

Transcript:

Databite No. 133: Fellows Talks

Michele Gilman, Anita Say Chan, Dan Bouk

Recorded: 06/17/20

**Sareeta Amrute:**

Hello everyone, welcome to Databite #133 featuring Fellows Talks by Michele Gilman, Anita Say Chan, and Dan Bouk. My name is Sareeta Amrute, Director of Research here at Data & Society. I will be your host for tonight, supported by my team behind the curtain: CJ, Rigo, Natalie, and Angie. For those of you who don't know us yet, Data & Society is an independent research institute studying the social implications of data and automation. We produce original research and convene multidisciplinary thinkers to challenge the power and purpose of technology in society. You can learn more about us through our website at [datasociety \[dot\] net](http://datasociety.net).

I ask you to join me in acknowledging the Lenape community, and the exclusions and erasures of many Indigenous peoples, not just the Lenape community, on whose land Data & Society is located, in what we now refer to as New York City. This acknowledgement demonstrates a commitment to beginning the process of working to dismantle the ongoing legacies of settler colonialism. Acknowledgments invite us to ask, what does it mean to live in a post & neocolonial world? What did it take for us to get here? And how can we be accountable to our part in history?

I also ask you to join me in recognizing the activists protesting in the streets, in city council meetings, in their workplaces and around the kitchen table to support Black liberation and end racism in all its forms. I'd now like to turn it over to our featured presenters, 2019-2020 Fellows: Michele Gilman, Anita Say Chan, and Dan Bouk. Faculty Fellows help ensure that new connections and perspectives deepen and expand our community's understanding of the challenges and opportunities society faces in a data-centric world. I'm especially looking forward to learning from each lightning talk as the fellows contextualize their work in relation to historic and current racial justice movements. We'll then have time for a Q and A. First up, Michele Gilman with her talk, "The Class Differential in Data Privacy." Take it away, Michele.

**Michele Gilman:**

Thanks, Sareeta. Hi, I'm Michele Gilman. I'm a law professor who directs a law clinic in Baltimore, Maryland. That means I teach and supervise law students who provide pro bono or free legal representation to people who cannot afford an attorney. When I started practicing poverty law over 20 years ago, it was immediately clear that my clients had far less privacy than my suburban upper middle-class neighbors. My clients not only lived in densely packed neighborhoods under the constant eye of the police, but they had to turn over extremely personal and sensitive information in order to access government assistance. Over the last two decades, I've observed how emerging technologies have been deployed against low income people in ways that add scope, scale and speed to the privacy deprivations. They have long suffered. These deprivations are fronts to dignity and liberty, and also barriers to economic equality and racial justice.

These dynamics are heightened in the current moment. I'm going to focus tonight on the pandemic and the police response to the Black Lives Matter protests. But I'm not talking about issues that you've seen on CNN, Instagram, or Twitter. I'm talking instead about processes that are largely invisible to the general public, but that have real world impacts. My prediction is that the pandemic and the protests will result in digital profiling that will in turn harm marginalized communities and extend the long tail of these crises. First, let's consider the pandemic. We are in the grip of a dire public health and economic crisis being suffered most acutely by Black, brown, and indigenous people of color. Even when the public health emergency recedes, these groups will struggle to regain their financial footing, in part as a result of digital profiling. So, what do I mean by digital profiling?

There are vast networks of data extraction that gather our personal data. Every time we turn on our computer, pick up a smartphone, or engage with a government agency. These interactions generate millions of data points used to profile us. This is not unique to marginalized people, but the outcomes of this data scraping are different for the poor. This is because digital profiling is used not only to target people with ads, but also to screen people: by landlords, employers, lenders, and universities. This is problematic in the best of economic times, but it is particularly fraught. Now, as of today, over 41 million people have lost their jobs. In most jurisdictions, there are temporary moratoriums on evictions. However, these moratoriums will expire soon and the accumulated rent bills will become due. Prior to the pandemic, one in four renters were already spending more than half their income on rent. The housing unaffordability crisis will deepen as people remain out of work and unemployment insurance benefits, timeout. That's how housing experts predict an avalanche of evictions if the government does not cancel rent or provide tenants with rental assistance. Having an eviction on your record can make it hard to find future housing.

That's the digital profiling dilemma. In selecting tenants, most landlords will search court records, or purchase a tenant screening report from a data mining company. Landlords disfavor applicants with prior eviction cases on their records even though these records are often inaccurate, confused tenants with similar names, and can fail to show the outcome of a case—meaning that even if a tenant ultimately won the case, they have a permanent negative mark on their record. Consider also that people who catch coronavirus may be facing large medical bills. The estimated cost of COVID-19 care for patients who are hospitalized ranges between 20 and 70 thousand dollars. Even before the pandemic, medical debt was the primary driver for two thirds of personal bankruptcies.

In addition, across the country, people are accruing debt in a scramble to pay for food, to cover utility bills, and to maintain internet access for schools and work. The medical and other debts from coronavirus will feed into the big data system. In turn these debts will lower people's credit scores and appear in other digital screening profiles. As a result, millions of people will be denied loans or have to pay higher interest rates for them. They may struggle to meet basic needs. They may even find it difficult to find employment because most employers use digital background checks in order to select employees.

Now, Let's turn to the protests. Here my concern is that the police response to the protests will serve as a digital barrier to future opportunities for many protestors. Since the protests began, more than 10,000 people have been arrested. Whether or not these protestors end up with a conviction, and even if they are acquitted, these arrests will appear in their court records and be scraped into their digital profiles. Police departments also release mugshots onto the internet, making them easy to find with a quick Google search. The collateral consequences of having an easily searchable criminal record are serious. A criminal record operates as a barrier for people to get loans, find jobs, enroll in higher education, qualify for public benefits, and obtain professional licenses. In some cases, protestors will be able to take advantage of state laws to expunge their criminal records—meaning taking legal steps to delete them from state databases. But this process can be hard to navigate. And once these public records have been scraped, the data has already been captured and is almost impossible to remove from the hundreds of profiling companies that profit from selling the data. Even people not arrested on the scene might face criminal charges in the future due to extensive surveillance of the protests.

Law enforcement has been monitoring protestors, text messages, and locations. Other surveillance tools include: license, plate readers, social media, monitoring, police body cameras, drones, and facial recognition technology. This is nothing less than the criminalization of constitutionally protected conduct, the right a free expression. I should

note here that data itself is not the problem. The problem is when data is deployed to oppress and stigmatize people. In the right hands, access to data can also serve justice. As we saw last week with the repeal in New York state of 50-A, a statute which long protected police disciplinary records from public view. This recent law reform is a reminder that controlling data is about power.

Clearly the pandemic and the protests are amplifying longstanding dynamics in which low-income people and people of color face greater harms from state surveillance and data extraction technologies. What does law have to say about this? The short answer is that privacy law is generally more concerned about protecting the interests of the powerful than the poor. I will highlight briefly three limitations in American privacy law. First in the United States, we do not have a comprehensive data privacy statute to protect our personal data. Rather our legal regime hinges primarily on the concept of notice and consent. This puts the onus on you to protect your personal data, yet my guess is that you don't read the convoluted, take it or leave it terms offered as a condition for accessing websites and apps. Of course not. You can't negotiate the terms. Even if you want to. This individualistic notion of privacy gives companies and governments free rein to use your data as they see fit. It ignores the ways that data is used to collectively sort, score, and segment groups of people. It also puts an undue burden on already strained communities. Not surprisingly a Data & Society survey found that low-income people have less confidence and greater concerns about their ability to protect their digital privacy and security. Second, courts protect privacy interests that are deemed reasonable. However, the measure of reasonableness is drawn from a white, middle class, and male norm. Interests outside that norm are not reasonable and thus not protected. That's why the government can legally search the homes of welfare recipients, but has never sought to conduct home visits for people claiming a mortgage tax deduction, which is just another form of a government benefit. The reasonableness standard also underlies the premise that information shared publicly or with third parties is not protected from government surveillance.

Modern technology is straining this doctrine and we can expect to see many legal challenges to the surveillance of the protests in the months to come. Third, discrimination law can play a role in limiting some of the disparate aspects of technology. However, as scholars have explained, civil rights laws written for an analog world are a tough fit for the digital world where computer algorithms now perform tasks that were previously performed by people. Discrimination law is also a limited remedy because poor people are not a protected class. They can be discriminated against with impunity. Finally, discrimination law is about generating an equal playing field, but it does nothing to give people the equipment, uniforms, or cleats needed to play the game. That requires an equity

approach grounded in social and economic rights. In conclusion, the protests have highlighted so many needed actions throughout our society to achieve racial justice. To this list we can add two more: abolish the data extraction economy, and abolish state surveillance.

Thank you.

**Sareeta Amrute:**

So, I'm going to introduce Anita Say Chan, Anita will be speaking to us about feminist data futures and relational infrastructures. Please take it away Anita.

**Anita Say Chan:**

Thanks, Sareeta. Let me introduce myself. I am an information science professor at the University of Illinois at Urbana Champagne, and I'll be speaking tonight about the relational infrastructures of protest. Since, of course, protest has been on a lot of our not minds, and perhaps unlike any other time since the 1960s in the U.S., protest has been propelled into the national consciousness. In mainstream liberal news sources, in particular, we're likely to have heard about it framed as an outcry of shared grief, frustration, or defiance. But in my comments today, I want to suggest that protest serves another key function that's largely been overlooked, and that is for it to serve as a key means for populations, namely vulnerable populations, to publicly assess and speak back to knowledge practice.

Protest in other words, it's not just a means to mobilize around a shared sentiment. I want to suggest it's also a means for marginalized populations to use the resource of public space, to call for new forms of knowledge work, and to do this in three ways. Firstly, by explicitly exposing the insufficiency of how dominant forms of common knowledge come to be not just defined, but stabilized in mainstream institutions and infrastructures. Secondly, by producing evidence on a narrowness of what it is that's presumed to be given knowns when it comes to vulnerable populations, and thirdly, by suggesting new infrastructures that can test and prototype empirical conditions for building other possible worlds. Protests may be read, in other words, as mobilizations around what feminist science studies scholar, Donna Haraway called "situated knowledges" that insist upon the need for other perspectives to offer better accounts of the world in order to "live in critical relationship to practices of domination."

That's arguably something that projects like Data for Black Lives founded by Yeshimabeit Milner underscore, and it's something too that's echoed by the NiUnaMenos movement that's mobilized since 2015 to combat gender-based violence in Latin America -- that's been my research focus. Their efforts have called attention to femicide as hate crimes against women -- and the striking absence of data around it as a matter of shared regional concern. Two years before the U.S. Me Too Movement exploded, this is what streets of Buenos Aires looked like. And this was too, since 2015 NiUnaMenos (not one less – in English) has drawn together some of the largest demonstrations in the region in Argentina. These began with a March after the murder of 14-year-old Chiara Paez who was found buried under her boyfriend's home, beaten to death, and just a few weeks pregnant.

Parallel were launched all across the region. In Peru, more than 50,000 filled highways in Lima. In Chile, more than 80,000 joined. And marches since 2016, have not only shut down streets in Santiago, but also closed university campuses, with more than 25 closing in 2018, including high schools --where protesters called not only for accountability around harassment cases, but also decried the exclusion of women in leadership, faculty, and in assigned syllabi. And most recently mobilizations in Mexico this spring captured widespread attention when more than 100,000 marched the Palacio Nacional where parallel protests were held in cities from Merida to Tijuana. But while protesters, that is while national media tended to emphasize the unique and local cases as catalysts for the protest, protestors organizers themselves pointed instead to global feminist researchers who have been pointing to data as a critical resource for their own efforts. And NGOs like Femicide Watch, for instance, estimate the number of women intentionally killed in Femicide cases (87,000 in 2017) was near equal to the victims killed in armed conflicts worldwide (89,000 that year).

With the key difference here being that the vast majority of Femicide cases are killed by people that they know. The underscore still is the widespread inaction of states would few collecting any data on Femicide and even fewer compiling official countrywide data. This entrenched invisibilization of Femicide brings to mind U.S. feminist data science scholars, who could critique the long silence of bodies missing from the archives states and corporations do deem worthy of producing. Organizers have pushed back though, and are not only naming state inaction as complicity and what they called the uncounted epidemic of Femicide, but have begun to build their own infrastructures to register data. In Argentina, citizens launched the first national index on gender violence, designing a 186-question survey for women and transgender women nationwide. Crowdsourcing projects have also begun by researchers elsewhere like Anita Lucchesi, who launched an archive for murdered and missing indigenous women in Canada and the US.

The WomenCount project was also started by Dawn Wilcox to draw together data in the US, with a network of volunteers, researching cases, state by state. But the Argentinian efforts, in particular, have been uniquely recognized for creating some of the largest global data sets on Femicide. Run entirely by volunteers, the networks first survey received over 60,000 responses with results showing over 97% of respondees had suffered some kind of gender violence. And despite being approached by institutions like the state to partner for future research, the network remains independent. As they've put it violence against women is indeed domestic violence, but it's also in the violence of the state, market, and capitalist property relations. It's in the violence that results from discriminatory policies against LGBTQ people, for mass incarceration, criminalizing migratory movements, and from abortion bans. Data worked for them, this drives to empower grassroots infrastructures and local accountability.

And among the new infrastructure, collectives have grown include those fostered in Vera Cruz, Mexico, or new hotlines for domestic violence for finding refuge, networks to provide accessible legal aid and a system of food donations for families hard hit by quarantine have emerged. Existing infrastructures, like the streets of Mexico City have also been repurposed and covered with victim's names to publicly evidence their loss. An online letter to the president and a campaign, #Nosotras Tenemos Otros Datos (we have other data), was also launched to highlight records like the number of domestic violence calls during quarantine at 155 calls per hour, all to respond to the president's dismissal of such reports as disinformation. Of course, science and technology scholars amongst us will be aware of the way our attentions have been turned to the hidden power of infrastructures to stabilize dominant framings of reality. the so called "epistemic infrastructures" that Michelle Murphy writes about, for instance, the buildings, bureaucracies, and technologies, that create dense numbers and data about previously unrelated objects, like population and the economy, came to turn life into something newly calculable, and they transformed them - what were once experimental research methods to quantify life into pervasive 20th century givens.

Here though, we consider not just the epistemic infrastructures that stabilized dominant givens, but the overlooked work of what I'd call the relational infrastructures of data justice networks to remake shared imaginaries and spaces of public life. These embodied forms generate other means, and materialities for intersectional knowing and connective accountability to remake worlds in need of transformation and redress. But I'd be remiss to suggest that relational infrastructures are new. And I want to close by quickly gesturing to longer history of feminist data and the work of sites like the Hull House project founded in 1889 in Chicago is a multiethnic 19th ward by Jane Addams and her partner, Ellen Gates

Starr. Many might remember its leading role in the US "settlement house" movement, and its unique success in advancing such key legal reforms in the US as the eight-hour workday, minimum wage, and outlawing child labor.

But it was also known for prototyping, various approaches to social science and community research that championed what amounted to a new ethical paradigm for understanding poverty. One that emphasized collective responsibility, change, and social justice over dominant ideas of the day -that's how poverty as a result of individual failing or even biology. Part of this involved Hull House's infrastructural design that blended a community center with an educational campus in Chicago's West side, where local families, large immigrants and ethnic minorities could access classrooms and libraries for free courses, kindergarten and daycare, theater and art studios, gyms and athletic programs, and coffee houses and meeting rooms. Such spaces were also used to foster though novel approaches to research by residents, and to document data on local labor conditions. Hull House maps and papers published in 1895 was an exemplar and quickly placed it at the forefront of new social science techniques that later established fields for urban sociology to social work.

Chapters on the sweatshop system and child labor written by ten authors, eight of which were women, and three who identified as US immigrants, thus featured the use of surveys and statistical data alongside direct testimony from 19th ward residents to newly capture how gender, ethnicity, and age impacted the lives of working residents. Color-coded wage maps were also paired with essays and showed how statistical and visual techniques could be blended in compelling ways with qualitative narrative and be used in this case to evidence a scale of wage exploitation in the 19th ward, where monthly wages adjusted to today's values would be just \$500 to \$2,000 a month for the wages of an entire household that then often included working children.

Hull House is a relevant reminder then of the longer legacy of research infrastructures that use data as the basis for local organizing and that engaged amateurs as well as professional social scientists. While struggles over the terms of research on poverty, however, would begin to shift towards a more detached technical paradigm decades later, those credited with founding US sociology decades later at the University of Chicago would draw amply from the techniques innovated by Hull House and social reformers, but would train their students to detach themselves from local communities.

By the late 1920s, this shift reinforced two growing gender divides where Hull House researchers would be framed as social workers rather than social scientists, and where research professions increasingly excluded groups most vulnerable to poverty themselves.

That is minorities, women, and the working classes. And while most Hull House alumni never became household names, it was still obviously generative with many going on to serve key leading roles in history making organizations from the NAACP to Legal Aid, to National Consumers League and the Progressive Party. To close then, we might say that in contrast to the empirical infrastructures around objective social science that came to take hold, that the relational infrastructures of Hull House and data justice protesters today, aim for new means to foster engage research. I would strive for these methods as key to transforming material world and future knowledge practice alike. Thanks. Now let me close and pass on the mic to my colleague, Dan Bouk.

**Dan Bouk:**

Thank you all for coming out. I want to thank also CJ, Rigo, Natalie and Angie, and everyone else behind the scenes and making this all possible for us. Today I'd like to talk about reading data, reading data closely, reading data deeply. My text is the 1940 U.S. Census. Early 20th century censuses were old school, big data. In 1940, we see billions of pieces of personal data, about 130 million Americans from all walks of life. It's also a remarkably open and transparent kind of data. After 72 years, all responses by individuals are made to be non-confidential. And so, as a result, we see everything that people enter as data. And we also can see exactly the methods by which they were produced. So, we often think of data making as a linear process, like an assembly line. And that's where I'd like to begin, with a process by which a person that became census data in 1940.

We'll begin with the form. My colleague, Caitlin Rosenthal, the author of "Accounting for Slavery," has a great maxim that she uses to talk about this. She talks about "reading the frame, not the data." We read the frame and not the data. Here we're looking around here. We're not looking in here. We look at the categories, the columns, the labels, and that gives us a sense of the values and ideas of those who designed that form, the set of the society that produced it. If we read the frame here, we find that race is fundamental and closely prescribed. These are the official racial categories for 1940. They're determined in Washington, D.C. We see them listed here, white, Negro, Indian, Chinese, Japanese, Filipino, Hindu, Korean, a possibility for other. These categories have long histories. They vary by time and place. I would point to the work of Melissa Nobles, books by Melissa Nobles and Debra Thompson, for examples of how it is that those vary across time and how they move through nations.

Next, after the form, we get to the data. Forms go out, they're carried by enumerators like this one, off to be questioned for the questioning of the nation, where all the answers are written down. This is a photo from a test census in 1939 in Indiana. It highlights the

centrality of whiteness for the census, first in selecting a predominantly white county for the test. And then we look at this publicity shot, it is all the white folks and the only white people in any of these publicity shots. Here on the right we see two more from the trial census over here, this is a staged photograph of someone enumerating the heartthrob Tyrone Power. The encounter of the enumerator and person is filled with mystery. People assert their identities, and enumerators ascribe identities onto people. It's hard to know whose idea, whose assertions went out.

All we know is that the enumerators write something down. Langston Hughes, dramatized that mystery in looking at the struggle of a black woman to be seen as she chose in a poem called "Madam and the Census Man." A poem that I wish I could rehearse for you right now, but I don't have time. But I encourage you to look at that. Here is how Hughes and his household recorded in the census in 1940. In this case taken down on April 17 by an African American enumerator in Harlem named Thomas W. Mosley. Step three, the completed census is shipped to Washington, D.C., to be processed. All of those forms are put in leather binders, shipped off, and the first step is editing. Here we see what is labeled here as the Negro section, a segregated statistical workforce within a segregated federal government, within a segregated city in Washington, D.C. For more on the way that surveillance was racialized in 1948 and the 1940 census, I point you to the work of J.D. Schnepf. It might've been a one of these editors, one of these workers, who edited this column of this census sheet. So, what we see here, these people are entered, any "Neg" stands for Negro, "W" for white, and then here what's written as "Mexican," written out fully is "Mexican." Mexican was a category that wasn't supposed to exist in the 1930 Census. It had been introduced in 1930 when Mexican American leaders became convinced of the wages of whiteness in America, and recognized they preferred not to be labeled as an other race. And so those Mexican American leaders were able to work with the Mexican national government to have the Mexican label removed and to be reclassified as white. There's more of this in the paper by Brian Gratton and Emily Merchant. Now, for some reason, an enumerator in this case wrote Mexican.

This in fact happened thousands of times in 1940. Then, an editor comes through and has erased all of those "Mexicans" and writes instead that number one, one being the code for whites, again, centrality of whiteness in this system. So, these folks become, they were first labeled Mexican and then become white, whether they knew it or not, believed it or not in all of these cases. Next, those edited responses are sent to be punched into a punch card like this one. There's plenty of a frame here to read in a bunch of cards and we can see the way that race is structured into it, sex is structured into it. The power of patriarchy is structured

into the card itself. So, for instance, in the labeling the possibilities for relationship within the household, there is head and on the next entry is WIF or wife.

So, the assumption again of a male head of household, and then that the spouse must necessarily be labeled “wife” built into the punch card itself in 1940. Finally, step four, punch cards are sent to be tabulated. The feminized work of tabulation is often obscured as in this description of a woman worker and an electronic tabulator that neglects to mention the worker at all. Here Mar Hicks' "Programmed Inequality," the programmed inequality has shaped my thinking. Tabulated results are distributed and used by private and public actors to make other visualizations like this one, or like this, a study showing that about one-eighth of African American population was probably missed. It was probably not counted in the 1940 census. Now in thinking about this process, I see a missing axis. When we talk about data as if it were purely a linear process, we fail to read the depth of that data.

I'm inspired here by Karen Barad's book, "Meeting the Universe Halfway," which I was lucky to read this January with Data & Society colleagues in a seminar that was led by Janet Vertesi. And Barad in this book of feminists' SDS is talking about quantum mechanics, but it has, I think something that at the very least at a metaphorical stance can be useful for thinking about data. Data is sometimes very tightly controlled. It's in Washington at the moment when the form is constructed, it's in Washington at the moment that the editing brings unruly responses into order. These are the nodes in the waveform when it's later edited. At other moments, data can be mysterious, unsettled the product of negotiation, and as various as there are ideas of the millions who answer questions—or as many as the different varieties of charts and arguments that can be you printed from data after it has been tabulated and released. Data, in other words, isn't the result of a linear process. It has depth. It is a wave and particle both at once. We only observe it when it is in one stage or in the other, depending on how we look at it. The upshot is to really understand that data and how it works in the world, we have to read data as deeply as possible, attentive to attempt to control and constrain it, and also its inherent messiness and bubbling variety.

Thank you very much.

**Sareeta Amrute:**

I'm going to take a few moments to ask a few questions and then draw some questions from the Q and A. So those of you in the audience who have questions, please do populate the Q and A. What I was thinking about as I was listening is that all of your work in some way, tackles the topic of race, data and oppression, but from very different vantage points. And I

would love to hear from you how the current movement for Black liberation, the movement for Black lives has affected the way you approach your work. What parts of your findings do you specifically want to highlight in light of the movement for Black lives and the call to defund police departments?

**Michele Gilman:**

All right. I'll jump in. Well for me, the current movement is really an inspirational model for effectuating social change. And so, I'm very interested in learning from this movement, listening to this movement and to work with grassroots activists, to advocate for changes that would advance justice for Black communities in Baltimore, along with my students as well. In terms of my work, one thing I'm hoping to highlight is how you can't cabin off the criminal justice system as something separate, it's deeply intertwined with people's ability to interact with all aspects of society. Entanglements with the criminal justice system impact where people can live, where they can work, whether they can vote and, and just so much more. So, once we defund the police we need to make sure that digital markers of difference don't limit people's ability to flourish and live up to their potential.

**Anita Say Chan:**

Yeah. And just to build on some of Michele's comments. So, thank you Sareeta for that question and for that really sort of timely and gorgeous question. So that, this is not just the protest or sort of a moment, but they're also a diagnostic on the kinds of problems, the larger problems within the system. This is about system reform. And we can think about the police and criminal justice system as one system. But what I think has been so illuminating and really inspiring for many of us is also the kind of diagnostic of the degree to which many, many systems, not just limited to police and criminal justice, but many systems have been in long and deep need of reform. And that we've now been all being brought to the kind of moment of reckoning with the evidence of just how much work we allowed to not get done sort of collectively.

And now how much work is before us, to be challenged, to dare, to imagine large institutions differently from criminal justice, to academia, to research institutions, to, I mean, you name, different kinds of workplaces, hopefully the tech industry as well. But really being challenged by protesters to really get to imagine and to know our futures slightly differently. This moment, I think is also a nod to the kind of breadth of the fellows' research and my colleagues, Michele and Dan, as well this year to the kind of work of how and why research is so instructive and also why history is so instructive. And this moment I think has brought many of us to start looking through past archives. And for me to look at

past archives of the intersection between protest and research again, not just think about protesters as sort of just protesters, right?

Sort of rabble-rousers on the street or these mobilizers on the street, but really to take seriously the deep layers of research and knowledge work that go into protest practices and the layers of kind of protest infrastructures. That's not just work on the street, but also work behind the scenes. So Sareeta, you know, you and I have been having these conversations as well about the sort of founding of the NAACP and looking through some of these old documents that heavily evidence not just the kind of allyships that were built across different kinds of justice networks across the country, internationally as well, that represented different walks of life, different races, different professions, different genders, et cetera. But that also demonstrated deep statistical work. Some of the first protests of the NAACP that were against exactly the violence on Black bodies, lynchings that were happening at the turn of the century were exactly ones that were, I mean, you can look at the kind of map of protest posters and they're amazing.

I mean, they excite a huge range of different statistics about the participation and the contributions of Black labor to different sectors: to agriculture, to industry, to education, et. cetera. But and then also of course the number of Black lives being lost over the decades without any action, the criminal justice system, lost to lynchings. And so again, you know, this sort of moment is bringing back again to the long history of evidence of protest work as research work. So, I've been really grateful for that.

**Sareeta Amrute:**

Which brings to mind Dan's point about the messiness of data and its interpretability, right.

**Dan Bouk:**

Yeah. Yeah. I'll jump in just really quickly. I am, I loved it the way Anita's talk began because it's been so powerful seeing just again in real life, how research agenda is set up on the streets and the kind of intellectual work that happens in protest. It is just invigorating. I wanted to say one thing about the census, which is interesting, right? Melissa Nobles and her kind of periodization of the census makes clear that throughout most of the 19th century the census acts as a tool primarily for creating scientific racism or justifying racial hierarchies. From 1970 on, it becomes a tool mobilized and used primarily by the civil rights movement, the racial categories. So, the same categories, change a little bit, but can be mobilized and used by different groups for different purposes all the time.

And they have, it's interesting to see that flexibility and how much that political power and force can then shape them. I was at a rally on 168th Street the other day and heard a Harlem pastor talking to the people behind me saying, yeah, we're out for Black Lives Matter, and we're out organizing for the census because when you vote, if you weren't counted in the census, then your votes don't count for your representatives. And when we talked about defunding the police, it's so interesting, right? Because it's about pulling funding out of the police, but then being able to redirect that funding to health, education, and other social services. When you need those funds, it would come in through, there are so many of those funds are directly funneled via automated systems that are dependent on census data.

But that said, I love also what Anita's showing, right? These ways in which you can imagine other kinds of data making, and talk about census sometimes as democracy's data. And if you think of the kind of democracy that we have, if this is its data, it's a kind of lame democracy, or it could be a much better. And I would love to think about the kind of work we can all do together, imagining what a democracy's data should look like in a more equitable democracy. One that doesn't center normative, white heterosexual, male lives, and instead is much more like an imaginatively creative.

**Sareeta Amrute:**

Thank you. I'm going to pick up on that phrase, "lame democracy" and its, and its attendant data is to ask each of you how you're thinking about bringing your work forward. So how, for each of your projects, are you thinking about the move, the necessary move between from "lame democracy" to "thick democracy," "thick data"?

**Anita Say Chan:**

I can jump in and try a quick response at this. And so, a lot of it I think has to do with kind of gesturing towards and highlighting the amount of work that's happening, the kind of research work, intellectual labor that's happening, not in traditional academic spaces and scholarly spaces. So, you know, bringing it into the street and into the, into publics and into the spaces of community engagement that again are not the ivory tower. And I think that challenges a lot of academics. But also, hopefully it's an exciting invitation because I think we've got so many allies and so much new energy that's coming from different kinds of research networks that are so demonstrative of data's democracy, democracy's data, or trying to reinvent new practices. So, in that sense, for me, it's a little bit of a continuation of what some of the work I've been up to in terms of collaborating with free software networks in Latin America and different open technology networks in Latin America had been about.

But, so this summer I'll be working on putting together a workshop for the Allied Media Conference and hope to see many of us in our space. I know I'll see many of us in the Data & Society space and hopefully we'll also get to see many and get to meet with some of the folks and our audience as well for that. So, there'll be a feminist data justice panel that we'll be organizing. A network of us who were part of a Feminist Data Manifest-no document that I encourage folks to look up as well as Lisa Nakamura and her work from visual media studies programming. You mentioned also Catherine D'Ignazio and Lauren Klein who are part of the feminist data and data feminism work that I think hopefully many in our audience are also familiar with. The work of #unsettle and collaborations with the #unsettle cluster that's also been seeded really terrifically at Data & Society with Sareeta Amrute and Rigo Lara, and Siera Dissmore as behind the scenes, but also in front of the scenes as well, leading the energies. And then lastly doing some work, we've been doing a lot of work this summer. We've been doing a lot of work in the past couple of years, but including over the summer with an effort that we've launched here at the U of I at the ISchool, the founding of Community Data Clinic. And so, we've been collaborating and launching some, fostering some growing research projects with some of the local civic organizations and community organizations that have been up to crisis response post-COVID. But also, in light of attempts to really address and redress the work of not just police reform, but trying to reimagine what security could look like from a community standpoint.

**Sareeta Amrute:**

Thank you. That's absolutely brilliant. I really appreciate those answers. I'm going to turn to some of our lovely questions from the Q & A, they're really very generative. I'll try to bunch a few together. The first set of questions is for Michele and it revolves around the question of protection. So, folks on the chat are asking what measures could be taken to further data protection for marginalized communities? And also, for those who are protesting, what are the possibilities or resources that you could recommend to those who wish to demonstrate, but might be concerned about their data collection associated with exercising their constitutionally protected rights in 2020?

**Michele Gilman:**

Sure. So, in terms of keeping your data secure, when you're protesting, that is an important question. There are some very good guides online that provide more detailed technical instructions than I can do here that would help people protect particularly all the information on their smartphones, if they're out at protests and keep that information out of the hands of police if those interactions happen. So, there are some good quality guides online, and I would direct people towards those. In terms of what can we do for legal reform? One thing that's been very exciting about my fellowship and a real honor has been,

I've been able to talk to advocates and activists all over the country who also do legal services work. And as legal services lawyers, we are people who help our clients gain access to basic needs of life: housing, food, medical care, education.

And we are seeing, as I spoke earlier, about all the ways that technology can impact our clients, and they're starting to be very exciting activism about these issues. So, in some jurisdictions, cities have outlawed facial recognition technology. But that shouldn't be a city by city determination. We should have a national law in that regard. There are other scattered laws that protect people's biometric information that provide certain workplace protections. But they're very, I would say scattered and isolated. So what we really need is advocacy on a national level and a national commitment to both recognizing people's individual privacy rights, but also our interests collectively in securing privacy for all of us.

**Sareeta Amrute:**

Thank you. Then we have some questions about how to read data, what happens when you read data? And the question is when, say, we read a nerdy and seemingly revelatory *Atlantic* story, larded with data and statistics, particularly demographic data. What should we be thinking about? How does Anita, your approach to relational infrastructures, and Dan, your approach to data, wave particle duality? How does that help inform our daily reading practices of our data saturated public culture?

**Dan Bouk:**

I'll do briefly. I know I really want to hear Anita's response. On Twitter, my thing is that I study things shrouded by cloaks of boringness. And so one of my problems with data often is that the sheen of objectivity or its numerical charisma makes it difficult to engage with it, or to feel like one can engage with it, but the kind of interpretive frames or the kind of intelligence that one deals with other forms of rhetoric. And so like at one level, when you're reading *The Atlantic* article, right, you don't know where any of these numbers are coming from and in many cases, so you treat them the same way you treat most of the other facts in *The Atlantic* article. But the key thing I want for any readers or many folks to think like, alright, well, I can engage with this data and these categories the same way I would engage with everything else in this, that data should be not understood as a separate category, but one part of culture.

And when we read it, then it becomes important to read it with the same kinds of critical lenses that we apply to all other parts of culture as well. I think that's the other important thing that reading data then forces us and invites us to think about race, class, gender, sexuality, and all the other kinds of ways in which we understand that on a very basic level,

when you read that, you should probably think when you see demographic data, things like demographic data is always, especially probably now too, it always is about placing people in boxes that they often have no choice whatsoever, whether they get placed in those boxes in the first place. And then whenever you see anything that looks like a very precise number, you should assume that there's a very large error bar.

**Anita Say Chan:**

Yeah. I want to echo Dan's beautiful response. And just to say that, you know, when you see this kind of gorgeous, beautiful compilation of statistics and datasets, you have to sort of pause and really push yourself outside, or beyond this kind of the instinct, that we have to just sort of believe and trust in the numbers. It's like, it's so natural when you see something that large, and that impressively compiled, and it's beautifully formatted, and it's all cleaned up, and all the categories are so stable. Look at those columns, and those lines, and those rules, they're so gorgeous. And then sometimes they're coming from these super authoritative spaces, right. Or places like a Facebook or Amazon that will say these are terabytes and terabytes and terabytes of information, it's all compiled gorgeously. How could it not be objective?

And to really step outside of that kind of notion of like just rote belief or trust, and that kind of sense and argument of objectivity, and push back and just remind yourselves that these were created by real humans, sometimes real humans in institutions that don't look like the rest of the world and who are embedded, and not just relationships, but deeply political and culturally shaped spaces and places with economic interests, with all kinds of different kinds of interests circulating around them. And so then when you see a compilation again, petabytes and petabytes of facial recognition data and all coming forth from, you know, a Facebook or Amazon or Microsoft saying that this is an objective compilation of different kinds of spaces and faces and parameters and forms of recognition of, of faces now digitized, right?

To be able to say, is it really, and is it, has it not been also compiled by, I mean, again, gesturing towards some of the past questions that African-American data scientists and researchers all provoked us, brought us to ask, which was, are those facial recognition data sets actually ones that are fully representative. And if given that they are, we know that they're not can we take a look at again, the backdrop of the cultural context that produced this, that the policies and relationships behind the dataset to really shine a spotlight on that, as opposed to just paying attention, keeping your gaze fixed on the numbers, as opposed to, again, drawing back the lens and seeing the kind of cultural politics and the bodies that were really behind it's making with real interest and politics in mind.

**Dan Bouk:**

If I could say just one more thing to add onto Anita too, then there's other, the question was about like the reader of *The Atlantic* article, the writer of *The Atlantic* article, or particularly anyone whose job involves dealing with a lot of data, I feel like that's where there's a really big change because there's a real responsibility for people who deal with data sets to understand where their data sets come from and to see them not as you see them as one part that wave particle duality, like one result, and they really need to understand the whole thing to know what the data really is. What am I going to make a shout out to an article called "Digital Natives" by Joanna Raiden, which is a wonderful example of what it is to take one data set and really try to do a trace of its history and how that shows the, like, what you really need to know about it.

**Sareeta Amrute:**

Thank you. I think that's all the questions we have time for. There was a quick question in the Q & A. Dan, what is your background? And maybe for Michele and Anita, what's a book on your lovely shelf that you're reading right now?

**Dan Bouk:**

Because these books aren't real, this is from 1930 and it's from the United States Census Bureau 1930 for an advertisement. And it is great.

**Michele Gilman:**

I think "Data Feminism" by Catherine D'Ignazio and Lauren Klein, who Anita had earlier referenced—great book.

**Anita Say Chan:**

I'm reading a bunch of historical texts on W. Du Bois' Atlanta school of sociology, the first American school of sociology fully invisibilized by lots of sociological grand standards much later but really recognized. And even the ASA is now coming back to recognizing the real place of Du Bois' is Atlanta school of sociology that had been completely dehisibilized. So, the work of Aldon Morris, "Segregated Scholars" is another text. So, I'll share the texts or resources with CJ to pass them on.

**Sareeta Amrute:**

Thank you. Please join me in thanking Michelle Gilman, Anita Say Chan and Dan Bouk for your wonderful talks tonight, we look forward to seeing you at future programming and we welcome your feedback. Thank you. And take care.