an extension [55, 61, 35]. Efforts to integrate these two subsets of data literacy seem to appear only in courses offered in higher education [32, 23, 30], making them inaccessible to people engaging in data work at an entry level due to time and financial constraints.

There is an aspect of power to data that is not commonly defined as part of basic data literacy, which Tygel & Kirsch labeled *critical data literacy* [64]. They adapted Paulo Freire's literacy method to explore digital inclusion and advocacy since the computing revolutions of the mid-2000s, citing data's ability to empower [64]. There are two key functions of critical data literacy. First, applying critical data literacy requires understanding the data's origins and acquisition, along with the positions and perspectives of everyone who touches that data—or the "infrastructural black boxes" [48]—to give a more nuanced view of what the data actually represents. For example, Loukissas explores how data is the *product* of place, but is commonly understood only as representations of place [38]; Taylor et al. further discuss the role of physical and social geography in shaping data [63]; and D'Ignazio and Klein show how implicit assumptions about neutrality and objectivity are baked into data's very definition, resulting in the neglect of important actors and prerogatives [11]. Critical data literacy can thus be used to understand data more fully as the product of complex societal factors.

Second, critical data literacy promotes finding, analyzing, and presenting data to substantiate or promote a new or clarified narrative. W. E. B. Du Bois' data portraits, for example, utilized data to disrupt the common conceptions of White Americans about their Black peers, and provide a more realistic and nuanced discussion of their experiences [2]. Similarly, Criado-Perez utilizes a more contextualized interpretation of common societal statistics to exemplify a consistent bias against women [47]. Other re-evaluated accounts of technical innovation from the lens of socioeconomic status [61] and race [42] exemplify this aspect of critical data literacy. Onuoha even uses a lack of available data to explore blind spots or gaps in our prevailing narratives about various social issues [44].

Altogether, our goal in the workshops described in this paper was to help the data workers to see data from both mechanical (data literacy) and sociotechnical (critical data literacy) perspectives. Through the mechanical perspective, we anticipated they would learn about the process—the series of steps—applied while interacting with data during the workshops. By pairing this mechanical perspective with one that is sociotechnical, we expected the data workers to gain insight into how social structures and factors may have impacted the data presented. We hoped that combining these two perspectives would provide the data workers with the heightened ability to examine data beyond its final format. As the implications of critical data literacy and its key concepts are relatively new, there exists limited work on how to build this skillset, especially in professional settings. For example, Hautea, Dasgupta, and Hill explore the critical data literacy skills of young students exploring data analyses in and about the Scratch programming language [29]. However, a key feature of this work is that it exists in an environment built for inquiry, while professional environments often are not. Our work addresses this gap by providing insight into what challenges a critical data literacy educational program faces, particularly in the professional context, by asking the data workers to collect and analyze data on topics that were personal and important to them, in an effort to engage them in this critical extension of common data literacy.

3.2 Leveraging Personal Data to Support Lived Experiences and Epistemic Agency

This research also draws from an emerging body of work related to data literacy from the field of the Learning Sciences that suggests leveraging personal relationships to data provides powerful ways to engage learners across disciplines including computer science. For example, as Lee describes, within a statistics education context, “Personal activity data (PAD) obtained from activity trackers has the potential to stimulate thinking about statistics in a way that other forms of data, even other real data, cannot. Because the data come from the students’ own activities, they are intimately familiar with them and able to reason about patterns and variations in the data based on their own experience” [37]. Similarly, Hautea, Dasgupta, and Hill illustrate, within a social computing context, novel designs for children that engage youth in critical data science by leveraging public data about children’s own learning and social interactions online [29]. Likewise, our prior work has developed a method called Mapping Self in Society (formerly Re-Shape) that allows students to collect, process, and visualize their own physical movement data in ways that support critical reflection and coordinated classroom activities about data, data privacy, and human-centered systems for data science [59].

In educational contexts, the motivations for this new body of work include calls for K-12 education to broaden learners’ epistemic agency within classrooms. For example, as Hardy et. al state, “we must broaden the sense of epistemic agency in science education to include students’ engagement with aspects of activity relevant to prior knowledge, identities, areas of expertise and desired learning trajectories” [27]. Miller et. al define epistemic agency as “students being positioned with, perceiving, and acting on, opportunities to shape the knowledge building work in their classroom community” [40]. Epistemic agency is important for learners to develop a personal approach to learning and acquire knowledge that is meaningful to them. An aspect of epistemic agency may be skepticism or criticality towards standard processes or content of learning.

Taken together, the workshops we describe and analyze in this paper leveraged learners’ alignments with personal data to recognize their lived experiences and provide them with epistemic agency to participate in meaningful knowledge construction outside of the K-12 education system. In particular, instead of applying the construct of epistemic agency in the classroom, we sought to apply it in the context of the workplace. Similarly, by centering the lived experience of the data workers and allowing them to set the topics and purposes of what data they selected and why, we intended to provide the conditions for epistemic agency. As data novices underrepresented in computing and data-related fields, we sought to examine how the data workers’ identities may or may not have affected their role as data workers, as well as how they engaged with data. Likewise, the design of the workshops described in this paper encouraged each data worker to bring the context of their experiences, along with their identities, to choose topics that were important to their personal or community interests. In other words, our work builds upon the aforementioned research, and positions the data workers as experts of their lived experiences, so they are the closest to a solution (in this context—a data solution) to the problem. In the context of the workplace, we sought to compare how the conditions of epistemic agency would help them to devise a data process of their own using data personally relevant to their own experiences.

3.3 Critical Race Theory

Our work in this paper is deeply motivated and informed by Critical Race Theory. In their paper “Critical Race Theory for HCI,” Ogbonnaya-Ogburu et al. discussed how CRT can be adapted in HCI research as a perspective for race-conscious efforts [43]. We draw directly from their work and recommendations to use

CRT as a lens to analyze, interpret, and question the workshops presented in this paper, the DataWorks program, and the very concept of data literacy. In other words, CRT is a race-conscious way to frame our reflections on what it means to support equitable computing and data practices in the context of everyday work.

Developed in the 1970s, Critical Race Theory (CRT) began to take root when legal scholars and activists noticed that the advances of the Civil Rights Movement appeared to be retracting [12]. Delgado and Stefancic describe CRT as a movement, specifically “a collection of activists and scholars engaged in studying and transforming the relationship among race, racism, and power” [12]. Not only did it focus on societal issues of race and ethnicity—CRT extended these issues into a broader perspective. It often related race to its position within the contexts of education, history, and self-interests.

With roots in critical legal studies and feminism, CRT analyzes and challenges the structures of what is proclaimed to be liberal and question its foundations, guided by several key tenets and themes. One of the founding CRT scholars, Derrick Bell, wrote about how critical race theorists do not necessarily condemn liberal ideology. However, he believed they were “highly suspicious of the liberal agenda, distrust its method, and want to retain what they see as a valuable strain of egalitarianism which may exist despite, and not because of, liberalism” [3]. Bell ultimately argued that the work of CRT would make way for legitimacy and justification that it deserved.

Contemporary themes of CRT include interest convergence, material determinism, racial realism, revisionist history, critiques of liberalism, and structural determinism [13]. Furthermore, these themes can be divided into more distinct tenets, which are acknowledged by many critical race theorists [12]. These tenets include:

- Racism is ordinary, not aberrational—when not acknowledged, there is little room to address it
- Interest convergence—notions of advancement for Black people are often agreed upon or sought after due to the advancement of white interest
- Social construction—race is a product built by societal thought and categorization
- Unique voices of color—race amplifies the experiences and thoughts of those who have been oppressed, so that their unique perspectives can be shared

Over time, CRT scholarship has expanded beyond civil rights work and legal scholarship and been applied to different disciplines—examining and challenging how these disciplines were constructed. The themes and tenets mentioned above have influenced disciplines including education and the social sciences. Recently, there is a growing body of work in HCI that leverages CRT as a lens to analyze research in computing. Ogbonnaya-Ogburu et al.’s recent work illustrates one such effort [43]. Another effort is the research of Hanna et. al in their paper, “Towards a Critical Race Methodology in Algorithmic Fairness,” which acknowledges and takes a deeper look into how race is embedded in algorithmic systems [25]. Namely, they demonstrate how oftentimes, we consider how race is amplified unjustly through these algorithmic systems, but rarely do we consider how the categorization of race can assist in amplifying these injustices. Therefore, it is important to recognize and be cognizant of the structural and institutional factors of race in algorithms, so we can adequately measure what fairness is. The authors share how race is more than just an attribute; it should be treated as the root of “a structural, institutional, and relational phenomenon” that highlights the prominent aspects of algorithmic unfairness.

Another area of HCI work that is influenced by CRT is the rethinking of intersectionality in HCI research. For example, Rankin and Thomas write about the importance of citing not only Black scholars or women, but to specifically cite Black women in scholarly work and acknowledging the work and challenges of Black women's intersectional identities. Coined by Kimberlé Crenshaw in 1993, the term intersectionality rose to draw attention to how Black women were experiencing discriminatory housing and legal practices [52]. Rankin and Thomas, in their paper, describe how recognizing the deep and true history of intersectionality is important when engaging in this analysis. They also share how intersectionality has been greatly informed by CRT, primarily in its stance that "when it comes to social inequality, people's lives and the organization of power in a given society are better understood as being shaped not by a single axis of social division, be it race or gender or class, but by many axes that work together and influence each other" [52].

As these examples show, CRT is central to the dialogues in HCI around race, racism, and equity. In our following analysis, interpretation, and critical reflection, we use CRT tenets to identify assumptions and dilemmas at the junction of lived experience, race, and data literacy that improve the DataWorks program. This work serves as an example of how CRT tenets can be used as both an analytical lens and a structure for equitably building future work and education programs.

4 METHODS

This section outlines the design and ethnographic methods used to gather and analyze data from the workshops that serve as the empirical basis of this paper. We include a discussion on the role of design workshops in research, limitations of design workshops as a method, and discuss how the methodology of design research impacted our workshops. To conclude this section, we acknowledge our position in this work and how it influences our subsequent analysis.

Our research combines design and ethnographic methods. We developed a series of workshops (detailed in section 5) in which the data workers identified a topic of interest, collected data related to that topic, analyzed the data, and produced and shared a presentation communicating their interpretation of the data. Throughout these workshops, we took field notes and documented workshop activities with audio, video, and photographs. In addition, we conducted semi-structured interviews with each data worker before and after the workshops, and we conducted a group interview with 3 of the 4 data workers.

Qualitative analysis was conducted by four of the authors, who read the interview transcriptions and watched the presentations, to identify themes and patterns. Subsequently, the research team discussed the themes and patterns collaboratively and iteratively, identifying and refining themes over the course of several meetings. Based upon this initial analysis and discussion, the team identified CRT as a lens for analysis. The first author then reviewed the interview transcriptions, videos of the workshops and presentation, and other documentation—such as images and drawings—using CRT as an analytic lens. Data was then presented to and discussed with the whole research team during multiple meetings using CRT as an organizing principle with an emphasis on the CRT tenets to further develop categories and themes. Three themes emerged related to CRT:

1. Tensions between epistemic agency, critical data literacies, and the work context
2. Encountering race, racism, and its ordinariness, through data work
3. The personal is communal and intersectional.

The first author created narratives for the four data workers that demonstrated examples of these three themes in each of the data workers workshop experiences. The authors returned to the interview transcripts and the presentations to identify how the themes were supported or refuted, and to reflect on the implications of the themes in the context of the workshop and the DataWorks environment. Narratives were refined to select examples that best articulated the themes.

Design workshops are common in HCI research, and often draw from practices of participatory design and codesign. The format of the design workshop brings about both challenges and opportunities. As Rosner et al. noted, “when design workshops act as research instruments, they shift the form and character of collaborative work” and move us “from knowing to doing” [54]. However, such workshops can also be exclusionary, either because of who is invited to participate or because of how the activities of participation are designed [28, 54].

With these challenges and opportunities in mind, we developed the workshops analyzed in this paper as casual activities, with the intent of providing the opportunity for the data workers to determine how they wanted to pursue these activities. While we did establish and follow a general structure through a sequence of workshops, we did not pre-determine the specifics of data practice. For instance, we did not dictate how data workers would source their data, what kinds of analysis they would do on the data, or what the topic of their inquiry would be. Moreover, to respect their status as employees, these projects were meant to enable the data workers to practice their skills in data work in an authentic environment, which was the basis of their employment.

Prior to providing a detailed accounting of the structure of these workshops, it is important to acknowledge that the workshops were conceptualized and facilitated by the first author of the paper who is a Black American Ph.D. student at an R1 university with an educational background in computer science. Additionally, this student collected most of the workshop data and was the closest in age to the data workers, being one to seven years older than each data worker. The additional authors of the paper are all white Americans, three of whom hold faculty and research positions at universities and one who is a Ph.D. student. Their areas of expertise include learning sciences, data visualization, and participatory design. These authors, with the exception of the Ph.D. students, are much older than the data workers, and while they did engage weekly with the data workers, it was primarily through management and training roles.

5 THE WORKSHOPS

The workshops consisted of seven two-hour sessions, with one workshop approximately every week during February and March of 2020. The workshops took place at the end of the data worker’s shift, so for that day, these two hours were the final hours of their four-hour workday. The data workers were paid for the time they spent in the workshop and the activities were treated as part of their weekly work. Six out of the seven sessions were held on a university campus in the DataWorks office space. The space was large, modern, and open, with workstations, a communal worktable, and whiteboards around the room. For group discussions, data workers gathered at the center office table while activities requiring computers was completed at their work desks. The final session, the presentations, were conducted virtually due to the COVID-19 pandemic. For most of the workshops, we focused on scaffolding the process by providing the data workers with a general structure and sequence to guide them. This sequence was outlined through four steps to 1) specify a problem, 2) find applicable data, 3) analyze and organize the data, and 4) communicate the data. At the end of this process data workers further communicated the data by leading a final presentation to an audience that included researchers, academics, and clients.

5.1 Activity 1: Specify a Problem

In the first activity of the workshops, the facilitator (the first author) introduced a multi-week project—using data to address topics that were meaningful to the data workers—by asking them to share issues that were important to them and why. Notably, within seconds, the data workers enthusiastically began sharing a wide range of topics. The topics at first reflected issues that often make the news, such as gun violence, traffic, and homelessness. With time, more personal topics, such as mental illness, wellness, and time management emerged.

After this initial discussion, each data worker was given a storyboarding worksheet with a series of questions to guide them in thinking through how they might use data to address a topic of interest. To scaffold this process, the facilitator introduced an example problem space: how to gather data to help determine where and how to build a grocery store for their community. The data workers helped complete the storyboarding worksheet, sharing what information they did and did not know about the problem, and what data could be useful to address the problem. The example was intended to enable the data workers to shift their mindset from the topic alone to how data could be applied to the topic.

Each data worker was then given ten minutes to complete the six-question storyboarding worksheet about an issue they had selected. Once they filled it out with their initial thoughts, we discussed the topics they chose and their answers to the remaining questions on the sheet. We asked each data worker “Why did you choose this problem?” and “How would you address this problem with the data?” While some data workers listed local community institutions (such as schools) as resources to gather data, others did not know how and where they could find the data necessary to address their selected topic. We paused, then allotted 30 minutes for them to do a quick online search for applicable data. They each initially started their search on Google’s search engine by typing in keywords associated with their topic along with the word “statistics.” After 30 minutes, we reconvened as a group and discussed the data they found.

5.2 Activity 2: Find applicable data

In the second week, we convened again at the communal table. Building from the prior week, the goal of this activity was to find data applicable to their chosen topic. What counted as “applicable data” was open to their interpretation.

To begin this activity, each data worker returned to their storyboarding worksheets from the week prior and shared a recap of what was written. The purpose of these recaps was to convey to the data workers the importance of iterating throughout the data process. The discussions which took place in the latter part of the first activity could have changed some of the responses written on the storyboarding worksheet during the beginning part of the activity. Therefore, we wanted to invite space to discuss these changes.

Listening carefully to each data worker’s responses, the facilitator wrote key points on the whiteboard about their projects, including what they knew about the problem and if anything had changed. The group also discussed what additional data sources needed to be added. For the next hour, data workers returned to their desks to find additional data and were asked to format that data so it would be readily accessible for presenting. After searching for data, the group reconvened at the common table for 10 minutes to discuss the data they found, the format in which they found it, and what process they used to find it.

This activity facilitated an important discussion about how the data was formatted (e.g., in Excel or text format). Of note, each data worker organized their data into a Microsoft Word document through the copy/paste feature rather than in what our team considered more “accessible” formats such as Excel. We hoped the data workers, who already had several weeks of training in data preparation and cleaning, would transfer that knowledge to their own projects. However, this did not happen. We realized that we, as researchers, must recognize how the social locations of the data workers (including economic status) influenced their choice to organize the data they gathered into Microsoft Word. Their choice could depend on several factors, like ease of use, time constraints, and level of knowledge about the tool. This finding helped us to recognize that in order for the data workers to be able to apply the skills they were strengthening during trainings, they would need further instruction on how to gather and format the data with a diverse set of tools, depending upon how they wanted to communicate or share the data.

5.3 Activity 3: Share this data

In week three, the facilitator emphasized to the data workers why organizing data is considered important, and how to organize their data in a desired format. While gathered at the communal table, the group revisited the roundtable discussion employed during the first activity. For ten minutes, the data workers shared many ideas about how to communicate data (e.g., through social media, newspapers, billboards, email, paper flyers, and television commercials). There was no mention of computational examples of how data was shared. The facilitator then showed online examples of how to communicate data, including typical news media use of graphics, video, along with narrative storyboarding that used data and murals that blended artistic interpretation with data in public spaces. These visualizations engaged the data workers, prompting questions about the data, and inspiring them to create their own visual data representations, rather than coping and pasting from existing sources.

For 30 minutes, the data workers returned to their personal workspaces to search for ideas and examples of how they would like to share the data. They were encouraged to think about what data needed to be gathered and how it needed to be formatted to fit how they wanted to communicate the data. During our 15-minute discussion at the end of the activity we found that most data workers chose to use charts and graphs to share their data. Creating charts and graphs was new for each of them, but they were excited to figure out how.

5.4 Activity 4: Communicate this data

Original plans were for the fourth week to be the final week of work, and the fifth week to be presentations. However, more time was needed for the data workers to engage in the experience of creating visualizations and presentations, which was a new experience for them. Therefore, we extended activity four from just one week to three weeks. During these last 3 sessions, the data workers were told to allocate two hours of their day further researching their topic, finding relevant data and reformatting it for the final presentations. The data workers would often alternate between their client work for the DataWorks program and their project. This self-regulation turned out to be efficient for the data workers. The facilitator was present during the usual workshop time of 11 AM-1 PM each week to answer any questions, and other team members were also present at other times.

Working on their projects individually while asking questions allowed the data workers to consider what would be important to share with the team after gathering their data. Condensing their findings into a presentation format and selecting certain data to share was challenging for them. Most of them needed to enter the text data they had gathered into Excel sheets, so it could be better understood visually through a bar or line graph, which caused them to reflect on how they collected data. They also used online resources to teach themselves how to make charts and graphs, and learn basic spreadsheet skills.

5.5 Final Presentations

To further communicate their data with others, each data worker was given eight to ten minutes to present their findings to an audience that included researchers, university professors, and clients. These presentations were conducted virtually through video conferencing software. These presentations were organized to provide the data workers experience with professional presentations, and give them an authentic motivation for gathering appropriate data and communicating it clearly. The presentations were also a way for the whole team—from incoming masters students to senior faculty—to see the data workers talk about issues they cared about, share advice, give encouragement, and build community across the DataWorks project team. We intended for these to be in-person presentations with a celebratory feeling. However, just before the data workers were ready to present, the university issued stay at home orders in response to the COVID-19 pandemic. Organizing and scheduling the online video conference was challenging, as several workers did not have reliable access to technology or internet connections at home. The presentations were held two weeks later than anticipated, and in two sessions rather than one. Three data workers presented in the first session, with a large audience of nine people. The fourth data worker presented two weeks later with a smaller audience of five.

Despite these challenges the presentations were successful. They started with the facilitator modeling professional skills, thanking everyone joining, and encouraging them to provide helpful feedback. Then the facilitator outlined the structure of the presentations, time available to present, and an overview of the workshop's objectives and goals. The data workers presented their work, expressing that they were nervous through words or actions, and the guest gave positive feedback while asking questions to encourage the data workers to look more critically at the data.

6 WORKSHOP FINDINGS

In this section, we discuss three primary themes that emerged from our analysis of data we collected during these workshops in ways that allow us to critically reflect on our research, our notions of data literacy, and the DataWorks program. For each theme, we use a detailed analysis of one data worker to make connections between their choice of topic and their experiences sourcing, formatting, and presenting the data. Then, we characterize other data workers' experiences, and raise questions and considerations about data literacy that we return to at the end of this paper. A primary motivation in doing so is to examine data literacy in relation to these themes, and present how the data workers approach data literacy from a different perspective. Notably, three of four of the topics selected by the data workers are imbricated with issues of race in the United States: gentrification, gun violence, and access to mental health resources. The fourth topic, bullying, is not so clearly imbricated with race, but does reflect the lived experience of the data worker who pursued this topic. These themes also reveal the tentative relationship between the skills the data workers learned through their training in the DataWorks program and their personal projects.

6.1 Tensions Between Epistemic Agency, Critical Data Literacies, and The Work Context

Chase is in her early 20s, and a resident of the English Avenue neighborhood. While this neighborhood is vibrant, it has also suffered systematic neglect since the "white flight" of the 1960s and 1970s. Chase is lively and vivacious, often engaging in animated conversations with the other data workers. Once the work began, Chase often adopts a leadership role by leading calls with clients. She is also the DJ, and enjoys setting a vibe for the office by playing old-school 90s jams. Chase has a self-assured presence, perhaps because she had worked at the university as a temporary worker assisting a lead administrator in a department on campus.

For her personal data project, Chase had many ideas to begin with, and settled on the topic of mental health. More specifically, she was concerned about the benefits of free therapy for teens and young adults, and the lack of such resources in her community. When asked why she picked this topic, she shared that members of her family dealt with mental illness, but they did not go to therapy. She set about collecting data on mental health in the United States, specifically looking at mental health issues in different populations. But, similar to many of the personal data projects, her focus shifted based upon the data she was able to locate. She ultimately focused on suicide as a particular mental health issue. Her presentation used data from several sources to make an argument for the pervasiveness of mental health issues, and the need for more resources in her community.

After the workshop, we asked the data workers about how they approached their personal project, and how it was similar to and different from their paid work. Chase's response was typical:

> "I feel like the steps that we did, like, look at what was in front of us, even though it was blank, like, we had you give us questions, and then, you know, find the data. That's what we did with our, you know, other, you know, stuff that we worked on. We found the data. Either we looked at other spreadsheets or we looked at, you know, government sites, and then we talked about, how did we want to present it, like, um, did we want to arrange it this way, that way, this order, you know, and then we presented it to them. So I feel like the steps were the same but I do totally agree on how we went about it is different."

Like Chase, another data worker, Jordan, told us the process she used for this project mimicked processes she used in her other data work. Jordan is in her late teens and a resident of the English Avenue neighborhood. Jordan is soft spoken, and exudes a creative spirit through sharing, writing, and art. While Jordan generally allowed other voices to dominate conversation, others listened attentively when she did talk. On Jordan's desk is a task board given by her much-admired sister—a student attending a Historically Black University—surrounded by cute, brightly-colored stuffed animals. In response to the same question we asked Chase, Jordan shared:

> "It's kinda similar I mean even though these were personal things, we still had to do research on it. It's like with [name omitted] somebody came in and talked to us about it and we still kinda like had to go out on our own and just look it up and find out like the different information and like how we stored the information we used PowerPoints in this one and the other one we used Excel spreadsheets but it was it's pretty much the same when you think about it."

What Chase and Jordan seem to be describing in their responses is that that the topic and presentation of their personal project, *but not the process*, was open to their interpretation. In general, they discussed their processes vaguely, without distinction or interpretation. The language they used matched the general process presented to the data workers at the beginning of the workshop to explain core methods of data literacy. Our goal was that the data workers would use this general process as a starting point and that together we would examine how to revise this process for their projects. While the workshops were able to teach this general process for data work, in this deeper exploration, the data workers did not deviate from or question the process. Thus it seems the structure of the workshop did not, in fact, provide the conditions for epistemic agency, such that the data workers could develop their own personal approaches to data work.

In addition to there not being evidence that the data workers developed their own personal approaches to data work, we found no evidence that the data workers brought their critical perspectives to the data they gathered. In all the presentations, they accepted the data as fact without questioning its source, methods for collection, or the analysis used to interpret it. In Chase's case, this lack of criticality was expressed in an exchange between one of the authors and Chase after her presentation on mental health resources. In the course of that presentation, Chase presented data on suicide rates, broken down by gender and race, saying:

> "This is a chart that I made, and this data came from the state of mental health in America website and this is 2020 data, so this is actually new or updated. So this is suicide rates by age and we have female ages and male ages and as you can see the males are higher than the females. Which was kinda surprising to me, but also not, because males tend to not get their feelings out, or was raised to not get their feelings out, or talk about their feelings to people. So they hold all their anger in, or they hold how they feel in so that carries on with them and it takes a toll on them. So it's not really surprising to me that its higher than females. And females tend to like carry on things and try to overcome things because we're very strong so I just feel like it's not surprising that their numbers are higher."

After presenting these statistics, one of the authors asked Chase if the data represented completed suicides or attempted suicides. Chase paused, realizing this was a perspective she had not considered, and recognized that those numbers might be quite different. Throughout most of the other presentations, we saw a similar pattern. In projects concerning gun fatalities, gentrification, and bullying (projects that will be described in the following sections), the data workers presented data without interpretation or understanding the context of the numeric findings. For instance, the data workers did not account for the difference between per capita and total numbers. As Jordan shared in her final presentation:

> "And 1 in 3 students are likely to be bullied in school, 30% of students surveyed have admitted to being the bully, and 70% of students recorded are bystanders... What I learned from this information is that there are more bystanders than there are of those being bullied. So if [we] were to like make bullying more aware to the 70%, we can make a safety net for those who are being bullied and ... be able to stop it. At least that's what I believe."

Accepting numeric data as a fact without questioning if it is ethical, accurate or distorted is common. For instance, Acker et. al [1] discuss how young people understand the data lifecycle through mobile phone usage without insight to how their data is created, managed, curated, and preserved in the background. In other words, only encountering data in its final format shapes the way they learn about the data. Loukissas makes a related argument, and advocates for understanding "data settings"—not just data sets—as a component of a critical approach to data [38].

Data workers in our workshops thus received and accepted their project data in its final format without acknowledging its source or the processes that made it. By accepting the data in its final format, we mean that the data workers presented the data as they found it, which was mainly in the form of text (or numbers), and used those statistics to create graphics such as pie, graph, and line charts. This is problematic because it grants data undue authority, and overlooks the social, cultural, material, and technical conditions of data. In order to be critical and present accurate narratives related to the data, data literacy emphasizes the importance of understanding the context of the data and its origin—including who was involved in the data lifecycle process.

We at first interpreted these findings as a lack of engagement with the process of data work and a lack of criticality in the exploration of data they sourced as a lack of epistemic agency. Epistemic agency would have been demonstrated if the data workers felt empowered to be skeptical of the process and developed critical perspectives about the data. Upon reflection, this should be unsurprising given the work context. As previously mentioned, the context of most critical data literacy programs are K-12 education and community engagement events. These contexts are often *already critical*, or at least already set within a context of inquiry. This is not the case for the context of work. In a work context, it is important to acknowledge and respect work subjectivities, power dynamics, and cultural context including racial identity, stereotypes, and assumptions about who can and cannot question authority. We return to this finding to illuminate broader implications of our work at the end of this paper, and discuss how issues of epistemic agency could foster a more diverse data science. This issue was also salient in the ways race was treated in the data presentations.

6.2 Encountering Race, Racism, and its Ordinariness, Through Data Work

The experience of another data worker named Valen demonstrates how experiences and work with data revealed the ordinariness of race and racism. Valen is in her mid 20s and the proud parent of a young child. Like Jordan, Chase, and Riley (who you'll meet next), she lives in the English Avenue neighborhood. In addition to being employed by DataWorks, Valen is also a security guard at the university. Her shifts as a security guard start mid-afternoon, so she often went straight from one job to another, and also picked up security shifts over many weekends.

Valen has a commanding presence. She stands out among the data workers by already demonstrating an interest in taking on more responsibility. Her voice captures the room when she speaks, as she shares her thoughts with passion. Yet, even when appearing to be confident, she often confided her anxiety by mentioning "Whoa, I am so nervous" when asked to speak in front of the group. She often speaks of her dream to go to a Historically Black College just a short walk from her home, but for now has put her creativity toward other efforts. For example, she excitedly talked about her do-it-yourself projects, taking budget-friendly furniture and décor and creating masterpieces to "spruce up" her room. Valen is considering real estate as another way for her to make money, and something she might be good at noting the large number of houses for sale in her neighborhood, but that most of them are not affordable to her or her family and friends. Perhaps not surprisingly, Valen decided to look into the topic of gentrification.

Valen began by looking for data about her neighborhood and the homes within it. She was surprised when her search results returned information about historic homes just a few blocks from her own that were owned by civil rights activist Maynard Jackson and Reverend Dr. Martin Luther King Jr. She had no idea that Jackson and King had lived in her neighborhood, and was emotional when she found out. Her emotions were complex, both proud and distraught. She was proud that she lived in this neighborhood where civil rights leaders had lived, but distraught that she did not know that these homes—and the legacy of her neighborhood—might be lost. As she continued her research, she continued to express surprise. Though she knew rents and the cost of homes were rising in her neighborhood, she had not known how widespread this was across the city she lived in, nor that so many others were concerned.

Through her search for data, she became aware of some of the conditions and complexities affecting real estate and gentrification. One such insight concerned the development of a walking path through an adjacent neighborhood. This walking path would make the neighborhood more appealing, and connect it to other neighborhoods across the city. Valen realized that this amenity would mean that the cost of housing would increase further.

Valen encountered racism in different forms through her data work. To be sure, as a Black person in a southern state in the United States, Valen saw racism every day. Data did not suddenly make racism apparent to her. What her search for data—and the data itself—prompted was an additional perspective on race and racism in her life, neighborhood, and city. She became aware of what she had not been taught about her own neighborhood, and about how the housing conditions affecting her neighborhood were part of a larger pattern. She began to realize the range of actors involved in this process, including those involved in real estate. For instance, in her presentation Valen stated, “Many developers, to me, think they’re just old buildings but they’re big artifacts to the neighborhood.” She also connected housing insecurity to class, through the data she sourced, saying, “This data goes back to the employment rates, which show people within the community that have minimum wage jobs and not career balance income.” Valen spoke about income, “outsiders” coming into the neighborhood, and issues related to the historical quality of civil rights workers’ homes. However, she did not explicitly name racism as an issue of gentrification of her neighborhood and did not use race as a way to dissect the data she presented.

Throughout all of the workshops, we were surprised that race and racism were not more present in the data workers discussions, particularly since the issues they chose were entwined with race. This lack of articulation may be due to the very ordinariness of racism in their lives. Valen’s initial lack of awareness of their neighborhood’s connection to the history of civil rights and the racist undertones of real estate, is not surprising. In fact, it is indicative of a tenet of critical race theory: racism is ordinary. As Ogbonnaya-Ogburu et al. note, “Racism is pervasive and ordinary in our society’s digital platforms and the larger socio-technical systems in which they are embedded… we have to shed light on the hidden traces of racism and acknowledge them” [30]. What Valen encountered through her data work was another manifestation of racism made ordinary. She did not need to call out race or racism, as it was obvious to her and to all who saw her presentation that the predominantly and historically Black neighborhood she lived in was becoming gentrified by white people. While the data she sourced did not identify race, racism was implied with each article that spoke about class and economics.

Valen is an example of how the lived experiences of data workers in the DataWorks program influenced their choice of topics, and also how the data they worked with caused them to think about their lived experiences differently. For instance, race and racism were apparent in data presented by Chase and another data worker named Riley concerning mental health resources and gun violence, respectively. But none of them called out racism in the data they presented. It was so embedded in the issues that it was an ordinary part of the issue, not a unique finding. In addition, related to the finding presented on the work context above, the racial makeup of the workplace (the senior members of the program are white) might have shaped whether the workers felt comfortable talking about race and racism.

Building from our previous section, Valen is also an example of how, on one hand, the data workers did not develop critical perspectives on the data. However, on the other hand, they did develop critical perspectives on the topic—the conditions the data is meant to represent. The discovery of the relevant data that Valen found led her to reflect on the sociotechnical factors she had personally experienced. Recalling concepts of data literacy and critical data literacy, the data workers were developing data moves that brought them to an understanding of the sociotechnical perspective, rather than the mechanical perspective [41]. Of course, the mechanical and sociotechnical perspectives need to be combined to achieve critical data literacy. Therefore, what is needed is to bring those mechanical (or technical) concepts into relation to the sociotechnical insights.

6.3 The Personal is Communal and Intersectional

The final theme we introduce in this section highlights that what is oftentimes considered to be personal is also communal and intersectional. Data worker Riley chose the topic of gun violence and related his own experiences to this theme. Riley is also in his early 20s and from the English Avenue neighborhood. From day one, Riley arrived at work dressed to impress. He wore pleated pants, a pressed shirt, a suit jacket and, at times, a blue high school letterman jacket—still proudly showing off his high school marching band affiliation, though he graduated years ago. Unlike the others, Riley was not one to carry a book bag and opted to travel light to and from the workplace. Riley speaks rarely and softly, and it could be difficult to hear him in meetings over others who were more boisterous. But he is also observant, always watching and learning. If you want to learn about anime, he is your go-to person, stashing anime figures under his computer monitor next to a picture of Kobe Bryant. Some days, as he worked, you could hear the sounds of HBCU bands coming from his computer speakers. Though he didn't still perform in marching bands, it was still meaningful to him.

The topic Riley chose to pursue was gun violence. He had previously mentioned his aspiration to become a police officer, and he connected that topic to this aspiration:

> "So why did I choose this topic? Well, one of the main reasons is because I want to be a police officer. And why I want to be a police officer is because in my community I see a lot of things I don't like and I feel like I can make it a safer place. And it's also like gun violence is a topic that needs to be spoken about and don't have a solution."

We began the workshops with prompts for data workers to consider topics of interest to them, and connect to their lived experience, so the connection to personal experience and aspirations was to be expected. What is notable is how Riley framed these topics from the perspective of his community. Community, as he used the term, referred to neighborhoods. His neighborhoods were defined both geographically and racially; they were a distinct bounded area, and also historically Black—a factor that was significant to his data project. A defining quality of his project, then, is his communal and intersectional perspective united his individual subjectivities with attachments and commitments to the social groups he saw himself being part of.

In other words, Riley shows why we, as researchers, should not assume that “the personal” is necessarily individual. Such an understanding echoes work emerging from personal informatics, which argue that “individual concerns become inevitably entangled with the lives of others—partners, children, colleagues, and employers” [16]. Emerging theories and principles of data feminism [15], as well as intersectional perspectives, have long argued for an understanding of subjectivities that are irreducible to singular identities [37].

More generally, this notion is also evident in how Riley saw gun violence as a topic that needs to be addressed through his potential future career in criminal justice. He explained that gun violence was “a topic that needs to be spoke about, like it’s something I mainly see in my community that’s like a number one problem for us.” By framing the issue this way, Riley was simultaneously looking toward the future with a career in criminal justice, while also looking at his past.

Riley’s broader experiences in relation to his choice of topic further highlighted how the personal could be intersectional. Specifically, we primarily noticed the intersection between race and gender. Riley shared with us that the actions of youth and his peers often concerned him, and he wanted to help contribute to a solution. After experiencing the loss of his friend and another friend to incarceration, he shared that he, “just felt like what’s the point? Why are we doing this?” He also shared that “we’re dying and we’re dying young... it’s going faster than what it should be.” Riley’s remarks expressed his sincere concern about how the issue of gun violence has not only affected his personal lived experiences, but also the experiences and lives of other Black men whom he was close with.

Jordan expressed similar sentiments when discussing her topic about how bullying affects mental health. She chose the topic for two reasons: because it was familiar to her and she saw that existing efforts to address the topic were lacking. She stated, “I knew more about [bullying] than any other topic that I wanted to do so that’s why I chose it.” She also described how, in high school, the anti-bullying campaigns were ineffective for her and her peers. She believed that data might be a way to convince others in her community of how pervasive and damaging bullying is. In this way, data might foster greater awareness. For Valen, her concern with gentrification was grounded in concern about her and her family’s homeownership, the effects of gentrification on others who lived in her neighborhood, and the effects of gentrification on the legacy and future of Black activism in her city. Her desire to become a real estate agent was, similar to Riley, both a personal aspiration and a means for supporting Valen’s community.

Through their projects the data workers present issues as both personal and communal, relating to them as members of a community. This is similar to what Peck et. al [26] refer to as “social framing” in their research into how novices understand and make use of data visualization: “Instead of ranking [data] charts and graphs based on which forms were most effective to them, these workers were concerned about the effectiveness of charts and graphs for other people” [26]. At the same time, a critical perspective might find both Riley’s and Valen’s perspective difficult, because their career aspirations were implicated in the very issues they wanted to address. The real estate industry enables and profits from gentrification, and the police are entwined with gun violence. But this complexity of subjectivities, this intersectionality, is precisely the condition of these data workers’ lives and the neighborhoods they seek to affect. This is not to lessen that problematic interplay, but to recognize and respect it as a very real set of intermingling circumstances, desires, and futures that should be attended to.

Altogether, our third theme begins to highlight that as we (collectively as a research community) develop concepts of critical data literacies, we should attend to how data (and its representation) not only connects to, expresses, and moves individuals. Data also connects to, expresses, and perhaps even moves communities and people whose race and gender (amongst many other axes) can influence people's experiences with social issues. Schlesinger et. al [57] challenged researchers to "increase our attention to the ways multiple facets of identity interact with one another when framing users lived experiences" [57]. This insight contributes to research attempting to design for effects beyond the singular or group category of "the user" or "users," and to consider more expansive collectives. One way to conceive of data literacies informed by critical race theory, broadly construed, is that the interpretation and use of data should be framed "beyond the individual." Criticality is not solely indexed by how an individual interprets the data in relation to themselves, but how they interpret the data in relation to a communal or intersectional experience or situation.

7 CRITICAL REFLECTIONS ON DATA LITERACY

Stepping back from the experiences of the data workers and reflecting on the workshops and DataWorks more generally brings into focus further issues. One of the core concepts of critical race theory is that liberalism as a political and legal philosophy may, in fact, hinder anti-racist agendas, perhaps even inadvertently advancing implicit bias [13]. This possibility gave us researchers pause. In what ways might DataWorks as a program be ensnared in problematic ideals and subjectivities? Upon reflection, one concern is the focus of DataWorks on traditional notions of data literacy. What is usually meant by data literacy are skills in the discernment and application of numeric data as information. Of course, there are many variations of data literacy, but many of them are reducible to this general concept. These skills and how they are measured tend to be grounded in traditional Western notions of what constitutes data. Even more broadly, they are anchored in assumptions of what constitutes veracity and objectivity: understandings of accuracy and legitimacy, coupled with values of detachment and neutrality. But the social sciences and humanities have demonstrated that concepts such as veracity and objectivity are not universals. Rather, they're perspectives instantiated through social and cultural practices [26, 5, 4]. As such, we need to consider how our focus on DataWorks as a program to develop data literacy might be reproducing and reinforcing hegemonic values and practices.

One way hegemony manifested was through our analysis of the data worker's projects. In the discussion of the final presentations by the data workers, we found ourselves calling out moments when they misinterpreted statistics or presented incomplete data as if it was comprehensive. In both the discussion of the workshops and the presentations, and throughout our analysis and interpretation, we also found ourselves calling out moments when the data workers seemed to be decidedly uncritical in their interpretation and use of data, accepting data as a given and applying it without reflection. Herein is a conundrum. On the one hand, we need to be attentive to these moments when the data workers make what are commonly referred to as "mistakes" with the data. We especially need to be attentive to these moments in which the data workers *seem* to not be perceiving the social, cultural, and political consequences of data. However, we also need to question whether such moments are really indicative of data literacy or critical data literacy. Moreover, we need to attend to what other qualities and markers might indicate data literacy, beyond those that are common to our current definitions. For example, Loukissas [38] and Klein and D'Ignazio [11] cause us to consider how examining the actual settings of the data can be a marker of data literacy, because the absence of this knowledge can lead to inaccurate analysis. If we are truly committed to a culturally situated program, then we need to begin to work together with the data workers to question what we defined as data literacy and on what basis we judge certain interpretations and uses of data as “literate.”

Furthermore, we have a nagging question about the very idea of data literacy as a meaningful outcome. Discussions of data literacy, including those we engage with, suggest that data literacy is an important skill for financial gains and civic engagement. That is, there is an assumption that data literacy leads to well-paying jobs and provides agency to make decisions about politics. However, that evidence is itself racially biased, because those financial gains and civic engagement are not equally distributed. In pursing data literacy, critical race theory and the concept of the limits of liberalism charges us to question whether data literacy is an outcome that advantages workers, or simply continues a liberal myth about equality and individual agency, and the capacity of self-determination in economics and politics. CRT causes us to question what is data literacy to the data workers, and how they can help us reimagine what data literacy should be.

Another central tenet of critical race theory is that advances in anti-racism often occur only through interest convergence—through programs and encounters that align with and advance the agendas of whites and other dominant subjectivities and positions. Is the context of work such an example of interest convergence? Likely so. But, as with the questions concerning the problematic aspects of liberalism in data literacy, we do not believe that to necessarily be so. Echoing earlier reflections, what is needed is to approach the context of work critically. Our efforts to explore the context of data work may be an opportunity to address racism through interest convergence and chart a path forward that will examine the impact of racism in the work. Rather than simply asking how a program like DataWorks can provide job training, we must also ask how a program like DataWorks can contribute to new imaginaries of contemporary labor, imaginaries, and practices that do not simply reinforce and reproduce the interests of the dominant Western, white, male, heteronormative cultures of data and computing.

8 IMPLICATIONS AND FUTURE WORK

In this section, we summarize four additional implications from our work in this paper. The first implication focuses on what it means to critically engage with the context of work to develop more expansive and equitable understandings of data literacy. The second, third, and fourth implications are related to the design of future workshops and programs centered around data or novice data workers. Through this work, we take first steps in using Critical Race Theory to support data literacy by providing concepts and ideas that can be used in future work, but also necessitate further expansion in future research.

8.1 Implication 1: Critically Engaging the Context of Work

Our analysis of the workshops shows how data workers were not called upon to question the data they were processing in their everyday work. They were asked to (and paid to) process data as directed by the project. Any lack of criticality in these workshops, then, should reflect back not upon the data workers, but upon ourselves and our structure of the work environment. As Philip et. Al [50] share:

> “Educators rooted in critical pedagogy argue that our contemporary schooling system tends to highlight the interests of groups such as corporations, the middle-class, and Whites, while obscuring the histories and perspectives of unionized labor, the poor, and racialized groups that struggled against slavery, conquest,

and exploited labor. Such partial portrayals reinforce inequitable and unjust processes in society and limit the range of alternative possibilities with which people might imagine and engage with the world."

This situation is exacerbated in the workplace, the very context that tends to, as Philip notes, problematically dominate education. When we conceptualize a critical data literacy, especially from the Frierian perspective that informs our work [64, 50], we must account for the dominant conditions of work that obfuscate the very histories and perspectives a critical data literacy is meant to bring to the fore. This challenge is further compounded in this particular work environment. Given the inherent lack of criticality in data science as a professional practice, as a form of labor, how should we introduce and support critical data literacy? In other words, though perhaps it is obvious upon reflection, it seems important to bluntly state that it is not enough to create the conditions of epistemic agency and scaffold data science skills. Criticality itself must be scaffolded, and the context of work stands as a barrier because its context is, historically and predominately, un-critical.

Following from one of the tenets of CRT, it may be our own liberal bias, as we discussed earlier, that leads us to expect that the data workers would develop a critical perspective to data or the process in the context of work [43]. As academics, developing a personal approach and skepticism are part and parcel of our work. We are rewarded for these qualities. But that is not the case in most work contexts. Certainly, it is not the case for most entry-level positions that have been the experience of these workers to date. For example, if Chase were to arrive at her new job and suddenly invent new approaches to their assigned tasks and question the content of those tasks, she might not keep that job for long. The assumption that Chase would develop these personal and critical perspectives in the context of work, particularly without scaffolding, is unreasonable. It is an expression of our own liberal bias about an inherent "good" of criticality, and our belief in the agency of individuals—a mismatch between theories of epistemic agency in learning and the realities of labor in the work environment and Black people in American society.

Thus, one conclusion that we draw from DataWorks is that we need to critically engage the context of work itself. Developing literacies, perhaps even data literacies, in the context of work tends to belie criticality. The phrase "tends to" is important because the context of work is not fundamentally oppositional to critical perspectives. There is a significant and growing body of scholarship that addresses alternative labor practices, some in relation to data work [58]. If we want to pursue a critical data literacy in the context of work, then the context of work must itself be approached as an environment and process open to critique and refiguration. This conclusion leads us to a consideration of the limits of liberalism, and how the context of work itself may be a trap of interest convergence.

8.2 Implication 2: Contextual support for novices developing critical data perspectives

As outlined in the literature review, understanding the power of data and those who shape the data we use is often not part of basic data literacy [64]. Instead, we need to look toward critical data literacy, and findings from the workshops that support assertions that understanding the context of work is a cornerstone to examine data critically [11, 48]. We argue that providing contextual support in future workshops and programmatic design will allow novices to perceive data differently than when interacting with it without context. For example, having knowledge about the harmful effects of being aggressively surveilled—particularly the aggressive surveillance of selected racial groups—can impact mental health. Knowing this context may lead novices to different conclusions drawn from quantified data produced by surveillance. The contextual knowledge of topics that data addresses, paired with multiple approaches to interpreting data, should strengthen the novices' critical perspectives. Moving novices' perspectives on data from impact through scale to impact through accuracy with context is a first step in critical data literacy.

8.3 Implication 3: Design for the presence of race

Although race was not directly accounted for when constructing the workshops, race had a presence and impact within the topics the data workers selected. Their stories about their lived experiences showed how race has a presence in how novices interact with data to solve various issues. The workers selected topics, which were all closely aligned with issues of race, yet never explicitly addressed race. CRT may explain this lack of acknowledgement of race by the data workers; the ordinariness of race in their everyday lives made it unremarkable. Or perhaps the data workers were concerned about protecting the white researchers and faculty, causing them to self-censor [43]. Alternatively, a critical data literacy perspective also suggests that the "objectivity" baked into data may have left them without room to consider race [11]. When designing for data literacy programs more generally, intentional considerations of race should be part of the process, and attention given to how race and racism may affect the topics learners select and they data the choose to investigate and express those topics.

8.4 Implication 4: Shift from professional to personal data practices

The workshops were intentionally shaped by the interests of the data workers to engage their epistemic agency in their learning processes [39]. By allowing each data worker to exercise agency over their topic and direction of their projects—which is different from the data projects they work on for clients—they were motivated to find, learn, and present data in new ways that pushed their skills. As outlined in the description of the workshops, we are surprised they did not transfer their day-to-day tasks cleaning and preparing data to these personal projects. But the personal projects did serve as a vehicle for the facilitator to teach about why data preparation was important, and to motivate the data workers to learn how to make visualizations on their own. This finding highlights two things. First, the data workers may have been building technical skills with their daily tasks, but they were not building transferable knowledge to other contexts in their lives. Second, their personal contextual interest in data was more motivational than the client pushing them to understand what they can do with data. Both professional data practices and personal data practices influence one another, which in turn strengthens them. Examining how novices learn about professional practices through examples relating to their lived experiences can arguably aid their approach to data work. Similarly, the novices' expertise of their lived experience can be useful in learning about professional data work practices.

9 CONCLUSION

In this paper, we described and analyzed a workshop developed for DataWorks, a work training program that leveraged principles of legitimate peripheral participation to develop a community of practice centered around paid data work. One purpose of DataWorks was to broaden our appreciation of what counted as data work beyond the commonly lauded practices of expert data scientists. Through DataWorks, we hope to draw attention to, better understand, and better support a range of labor practices and environments that involve working with data. In tandem, we sought to broaden participation in data work beyond the race, class, and gender biases that characterize most data science and technology work. This broadening demanded an intersectional approach that engaged with methods and theories that foregrounded issues of race, class, and gender. In our analysis, we have drawn on Critical Race Theory to attune us to these issues. From this, we identified three themes: tensions between epistemic agency, critical data literacies, and the work context; encountering race, racism, and its ordinariness through data work; and the personal is communal and intersectional. We have also offered suggestions for approaching designing data literacy activities, taking into account critical race theory: critically engaging the context of work; providing contextual support for novices developing critical data perspectives; designing for the presence of race; and shifting from professional to personal data practices.

We recognize and embrace the limitations of this work. Our cohort is small, and there are deeper connections yet to be made with Critical Race Theory and other scholars. We consider this paper to be the start of an extended engagement with both the practice of data work and an expanded field of theory. What we have shared in this paper are observations, themes, and implications that form the basis of ongoing work. We also believe these observations, themes, and implications make a meaningful contribution to the human-computer interaction community.

There is an important and growing commitment to intersectionality and (more generally) issues of justice and equity within the HCI community. One swath of that research attends to the context of technology work, and how current work practices and environments are exclusionary. Another swath of that research attends to data literacy, and how data and data practices are exclusionary. Our interest is in how these swaths overlap. As efforts to broaden participation and cultivate data literacy with historically minoritized and underrepresented groups in computing continue to grow, our work in this paper further shows why researchers should embrace critical and reflective perspectives. Such critical perspectives should be intersectional and attend to how race, class, and gender figure into patterns of exclusion and oppression with and through data. Such critical perspectives should also extend beyond data to include the field of computing itself, calling into question our assumptions of what computing is and is not, both as domain of research and profession. The issues that a truly critical reflective practice raise will be uncomfortable and will challenge conventions and expectations; however, we believe such work is imperative. To broaden participation in computing, we must do more than make its dominant culture more accessible; we must transform the very culture of computing.

ACKNOWLEDGMENTS

This material is based upon work supported by the National Science Foundation under Grant No. 1951818. We wish to thank the Georgia Tech Constellations Center for Equity in Computing and PwC Charitable Foundation for their generous financial, advisory, and in-kind support for our work. We also thank Dr. Amanda Meng for her contributions to the DataWorks program, our advisory board members, and our anonymous reviewers for their generous feedback on the paper.

REFERENCES

[1] Amelia Acker and Leanne Bowler. 2018. Youth Data Literacy: Teen Perspectives on Data Created with Social Media and Mobile Devices. In Proceedings of 51st Hawaii International Conference on System Sciences. Hawaii, USA, 1923-32. DOI: 10.24251/HICSS.2018.243.

[2] Whitney Battle-Baptiste and Britt Rusert. 2018. WEB Du Bois's data portraits: Visualizing black America. Chronicle Books.

[3] Derrick A. Bell. 1995. Who's afraid of critical race theory. University of Illinois Law Review, 1995(4), 893–910. Accessed December 15, 2020. https://sph.umd.edu/sites/default/files/files/Bell_Whos%20Afraid%20of%20CRT_1995UIllLRev893.pdf

[4] Ruha Benjamin. 2019. Race after technology: Abolitionist tools for the new Jim Code. Polity.

[5] Geoffrey C. Bowker and Susan Leigh Star. 2000. Sorting things out: Classification and its consequences. MIT press.

[6] Tamara Clegg, Daniel M. Greene, Nate Beard, and Jasmine Brunson. 2020. Data Everyday: Data Literacy Practices in a Division I College Sports Context. In Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems (CHI '20). Association for Computing Machinery, New York, NY, USA, 1-13. https://doi.org/10.1145/3313831.3376153

[7] Computing Research Association. 2019. CRA Taulbee Survey. https://cra.org/resources/taulbee-survey/

[8] Correlation One—Data Science For All / Empowerment. 2020. https://www.correlation-one.com/ds4a-empowerment/?utm_campaign=university_group2

[9] Kimberlé W. Crenshaw. 2017. On intersectionality: Essential writings. The New Press.

[10] Roderic N. Crooks. 2019. Times thirty: Access, maintenance, and justice. Science, Technology, & Human Values, 44, 1, 118-142.

[11] Catherine D'Ignazio and Lauren F. Klein. 2020. Data feminism. MIT Press.

[12] Richard Delgado and Jean Stefancic. 2017. Introduction. In Critical Race Theory: An Introduction (3rd. ed.). New York, NYU Press, 1-18.

[13] Richard Delgado and Jean Stefancic. 2017. Hallmark Critical Race Theory Themes. In Critical Race Theory: An Introduction (3rd. ed.). New York, NYU Press, 19-43.

[14] Betsy DiSalvo, Mark Guzdial, Amy Bruckman, and Tom McKlin. 2014. Saving face while geeking out: Video game testing as a justification for learning computer science. Journal of the Learning Sciences, 23, 3, 272-315.

[15] Betsy James DiSalvo, Mark Guzdail, Tom Mcklin, Charles Meadows, Kenneth Perry, Corey Steward, and Amy Bruckman. 2009. Glitch Game Testers: African American Men Breaking Open the Console. Paper presented at DiGRA 2009, West London, England.

[16] Chris Elsden, David Kirk, Mark Selby, and Chris Speed. 2015. Beyond Personal Informatics: Designing for Experiences with Data. In Proceedings of the 33rd Annual ACM Conference Extended Abstracts on Human Factors in Computing Systems (CHI EA '15). Association for Computing Machinery, New York, NY, USA, 2341-2344. https://doi.org/10.1145/2702613.2702632

[17] Shelly Engelman, Brian Magerko, Tom McKlin, Morgan Miller, Doug Edwards, and Jason Freeman. 2017. "Creativity in Authentic STEAM Education with EarSketch." In Proceedings of the 2017 ACM SIGCSE Technical Symposium on Computer Science Education (SIGCSE '17). Association for Computing Machinery, New York, NY, USA, 183-188. https://doi.org/10.1145/3017680.3017763

[18] Barbara Ericson, Shelly Engelman, Tom McKlin, and Ja'Quan Taylor. 2014. Project rise up 4 CS: increasing the number of black students who pass advanced placement CS A. In Proceedings of the 45th ACM technical symposium on Computer science education (SIGCSE '14). Association for Computing Machinery, New York, NY, USA, 439-444. https://doi.org/10.1145/2538862.2538937

[19] Virginia Eubanks. 2018. Automating inequality: How high-tech tools profile, police, and punish the poor. St. Martin's Press.

[20] Kate Farrell and Judy Robertson. 2019. Interdisciplinary Data Education: Teaching Primary and Secondary learners how to be data citizens. In Proceedings of the 14th Workshop in Primary and Secondary Computing Education (WiPSCE'19). Association for Computing Machinery, New York, NY, 1-2. https://doi.org/10.1145/3361721.3362120

[21] Juan E Gilbert. 2006. Making a case for BPC [Broadening Participation in Computing]. Computer, 39, 3, 83-86.

[22] Andreas Grillenberger. 2016. Bringing the Innovations in Data Management to Secondary CS Education. In Proceedings of the 2016 ACM Conference on International Computing Education Research (ICER '16). Association for Computing Machinery, New York, NY, USA, 267-268. https://doi.org/10.1145/2960310.2960341

[23] Loni Hagen. 2020. Teaching Data Science to Social Sciences and Humanities Students. In The 21st Annual International Conference on Digital Government Research (dg.o '20). Association for Computing Machinery, New York, NY, USA, 363-364. https://doi.org/10.1145/3396956.3396968

[24] David Hankerson, Andrea R. Marshall, Jennifer Booker, Houda El Mimouni, Imani Walker, and Jennifer A. Rode. 2016. Does Technology Have Race? In Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems (CHI EA '16). Association for Computing Machinery, New York, NY, 473-486. https://doi.org/10.1145/2851581.2892578

[25] Alex Hanna, Emily Denton, Andrew Smart, and Jamila Smith-Loud. 2020. Towards a critical race methodology in algorithmic fairness. In Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency (FAT* '20). Association for Computing Machinery, New York, NY, 501-512. https://doi.org/10.1145/3351095.3372826

[26] Donna Haraway. 1988. Situated knowledges: The science question in feminism and the privilege of partial perspective. Feminist studies, 14, 3, 575-599.

[27] Lisa Hardy, Colin Dixon & Sherry his. 2020. From Data Collectors to Data Producers: Shifting Students' Relationship to Data, Journal of the Learning Sciences, 29, 1, 104-126.

[28] Christina Harrington, Sheena Erete, & Anne Marie Piper. 2019. Deconstructing community-based collaborative design: Towards more equitable participatory design engagements. Proceedings of the ACM on Human-Computer Interaction, 3(CSCW), 1-25.

[29] Samantha Hautea, Sayamindu Dasgupta, and Benjamin Mako Hill. 2017. Youth Perspectives on Critical Data Literacies. In Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems (CHI '17). Association for Computing Machinery, New York, NY, USA, 919-930. https://doi.org/10.1145/3025453.3025823 .

[30] Malcolm Haynes, Joshua Groen, Eric Sturzinger, Danny Zhu, Justin Shafer, and Timothy McGee. 2019. Integrating Data Science into a General Education Information Technology Course: An Approach to Developing Data Savvy Undergraduates. In Proceedings of the 20th Annual SIG Conference on Information Technology Education (SIGITE '19). Association for Computing Machinery, New York, NY, USA, 183-188. https://doi.org/10.1145/3349266.3351417 .

[31] Birte Heinemann, Simone Opel, Lea Budde, Carsten Schulte, Daniel Frischemeier, Rolf Biehler, Susanne Podworny, and Thomas Wassong. 2018. Drafting a Data Science Curriculum for Secondary Schools. In Proceedings of the 18th Koli Calling International

Conference on Computing Education Research (Koli Calling '18). Association for Computing Machinery, New York, NY, USA, Article 17, 1-5. https://doi.org/10.1145/3279720.3279737

[32] Jeremiah W. Johnson. 2018. Scaling up: introducing undergraduates to data science early in their college careers. Journal of Computing Sciences in Colleges, 33, 6 (June 2018), 76-85.

[33] Yasmin Kafai, Kristin Searle, Cristobal Martinez, & Bryan Brayboy. 2014. Ethnocomputing with electronic textiles: Culturally responsive open design to broaden participation in computing in American Indian youth and communities. SIGCSE 2014 - Proceedings of the 45th ACM Technical Symposium on Computer Science Education. 241-246. 10.1145/2538862.2538903.

[34] Tibor Koltay. 2015. Data Literacy: In Search of a Name and Identity. Journal of Documentation, 71, 2, 401-15. https://doi.org/10.1108/JD-02-2014-0026.

[35] Sean Kross and Philip J. Guo. 2019. Practitioners Teaching Data Science in Industry and Academia: Expectations, Workflows, and Challenges.

[36] Jean Lave & Etienne Wenger. 1991. Situated learning: Legitimate peripheral participation. Cambridge university press.

[37] V.R. Lee. 2018. Personal Analytics Explorations to Support Youth Learning. In R. Zheng (Ed.), Digital Technologies and Instructional Design for Personalized Learning. IGI Global, 145-163.

[38] Yanni Alexander Loukissas. 2019. All Data Are Local: Thinking Critically in a Data-Driven Society. The MIT Press.

[39] Deborah Lupton. 2018. How do data come to matter? Living and becoming with personal data. Big Data & Society, 5, 2, 1-11.

[40] Emily Miller, Eve Manz, Rosemary Russ, David Stroupe, & Leema Berland. 2018. Addressing the epistemic elephant in the room: Epistemic agency and the next generation science standards. Journal of Research in Science Teaching, 55, 7, 1053-1075.

[41] Ashlyn Munson, Barbara Moskal, Alka Harriger, Tonya Lauriski-Karriker, & Daniel Heersink. 2011. Computing at the high school level: Changing what teachers and students know and believe. Computers & Education, 57, 1836-1849. 10.1016/j.compedu.2011.03.005.

[42] Safiya Umoja Noble. 2018. Algorithms of oppression: How search engines reinforce racism. NYU Press.

[43] Ihudiya Finda Ogbonnaya-Ogburu, Angela D.R. Smith, Alexandra To, and Kentaro Toyama. 2020. Critical Race Theory for HCI. In Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems (CHI '20). Association for Computing Machinery, New York, NY, 1-16. https://doi.org/10.1145/3313831.3376392

[44] Mimi Onuoha. 2016. The Library of Missing Data Sets. Accessed December 20, 2020. http://mimionuoha.com/the-library-of-missing-datasets.

[45] Evan M. Peck, Sofia E. Ayuso, and Omar El-Etr. 2019. Data is Personal: Attitudes and Perceptions of Data Visualization in Rural Pennsylvania. In Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems (CHI '19). Association for Computing Machinery, New York, NY, USA, Paper 244, 1-12. https://doi.org/10.1145/3290605.3300474

[46] Joan Peckham, Lisa L. Harlow, David A. Stuart, Barbara Silver, Helen Mederer, and Peter D. Stephenson. 2007. Broadening participation in computing: issues and challenges. In Proceedings of the 12th annual SIGCSE conference on Innovation and technology in computer science education (ITiCSE '07). Association for Computing Machinery, New York, NY, 9-13. https://doi.org/10.1145/1268784.1268790

[47] Caroline Criado Perez. 2019. Invisible women: Exposing data bias in a world designed for men. Random House.

[48] Firaz Peer and Carl DiSalvo. 2019. Workshops as Boundary Objects for Data Infrastructure Literacy and Design. In Proceedings of the 2019 on Designing Interactive Systems Conference (DIS '19). Association for Computing Machinery, New York, NY, 1363–1375. https://doi.org/10.1145/3322276.3322330

[49] Thomas M. Philip, Maria C. Olivares-Pasillas & Janet Rocha. 2016. Becoming Racially Literate About Data and Data-Literate About Race: Data Visualizations in the Classroom as a Site of Racial-Ideological Micro-Contestations, Cognition and Instruction, 34, 4, 361-388. 10.1080/07370008.2016.1210418

[50] Thomas M. Philip, Sarah Schuler-Brown & Winmar Way. 2013. A framework for learning about big data with mobile technologies for democratic participation: Possibilities, limitations, and unanticipated obstacles. Technology, Knowledge and Learning, 18, 103–120.

[51] Candice Powell, Cynthia Demetriou, Terrell R. Morton & James M. Ellis. 2020. A CRT-Informed Model to Enhance Experiences and Outcomes of Racially Minoritized Students. Journal of Student Affairs Research and Practice. 1-13.

[52] Yolanda A. Rankin and Jakita O. Thomas. 2019. Straighten up and fly right: rethinking intersectionality in HCI research. interactions 26, 6 (November - December 2019), 64-68. https://doi.org/10.1145/3363033.

[53] Gerda Roelvink, Kevin St. Martin, & J. K. Gibson-Graham (Eds.). 2015. Making other worlds possible: Performing diverse economies. U of Minnesota Press.

[54] Daniela K. Rosner, Saba Kawas, Wenqi Li, Nicole Tilly, and Yi-Chen Sung. 2016. Out of Time, Out of Place: Reflections on Design Workshops as a Research Method. In Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing (CSCW '16). Association for Computing Machinery, New York, NY, 1131-1141. https://doi.org/10.1145/2818048.2820021

[55] Jeffrey S. Saltz, Neil I. Dewar, and Robert Heckman. 2018. Key Concepts for a Data Science Ethics Curriculum. In Proceedings of the 49th ACM Technical Symposium on Computer Science Education (SIGCSE '18). Association for Computing Machinery, New York, NY, 952–957. https://doi.org/10.1145/3159450.3159483

[56] Milo Schield. 2004. Information Literacy, Statistical Literacy and Data Literacy. IASSIST quarterly / International Association for Social Science Information Service and Technology. 28, 7-14. 10.29173/iq790.

[57] Ari Schlesinger, W. Keith Edwards, and Rebecca E. Grinter. 2017. Intersectional HCI: Engaging Identity through Gender, Race, and Class. In Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems (CHI '17). Association for Computing

Machinery, New York, NY, 5412-5427. https://doi.org/10.1145/3025453.3025766.

[58] Trebor Scholz. (2016). Platform cooperativism: Challenging the corporate sharing economy. Rosa Luxemburg Foundation.

[59] Ben Rydal Shapiro, Amanda Meng, Cody O'Donnell, Charlotte Lou, Edwin Zhao, Bianca Dankwa, and Andrew Hostetler. 2020. Re-Shape: A Method to Teach Data Ethics for Data Science Education. In Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems (CHI '20). Association for Computing Machinery, New York, NY, USA, 1–13. https://doi.org/10.1145/3313831.3376251

[60] Lucy Suchman. 1993. Do categories have politics? The language/action perspective reconsidered. In Proceedings of the Third European Conference on Computer-Supported Cooperative Work 13–17 September 1993 (ECSCW'93). Milan, Italy, 1-14.

[61] Rong Tang and Watinee Sae-Lim. 2016. Data Science Programs in U.S. Higher Education: An Exploratory Content Analysis of Program Description, Curriculum Structure, and Course Focus. Education for Information, 32, 3, 269-290.

[62] Targeted Infusion Project: Data Science eXtension (DSX): Incorporating data science fundamentals in computing curriculum at Spelman and Morehouse Colleges. 2016. National Science Foundation. https://www.nsf.gov/awardsearch/showAward?AWD_ID=1623362

[63] Alex S. Taylor, Siân Lindley, Tim Regan, David Sweeney, Vasillis Vlachokyriakos, Lillie Grainger, and Jessica Lingel. 2015. Data-in-Place: Thinking through the Relations Between Data and Community. In Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems (CHI '15). Association for Computing Machinery, New York, NY, 2863-2872. https://doi.org/10.1145/2702123.2702558

[64] Alan Freihof Tygel and Rosana Kirsch. 2016. Contributions of Paulo Freire for a Critical Data Literacy: a Popular Education Approach. The Journal of Community Informatics, 12, 3, 108-121. http://www.ci-journal.net/index.php/ciej/article/view/1296

[65] Moshe Y. Vardi, Tim Finin, and Tom Henderson. 2003. 2001-2002 Taulbee Survey: Survey Results Show Better Balance in Supply and Demand. Computing Research Association. https://cra.org/wp-content/uploads/2015/01/02.pdf

[66] Michelle Wilkerson,Kathryn Lanouette, Rebecca Shareff, Tim Erickson, Nicole Bulalacao, Joan Heller, Natalya Clair, William Finzer, and Frieda Reichsman. 2018. Data Moves: Restructuring data for inquiry in a simulation and data analysis environment. In Proceedings of the 13th International Conference of the Learning Sciences (ICLS) 2018, Volume 2, 1383-1384. London, UK: International Society of the Learning Sciences.

[67] Andrew B. Williams and Whitney O'Banner. 2008. SpelBots: Using Autonomous Robotics to Inspire Women to Participate in Computer Science. Paper presented at the AAAI Spring Symposium: Using AI to Motivate Greater Participation in Computer Science, 108-109.

[68] Annika Wolff, Daniel Gooch, Jose Cavero Montaner, Umar Rashid, and Gerd Kortuem. 2016. Creating an Understanding of Data Literacy for a Data-driven Society. The Journal of Community Informatics, 12, 3 (Aug. 2016), 9-26.