CATALYST
feminism, theory, technoscience

Open Ethnographic Archiving as Feminist, Decolonizing Practice

Angela Okune
University of California, Irvine
aokune@uci.edu

Abstract

Dubbed Silicon Savannah, Nairobi has become a hot spot of tech development that promises to "save Africa." Qualitative research—carried out by a tangle of private, academic, and non-profit organizations—is part of the work, promising to reveal how people in Kenya are building and benefiting from a dazzling array of digital products. Amidst the enthusiasm, longstanding problems with ways in which research data in Nairobi is conceived, collected, and shared are easily glossed over. This article advances thinking about the politics of qualitative data, unraveling normative concepts like ethics and transparency by both examining existing data practices and modeling alternatives. I describe the sociotechnical infrastructure underlying the ethnographic project, detailing tactics for deploying an instance of open source software—the Platform for Experimental, Collaborative Ethnography (PECE)—to draw research interlocutors into collaborative effort to understand and build decolonized qualitative data infrastructures. Through such processes I learned that collaborating on data not only refreshes the social contract of qualitative work; it can also enhance its robustness and validity. I advise scholars to better document our own knowing practices in order to attend to the inevitability of margins created through all data practices, including our own.

Okune, Angela. (2020). Open Ethnographic Archiving as Feminist, Decolonizing Practice. *Catalyst: Feminism, Theory, Technoscience*, *6*(2), page 1–24.

http://www.catalystjournal.org | ISSN: 2380-3312

Introduction

"I recognize your voice," I was told at an academic conference in 2017. I had arrived at the budget accommodations organized for graduate students and those from the Global South and placed in a dorm with five other women whom I had never met before. It turns out that one of the women studied tech entrepreneurship in Kenya. She recognized my voice because she was finishing the transcription of an interview that her primary investigator had conducted with me when I was still working at Nairobi's flagship technology co-working space, the iHub. "I have been listening to your voice for hours," she explained. When I did that interview in 2015, I was preparing to move to California to begin graduate school and had a growing interest in why particular places and people are heavily studied. I had asked for a copy of the transcript and audio recording from the interview but never received it. This time when I met the project representative, I did not even bother to ask. But she mentioned it of her own accord: "I will send you a copy of the transcript," she promised the morning she left. Nearly two years later, I still haven't received anything. This recounting is not an accusation directed at a particular research group or individual, but serves to question why research data sharing practices in the social sciences are still such an afterthought. Why are they contingent on a chance encounter at a conference or reliant on a well-intentioned individual taking personal time out of her already overloaded and underpaid work schedule? What does this reveal about the values embedded in our current systems of research practice and how might we study and better address these data paradigms?[1]

My research project was born out of first-hand experience in Kenya as part of a group that felt "over-researched."[2] During four years of managing, designing, and conducting research as part of Nairobi's first technology innovation center, the community increasingly felt burdened by the gaze of various global interests. A self-described "open space for technologists, investors, tech companies and hackers in Nairobi" (iHub, 2011), the iHub saw itself as first and foremost serving a community of Kenyan entrepreneurs. With the launch of the research department in 2011 and other departments in subsequent years (Okune et al., 2018), the scope of who comprised the "community" expanded to include international development funders, non-governmental organizations, foreign researchers and their universities, media houses, and private Kenyan and multinational companies interested in learning more about the growth of technology startups in Kenya. As one of the first physical spaces on the continent for African techies and innovators to meet, iHub established itself as a requisite stop for those interested in technology in the region (e.g., CNN, 2012). While iHub's fame and reputation as the nexus of African technology entrepreneurship was crucial for its material success, the everyday experience of being part of the regular tour groups seeking to observe the "Africa Rising" narrative (Beresford, 2016; Nothias, 2014) was a drain on both Kenyan entrepreneurs working out of the space as well as staff.

We felt like our time was often wasted by media, researchers, government, and non-profit representatives all interested in learning about what made the heart of "Silicon Savannah" tick.[3] Everyone asked similar questions: "How has innovation in Nairobi changed over time? What kinds of innovations have emerged from the iHub?" We grew tired of giving the same responses.[4] The inquirers would of course inevitably want to also talk to iHub entrepreneurs, but I grew wary of making those introductions myself because I knew the entrepreneurs also felt research fatigue. "You can find their contacts easily online and reach out to them directly," I suggested, knowing full well that those emails would never be answered. But I wasn't willing to use my own social capital to coax an entrepreneur to take an interview with an academic I didn't know.

Dear [redacted]

I got your contact informations through [redacted], who is supporting me on my research in Kenya. I am a post-gratuade student from Germany right now based at [redacted] University research for my master thesis which is about mobile phone applications in Kenya targeting low-income users.
My study is focusing on technical solution, awarness/marketing and customer needs. Therefore I wanted request if it would be possible to do an interview with you. Also I wanted to req
provide me with any contacts to companies/start-ups or app-developers who are dealing with applications created for low-end users.
The interview won't take longer than 30 minutes and as a consideration I offer you free access to all my findings which might be useful for your project. Please let me know if this would
kenyan phone number is [redacted].

Best regards,

[redacted]

Figure 1. One example of the many emails I and my iHub colleagues received requesting research or media interviews.

We tried to develop alternative ways to communicate what everyone seemed to want to know: sending people to news articles that had already been written on us; research papers that we had written about our work and the entrepreneurs in our community; creating videos and multimedia to watch. But because of the privileging of first-hand accounts and "original" research, we were told each person would have to come and ask the questions themselves. "Because an economist has a different perspective and set of questions than a political scientist," I was told when I questioned why a researcher could not just use outputs that someone else had produced. But the questions they all asked did not feel significantly different to me. Figure 1 is an example of the kinds of email interview requests that I and iHub colleagues frequently received. After taking more than two hours with this European master's student, I explicitly asked him to "please share the transcript from our interview" so I could send it along to future research requests I would inevitably receive. I never received it; nor did I ever hear back about any findings or eventual resulting outputs, despite his original stated promises.

Each day was busy at the iHub; the bright space was always abuzz with twenty-somethings tapping away at their computers or energetically pitching their next idea to a friend over the noise of the barista preparing a steaming cup of house coffee. The iHub Research team was both a part of and apart from the technologists that were developing their code bases and building their businesses. We sought to conduct research that they would consider helpful, but our research agenda was more often directed by the interests of those who were funding our work. On any given day, an iHub Research team member might attend a breakfast meeting hosted by the Kenyan Information Communication Technology Ministry; be interviewed by a postdoctoral student from Oxford University; and then dash to a nearby informal settlement like Kawangware to interview cyber cafe users for some donor-funded research project. Slipping from being a researcher to a research subject on such a regular basis began to engender a reflective capacity within our team.

Interacting with several non-profit groups working on technology for development in Kibera, one of the largest informal settlements on the continent, I observed the same phenomenon of feeling "over-researched" also existed there. These were groups working just across the street from our offices on Ngong Road that were pitstops on the same "tour" that iHub was also part of. Except that these groups—working in environments of extreme poverty with mobile tech towards community empowerment—embodied the flip side of the coin, "Africa as Hopeless Poverty."

Much of the literature in medical anthropology, postcolonial science studies, and critical development studies frames research ethics in the Global South using well-worn caricatures of the white, Northern, wealthy researchers, and the non-white, poor, Southern "over-researched." Studying the problem of over-researched communities in the Shatila Palestinian refugee camp, Sukarieh and Tannock (2012) note that while over-research can arise anywhere, it has been most prevalent in communities characterized by three key features: "communities that are poor, low income, indigenous, minority or otherwise marginalised; communities that have experienced some form of crisis…and communities that are accessible to outside researchers, in particular, by being located in close geographical proximity to research centres and universities" (p. 496). While these indeed appear to be some of the factors that mark idealized research subjects, I also ask, what are the implications for research practices when relatively privileged techies like those at iHub also feel over-researched? Moving away from simple binaries like "poor Africans" and "rich whites" that often arise in discussions about ethical research in Global Souths[5] helps to focus attention on how caring for, about, and through data becomes a recursive practice that takes knowledge production itself as constituted by extraction and exploitation, regardless of who occupies the positions. It is important to look at the ways in which norms and standards of research worlds, such as collecting informed consent or sharing/not sharing research results, are co-constituted by positions such as race and indigeneity and have come to exist in the wake of histories of dispossession, extraction, and "development."[6] In this paper, I seek to refresh how we think about the politics of qualitative data, describing open ethnographic archiving as a feminist, decolonizing practice.

Research Data Politics

Instead of framing open ethnographic data around the concept of transparency, I align with longstanding work by Indigenous and feminist scholars who have been concerned with the enacting of more just and equitable approaches to research and data. Sociologist Victoria Reyes (2017) has developed a model of ethnographic data transparency for understanding the different ways in which ethnographies can be transparent. She defines transparency as being precise about what we are counting as data or information and how and over what period of time we collected them. She argues that if research relies on data for the accumulation of knowledge, then making ethnographic data more transparent should be of scholarly concern. While I share many of the same concerns as Reyes, I avoid the term "transparency" first because of its use historically in the Kenyan information communication technologies sector. As Lisa Poggiali (2016) has pointed out, since the early 1990s, when transparency emerged as a major discourse of governance in Kenya, the term has been leveraged by aid donors, Kenyan NGOs, and Information and Communication Technologies for Development (ICT4D) initiatives. More recently, transparency has been a key justification for open data advocates pushing for improved access to private and government datasets. Second, there is a growing paradigm of "transparency" in some of the social sciences, a pushing for greater publication and standardization of data in order to improve the claim-making of social scientists and heighten the perceived rigor of their evidence and integrity of the work. I avoid the term "transparency" here to distance this project from such discourse that promotes openness as a means towards greater "reproducibility" in social science and a way to reclaim scientific rigor and (re)consolidate its power.

Instead, I turn to work by Indigenous communities such as the San[7] groups of

Southern Africa and North American and Australian aborigine communities who have devised various ways to push back against exploitative research practices including the development of Indigenous research ethics codes and guidelines,[8] community-researcher contracts (Traynor & Foster, 2017), and community peer review (Liboiron et al., 2018), among other things. These guidelines may offer important lessons for understanding how "indigenous data sovereignty" (Kukutai et al., 2016; Lovett et al., 2019) could translate for communities not necessarily organized around a frame of indigeneity. These literatures also highlight the tensions between desires for repatriation of data (in digital form) to the communities from which they originate and worries that full access of digitized community data to diverse online publics may not in fact always be in the communities' interests (Christen, 2011).

Moving away from a quest for purity, in this paper I seek to unravel, rather than rely on, normative concepts such as ethics and transparency by documenting my own knowing practices and methods, foregrounding collaboration as situated, contingent, and ambivalent, yet showing how it operates as a stabilizing and important force in knowledge production. Feminist ethnographer Kamala Visweswaran calls for anthropologists to conduct more home-work or "anthropology in reverse," which she describes as "speak[ing] from the place one is located to specify our sites of enunciation as home" (1994, p. 104). This article is an attempt to reflect on how I have undertaken my project in order to partially address some of the same critiques of research practices that I am studying.

Nairobi's Research Landscape

There are at least two distinct veins through which qualitative research has become mainstreamed in contemporary Nairobi, although today the strands are increasingly interwoven. Qualitative work was established as a component of colonial development interventions that scholars such as Helen Tilley (2011) have described in detail. As critiques of big Development became widespread in the 1980s and 1990s, donors and agencies turned to qualitative work as an important mechanism to understand why development subjects did what they did and why changes occurred on the ground, usually to their surprise. This type of qualitative work is nearly always combined in some way with other methods such as surveys (Biruk, 2018) or randomized controlled trials (de Souza Leão & Eyal, 2019; Rayzberg, 2019).

Another notable strand of qualitative work emerged from companies developing products for African consumers to better understand their customers. In 2010 Kenya was put on the international map as an example of African-driven innovation after the incredible uptake of m-Pesa, a mobile banking platform actively used by almost half the population in Kenya to send and receive digital money. Since the rise of m-Pesa, and in concert with efforts by the government and local technology sector to develop Kenya into "Silicon Savannah," Nairobi has seen a spike in the establishment of regional headquarters for multinational technology companies like IBM Research, Google, and Microsoft. While this work often maintains development and humanitarian rhetoric about how digital technologies will help Africans, this user-driven strand of work also often critiques and distances itself from donor money, arguing that business relations are more egalitarian and less paternalistic than aid. This strand of qualitative work has grown exponentially due to investments of tech capital into the city and the introduction and continuing popularization of user experience research.[9] This mode of inquiry also holds ties to market research performed by corporate firms working for the likes of multinationals such as Coca-Cola and Unilever.

The processes of producing qualitative research in Nairobi both under colonialism and now under postcolonial "Silicon Savannah" have impacted certain city residents who feel over-researched: "After all this research, what are the findings like, what are the results, how is it going to help us?" they've asked me. In spite of attempts to use ethnographic work to improve the relevance and grounded nature of field research methodologies and approaches, the data question thus still lingers. Where is all of the data being stored? How is it managed and who has access? What guidelines, protocols, and review and publishing processes accompany the data?

Given Kenya's longstanding importance as a site of scientific research and the heavy density and saturation of research work, it is a particularly important place from which to begin to think about what sharing data across individuals, organizations, disciplines, and other boundaries might look like. I conducted fieldwork from January to December 2019 focused on commercial, non-profit, and academic research organizations in Nairobi and investigating what counts as ethical research for researchers producing data about communities and how that has shifted in the recent past. Given the project was born out of an interest in the effects and affects of research (hyper)-saturation, I was especially interested in capturing a horizontal slice of this work in research-busy (Neal et al., 2016) Nairobi. Thus, my ethnographic sites include three research organizations in the city, with a keen focus on day-to-day data practices.

Located next to Kibera, one of the largest informal settlements on the continent, the Z[10] research offices are split across two buildings. In the older building, lab staff work with recruited Kenyan research subjects, who mostly come from informal settlements like Kibera. Subjects are given research prompts on a mobile phone or tablet, which are then aggregated to generate insights into their preferences and behaviors. In the other building commonly referred to as the headquarters, over fifty project staff and directors analyze and write-up insights based on the lab data and manage client/partner/funder relations. Qualitative insights have not been a central part of the organization's work until this past year (2019) when an internal qualitative lab was launched to better contextualize and ground the organization's quantitative experiments.

A fifteen-minute walk down Ngong Road leads me to the much smaller team X, who recently upgraded to a new office on the thirteenth-floor penthouse that overlooks much of Nairobi. About six full-time Kenyan staff make up the X team, and their research work is often done in tandem with, if not secondarily to, policy and advocacy work. The X managing director is a long-time advocate of open data in Kenya and has well-established networks and connections amongst governments, donors, and grassroots organizations abroad and in Kenya.

It takes about thirty minutes to move from X research office to W research office by *pikipiki* or motorbike, my preferred means of transport in a city notorious for its traffic congestion. The W research company, although renamed in mid-2017 after being integrated into a multinational consulting group, has been based in Nairobi for over twenty years. Many of the Kenyan staff have been working there for over ten years and the European management duo have been based in Nairobi for over a decade with established roots and kinship ties to the region.

Despite different dynamics, staff across all teams are often stretched thin with multiple projects and a constant sense of not having enough time. Given that these organizations are largely fueled by donor and client projects, the three organizations depend on a steady flow of incoming projects. Staff are thus torn between a constant need to successfully deliver on existing projects while also generating leads and proposals for future projects. Juggling team capacity means that people are often moved on and off projects. Certain individuals on the teams travel around the globe frequently (although not all staff travel equally). And while the different groups have attempted to develop processes to share information across their internal teams, there remains siloed knowledge, information, and data within each of these research organizations.

Infrastructuring Collaborative Ethnography

One of the first things that I quickly learned during fieldwork was that it was early to be fostering conversations about sharing data across organizations because even within the organizations, they were not necessarily sharing data with team members. This was surprisingly true of both the well-resourced multinational company as well as the small startup. Despite the overall lack of established data sharing infrastructure, I nonetheless found qualitative research data[11] to be widely perceived as a legitimate object of concern shared by individuals and organizations working at various nodes within the Nairobi research landscape. Most if not all of the researchers—from short-term field officers to upper-level management—were already familiar with critiques about certain populations in Kenya feeling over-researched. All of the organizations that I ended up working with were keen to understand how, if at all, research data infrastructure could be established to better address this critique.

Leveraging this shared interest, I carved out a field where epistemic partners concerned about qualitative research data and its potential for sharing—across diverse interests, organizational structures, and backgrounds—have begun to dialogue. Sarah Pink argues that collaboration is "precisely about the relationality between persons and things; it is about process, agency, shared knowing, making and practice" (2014, p. 48). From this definition, collaboration then follows not as something to be achieved as a thing in itself but rather as "an emergent quality of the relationships within the encounters that constitute the mixed temporalities of the ongoing event of ethnographic research" (Pink, 2018, p. 206). Such collaborative relations reconfigure the ethnographer's role to include working with people to make things while also seeking to document and understand the structures and processes at play.

Tactics

Aware of the widespread lack of qualitative data archiving and sharing infrastructure in Nairobi research organizations, I was quickly drawn to the Platform for Experimental, Collaborative Ethnography (PECE) when I first learned about it.[12] I was particularly impressed by the explicit epistemological values embedded in its design.[13] In comparison to most qualitative data software, which are built on a coding paradigm, the platform is built through poststructuralist theorizing (Khandekar et al., in press). Using PECE enabled me to conduct analysis at varying scales and with multiple layers of resolution. Thus, as my knowledge of PECE and its possibilities grew, I decided to develop my own instance of PECE, which I have called "Research Data Share" (RDS) (see Figure 2 for a sense of how

RDS looks).

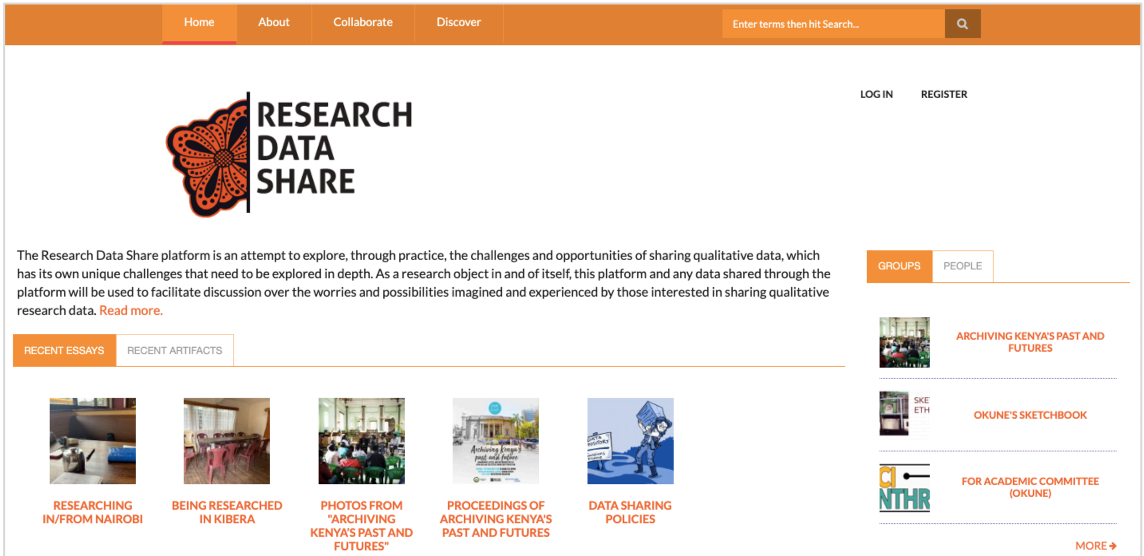


Figure 2. Landing page of Research Data Share (RDS) platform (www.researchdatashare.org)

I conceptualized the development of the RDS qualitative data archive under three distinct rationales. First, I saw it as an elicitation device and grounds for collaborative discussion and engagement, imagining that the deliberations about the archive that I would have with those in the field would be a basis for my learning. Second, it was an attempt to produce something of value to informants and respond to the ethical questions that I started with. At the very least, I could give a transcript and/or audio recording from the research encounter back to my interlocutor. Third, I anticipated that key questions would emerge through my own process of building and studying that would be valuable.

Over the course of the research, I found research data provided scaffolding upon which I could investigate the politics of knowledge production in Africa at its many scales. By intentionally forming an ethnographic data platform to both study and use myself, I reconfigured my relationship with the topic as well as my relationship with interlocutors, enacting a new form of collaborative ethnography that took my own complicity in the structures of knowledge as a starting point for theorizing how researchers might better navigate, organize, and remix existing collections of data. Focusing on an object—data—that resulted from but was not of the researcher/researched relationship opened up discussions beyond critique and set in motion a new set of social relations to study. It also situated me squarely as a participant in the production of the very things I was studying.[14]

As part of preparing myself for grappling with data sharing issues in the field, I developed textual devices I thought might be necessary for establishing some of the social infrastructure of data sharing. Inspired by the work of Max Liboiron and collaborators (2018), I developed two draft working documents—a data circulation form[15] (Okune, 2019b) and a collaboration agreement[16] (Okune, 2019a)—to reflect my current thinking and also with which to continue to think.[17] The intention behind the forms was not to create yet another moral procedure in a research environment already steeped in biomedical regimes of governing ethics. Rather, I imagined the draft protocols to act as shared objects of attention to focus on with interlocutors and through which to discuss expectations of each other. I saw these devices as part of what Rosa Castillo (Pels et al., 2018) has called a responsive review process where careful discussion of actual and potential ethical challenges and possible ways of dealing with them take place throughout the research.

After sharing the collaboration agreement (Okune, 2019a) with the three research groups within the first few months of fieldwork (January through March 2019), I was given access to a variety of qualitative data,[18] especially digital transcripts of one-on-one interviews and group discussions; photographs; coded summaries of data; final reports; and interview guides. Of this data, I selected one sample from each of the three organizations, anonymized the text (if it was not already), and uploaded it to the RDS platform with any available context as meta-data.[19] I then used this data sample as an elicitation device for initial interviews (Okune, 2019c) and found the exercise helpful to ground what can often end up as an abstract conversation about morals and ethics. Figure 3 shows the Organizational Data Interview Guide which appears as open data on the RDS site.

In addition to these types of organizational artifacts, I also use the RDS platform to upload descriptive event or meeting notes as well as field notes. Utilizing the platform's three privacy settings, I have private (only accessible to myself), group (accessible to a delimited group of users), and public (fully accessible to anyone on the web) artifacts and essays now up on the platform.

Global, interdisciplinary data sharing is a complex cultural system that assembles many actors, institutions, technologies, and frameworks. It is a system animated by a diverse set of forces operating at many different locations and across many different scales. To unpack such a system means learning how to simultaneously observe the multiple forces acting upon and interacting with researchers and the data they produce (K. Fortun, 2009). PECE leverages a shared set of analytic questions called "structured analytics" to act as focusing devices that order what the ethnographer pays attention to. Building off of a set of structured analytics specifically looking at data sharing developed by Poirier and Costello-Kuehn (2019), I use these layers to understand research data through multiple frames. Kim Fortun (2009) has described how, at its best, this kind of scale provides a way to see many types of action in motion at once, evoking a sense of the dynamism of the system at hand rather than locking it down and stabilizing it.

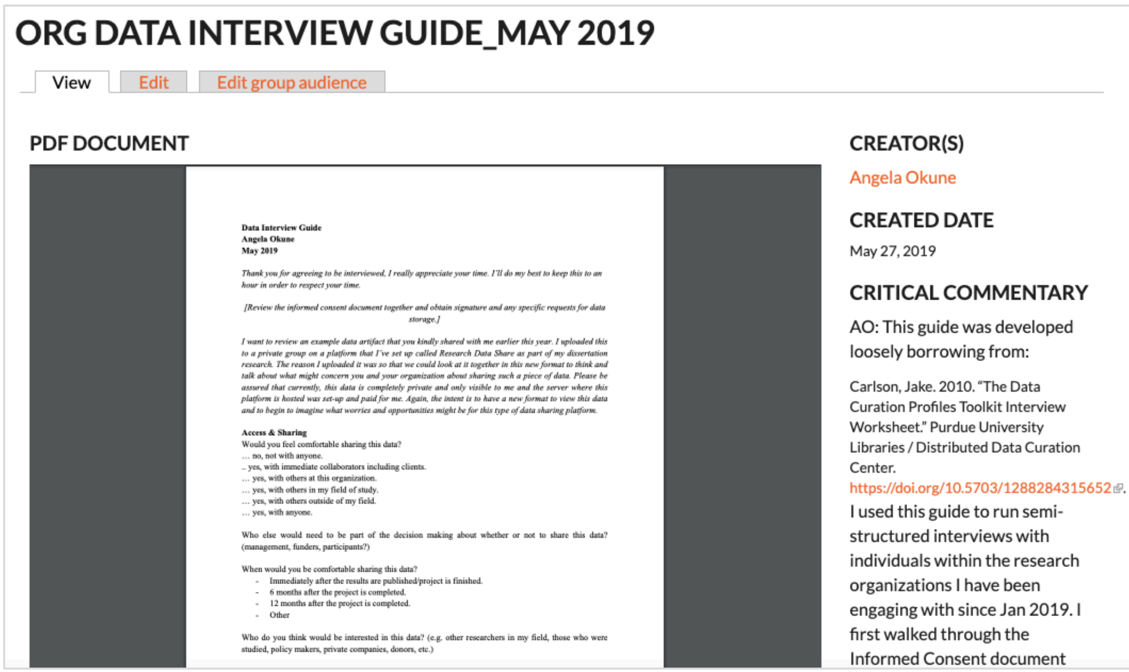
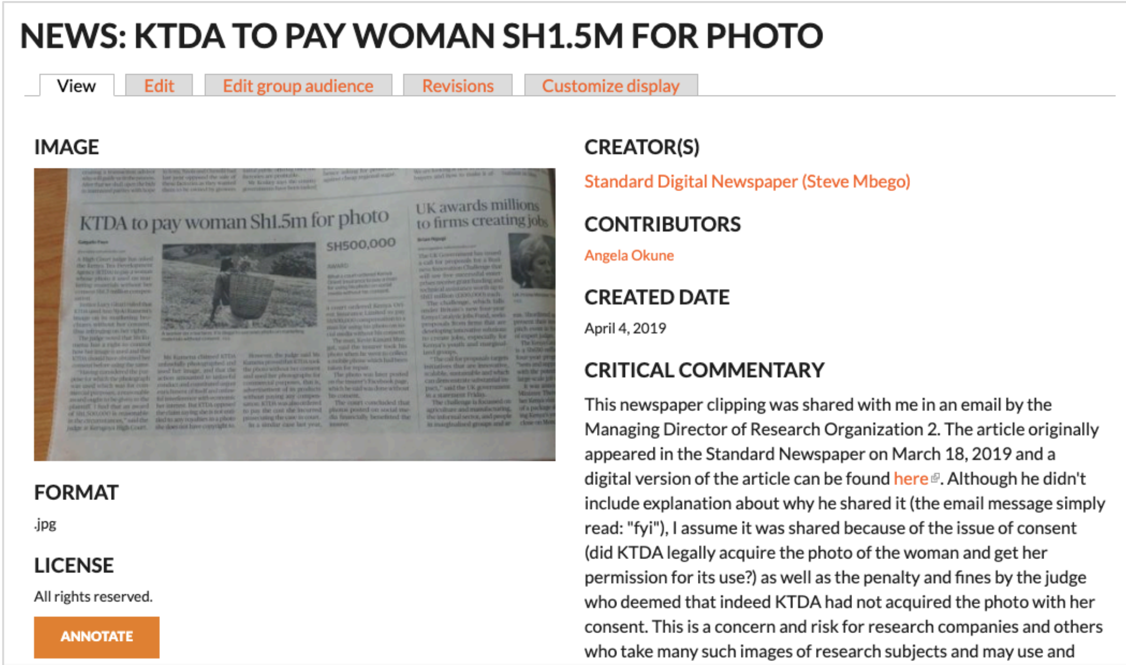


Figure 3. A screenshot of the Org Data Interview Guide in RDS. This is also an example of a PDF artifact. Source: Okune (2019c)



NEWS: KTDA TO PAY WOMAN SH1.5M FOR PHOTO

View | Edit | Edit group audience | Revisions | Customize display

IMAGE

FORMAT

.jpg

LICENSE

All rights reserved.

ANNOTATE

CREATOR(S)

Standard Digital Newspaper (Steve Mbego)

CONTRIBUTORS

Angela Okune

CREATED DATE

April 4, 2019

CRITICAL COMMENTARY

This newspaper clipping was shared with me in an email by the Managing Director of Research Organization 2. The article originally appeared in the Standard Newspaper on March 18, 2019 and a digital version of the article can be found here. Although he didn't include explanation about why he shared it (the email message simply read: "fyi"), I assume it was shared because of the issue of consent (did KTDA legally acquire the photo of the woman and get her permission for its use?) as well as the penalty and fines by the judge who deemed that indeed KTDA had not acquired the photo with her consent. This is a concern and risk for research companies and others who take many such images of research subjects and may use and

Figure 4. This photo artifact in RDS is of an image that was sent to me by the managing director at one of the research organizations. The context for how I came to have the artifact is part of the meta-data contextualizing the artifact. Source: https://www.researchdatashare.org/content/news-ktda-pay-woman-sh15m-photo

Figure 5 from the RDS platform illustrates how such structured analytics also provide the light structuring necessary for interdisciplinary collaborators to dialogue and generate new insights together. At a capstone event I co-organized at the end of my fieldwork (Okune, 2019e), I invited four researchers of diverse backgrounds to also take field notes of the event. The next day, we held a debrief session for two hours where we discussed what we observed and then answered a set of structured analytic questions (Okune & Fortun, 2019).

I am now in the process of transcribing, cleaning, and publishing ethnographic fieldwork data on RDS, care-fully. Since much of the data will become public, it is not enough to transcribe so that I can remember what happened. I share a concern with some interlocutors that open ethnographic data has a risk of losing contextual depth, potentially leading to misapplication or misunderstanding. Given that I am not the only intended audience, I am working through the data process care-fully[20] and as I do, realizing that seemingly mundane data work holds important analytic moments. For example, Figure 6 is an annotation I made on one of the consent forms I recently uploaded, which describes the mismatch between a research participant's stated wishes and the way they completed their consent form.

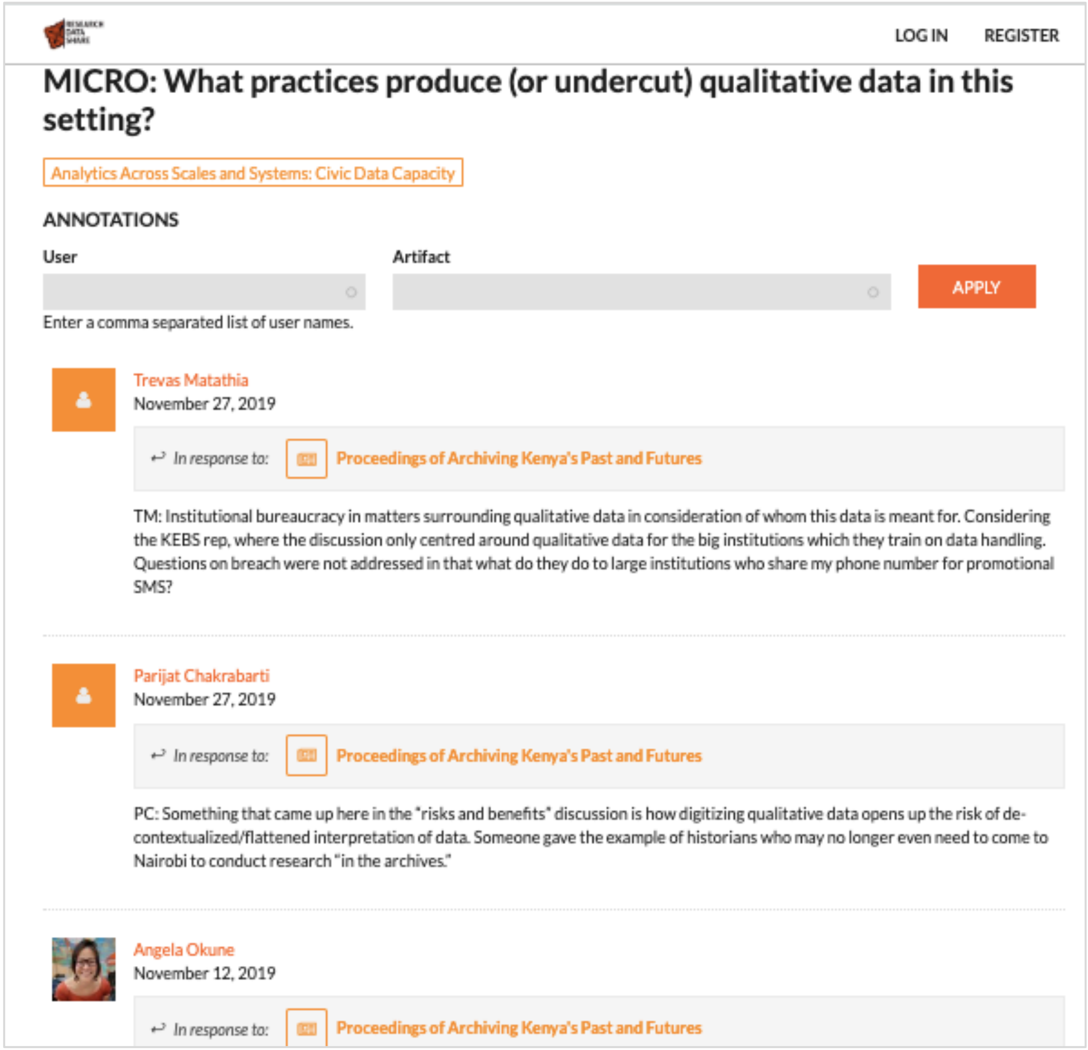


Figure 5. This screenshot illustrates how collaborative annotations can work with three different people responding to a shared question. The analytic question set can be found here: https://www.researchdatashare.org/structured-analytics-questions-set/analytics-across-scales-and-systems-qualitative-data-capacity

As I have been uploading artifacts to RDS, I have emailed the links to interlocutors and research collaborators, sharing outputs with them and inviting them to also analyze the artifacts using the set of analytic questions developed (which are continuously being amended and added to). Figure 7 is an example of an annotation that an interlocutor inputted into RDS after I shared a link to the research transcript with them, illustrating the potential of a platform like RDS to facilitate research participants to "speak back" to the researcher and provide their own interpretation of the data.

MY ACCOUNT DASHBOARD LOG OUT

RESEARCH DATA SHARE

AO. MISMATCH IN STATED DESIRES + PAPERWORK

View Edit

AO. Despite stating their desire to be anonymous, participant #KPKMU09 wrote out their whole name in a legible format on the consent form. (I had given verbal directions to those who wished to be anonymous not to write in their names and stated that I would write in a participant number instead afterwards). Ideally, this form should have been redesigned so as not to ask for names if one wishes to be anonymous. But as it was the officially approved IRB form, this example illustrates some of the cracks where formal mechanisms for protection may not be sufficient. It would not be hard to imagine how a larger sample size and less careful attention to explicit/implicit patient desires and individual/collective risks might lead to a bulk uploading of such consent forms and increased exposure to individual and collective risks. Thus, taking the practices as I attempted in this project as "best practice" and scaling them up in other projects may not in fact be advisable nor desirable. I hope this example helps reveal the entanglements caught up in concepts like Open Data and how such "experiments" within the ethnographic endeavor (asking participants for their desired level of sharing, uploading and sharing materials including consent forms) should be seen as emergent forms rather than fixed targets or clear criteria to be achieved within careful consideration.

ARTIFACT

Informed Consent_191029_001 FGD

ANALYTIC (QUESTION)

How is being anonymized or named in research described? Moralized?

Figure 6. A screenshot of an annotation on one of the consent forms. Find the full PECE essay here: https://www.researchdatashare.org/content/Okune-2019-fieldwork-consent-forms.

Thus, even as I transition back to "home-work" in California, my "field-work" is not over. As interlocutors send me their own interpretations and we continue to keep in touch, these interactions further inform and become part of my ongoing research. The benefit of having this on a publicly accessible data platform like RDS rather than in private email exchange is that I can cite the insights produced by the individual and explicitly acknowledge their expertise, time, labor, and authorship. Doing this collaborative analysis on a public site, looking at the same source material together from different perspectives, shifts the figure of the research participant from a quoted figure (at best), to an analytic actor and authorial voice in their own right (with ongoing ability for them to edit and amend thoughts over time if so desired).

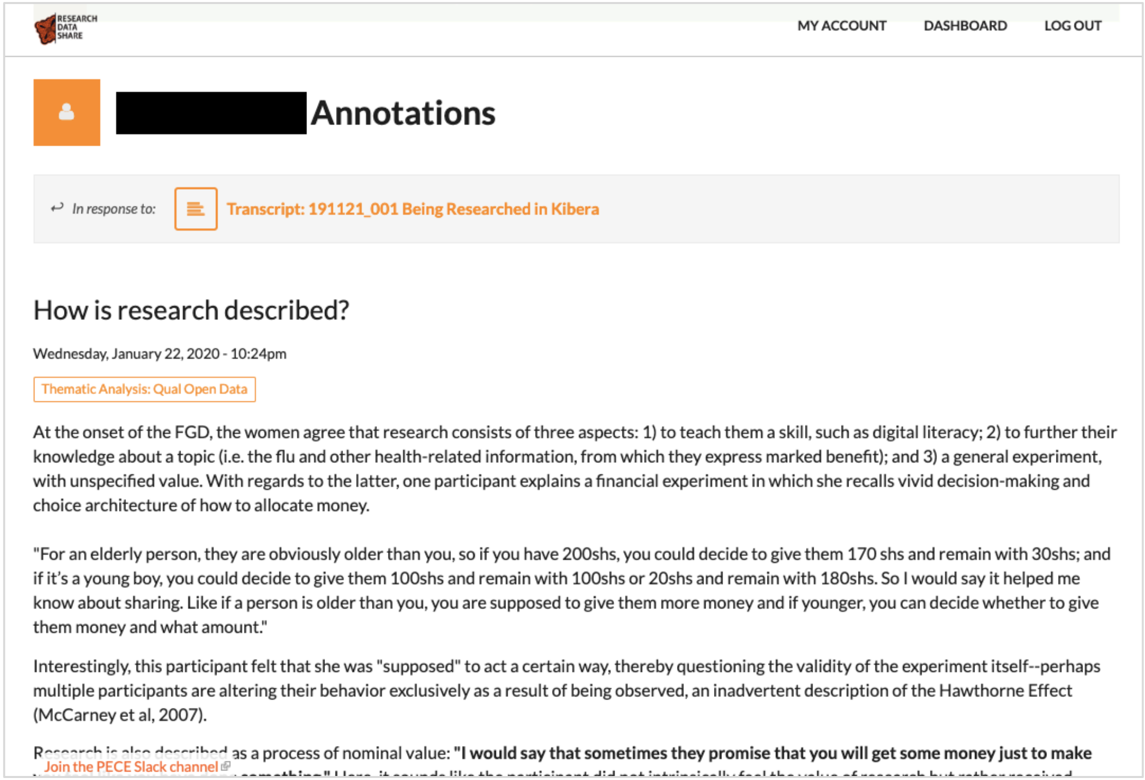


Figure 7. This is an annotation that an interlocutor uploaded to RDS based on a transcript on the platform.

Nonetheless, questions raised as part of the Writing Culture turn (Behar & Gordon, 1995; Clifford & Marcus, 1986) of who is able—based on material, social, educational, technical constraints—to contribute and hold an authorial voice still remain. As scholars of science and technology well know, technical platforms like RDS can be as exclusionary as they can be inclusionary. I am emailing the transcript artifacts out to interlocutors, but there are interlocutors who have no email addresses, some for whom reading a text-heavy document online that requires an internet connection is not easily possible, and others for whom annotating such an artifact is uncomfortable and unnerving. I am not advocating that *all* ethnographic data should *always* be shared; nor am I prescribing that platforms like RDS should be taken up across the board as the best or only way to share ethnographic data. My point, rather, is to unravel normative concepts such as ethical open data and instead call for better documentation of our own knowing practices and methods. I align with scholars like Luis Felipe Murillo (2018) and Eli Elinoff (2017) who have asked us to collectively imagine and begin to practice the creation of an ethnographic data commons. To do so requires close attention to the inevitability of margins created through all data practices, including our own. Instead of fixed targets, clear criteria to be achieved, or universal data platforms, there is need for cultivating expertise in listening and watching for the particularities of data in context. The tactics used should be nuanced and ever changing. If decolonized, transnational infrastructures that unsettle racialized hierarchies of knowledge are to be established, rather than ignoring the margins created by our data practices, we need to practice attending to them care-fully.[21]

Data work is a time-intensive process—tedious, undervalued work that requires the development of a kind of expertise often skipped in qualitative research classes and missing from commonly required coursework. Rarely is such data work framed as co-constitutive of research ethics. I encourage practices of ethical care for our data to be foregrounded and cultivated. Set aside sufficient time to think seriously about how you intend to care for your project data:

- What kinds of data will you collect?
- What organization built the data platform that you plan to use and how is their data ideology reflected in the technical functionalities of the platform?[22]
- Where will the data (physically) sit? (E.g. What kind of entity owns the server? What country is the server in? Under what legal jurisdiction does the server fall under?)
- Who will have access to what data?
- What meta-data[23] will be needed?
- How will you decide the ontological categories that are relevant? With whom and when?
- What activities need to be built into fieldwork in order to have datasets that make sense not only for your individual project but also for those that you conduct research with?
- (How) would your interlocutors like to use the data?
- Should you produce and package the data so that it can also meet their needs? Why/why not?
- What are the costs associated (time and money) with the data practices envisioned and how will you build those in?
- What skills will you need to do this work? How/when will you obtain the necessary training?
- Will you need to bring others on board in your field site to help specifically on data?
- Are there others outside of your field site that you need to consult with?

Validity Redoubled, Collaboration by Design

My project follows Joan Scott's call "to take as [my] project not the reproduction and transmission of knowledge said to be arrived at through experience, but the analysis of the production of that knowledge itself" (1991, p. 797). A key aspect of analyzing such production of knowledge involves recognizing the multiple actors involved in its creation beyond the individual researcher herself. Research collaboration has been a keyword—even a buzzword (Cornwall, 2007) – that, despite problematic uses and warranted critique, nonetheless continues to be indispensable. Luke Lassiter (2005) has described longstanding models of collaborative research and argued that the disparity between the academy and communities in which we work could be narrowed further through collaborative ethnographic practice and writing. Questions of reciprocity and benefit sharing from research continue to worry many analysts, especially those working in contexts with stark economic inequalities (Aellah et al., 2016; Jentsch & Pilley, 2003) where research collaboration is therefore often pursued primarily for its ethical value. In this section, I advance that collaborating on research data not only holds ethical value but also epistemological value.

The previous section detailed the experimental tactics and data infrastructure established for my project. These included installing my own instance of the Platform for Experimental, Collaborative Ethnography; developing textual devices such as a collaboration agreement and data circulation form; reviewing and archiving existing data held by research groups in Nairobi, which were in turn used as an elicitation device to produce new data; and co-hosting a public event in one of the oldest libraries in Nairobi on the politics and infrastructures of research. The proceedings of the event, which I will describe in more detail in the next section, were not only archived on the RDS ethnographic archive, they were also analyzed and publicly annotated with interlocutors (see Figure 5). Through these processes, I learned that collaborating on data not only refreshes the social contract of qualitative work, it can also enhance the robustness and validity of the research data. For example, after emailing a non-Kenyan anthropologist I met in Nairobi a link to a public data artifact from my research, I received the following response:

> I went through the transcript and I really find it fascinating how respondents link research with development. It took me several hard-on collisions with scholars who work in political sciences, agricultural sciences and similar disciplines to realize how much this causality is assumed even by scholars themselves. Being raised in a postmodern science, such blatantly modernist assumptions were like a shock to me. That being said, I could not find much of it in my own interviews which were mainly done with men who mostly linked their participation with [redacted Org Name] with a concept of "work" or almost a contractual situation where they participate in order for

> [redacted Org Name] to develop the community as such (email correspondence, February 20, 2020).[24]

This example helps illustrate how sharing ethnographic data can increase our understanding of complexity in a given site. Since his own interviews were mainly done with men "who mostly linked their participation with.. a concept of 'work,'" looking at my data from an all-women group discussion offered a different starting point, which he may not have had access to otherwise and which seems to have led him to think about the issue in a different way. Equally important, adding his comments to the data via an annotation allows others (including myself) to also build off of his perspective. When I shared this annotation back to him, he responded that it is probably very common for people being researched to assume that it is part of a developing or modernization process, so perhaps it is a bit of a stretch to say that the additional interview data helped reveal that.

> What they [the transcript/conversation] did, however, was to remind me of the importance [of] how such demands are made and how they are framed politically, i.e. are they framed as unfulfilled collective rights or as offers delivered to individuals. I find it fascinating to start thinking through how research as delivering a neoliberal service to consumers clashes with research as advancing and experimenting with collective entities. (email correspondence, February 21, 2020)

I added to the public annotation based on his second email response to clarify his initial statement.[25] This dialogue highlights an additional strength of open ethnographic archives. Beyond just the data, it is the opening enabled by the sharing of annotations and data back and forth. In this case, by sharing an annotation to ask, "is this what you meant?," the earlier statement could be further nuanced and qualified. Again here, through such ongoing processes, ethnography is made more robust, increasing not only its ethical validity (by increasing researcher accountability and representations of those we engage with), but also strengthening its research validity.

The analytic structure that undergirds the collaborative analysis of research data in RDS further highlights this double validity of data robustness and research ethics. Griffin and Hayler (2018) note that collaborators within an endeavor can become marginalized through the denigration of certain kinds of expertise, since "power structures both within and beyond the immediate interactions can lead to the work of one or more collaborators being reduced or going uncredited, and to the detriment of their institutional and subject standing."[26] This echoes sentiments expressed by Kenyan and non-Kenyan researchers working in Nairobi who shared stories of having their refined arguments taken up and used by supervisors and peer reviewers. Many of these individuals expressed worry that an open space for data sharing would therefore only increase potential for such harmful practices. The analytic structure attempts to address this potential for marginality produced in collaborative research by establishing a public mechanism for collaborators to chime in directly with their thoughts. The structure also allows for insights produced by the individual to be explicitly credited, validating their expertise, time, labor, and authoritative voice. This can be important for researchers and non-researchers alike. In a discussion with researchers/research subjects who reside in Kibera, they mentioned, "You know, it's also important it reaches a level for us as researchers, in Kibera, or leaders, where we would prefer that some documents are quoted with us. You know, because it's professional, it's information based. And as a leader, if a document for example, raises about the level of sanitation, the level of governance is quoted on your name, in Kibra it gives you credibility. You know, on your leadership. [Others make sounds that show agreement]" (KPKMU09 1:03 October 29, 2018).[27]

Scholars are increasingly realizing that for data sharing to be useful, it cannot be thought of as simply depositing decontextualized results to tick a transparency box (Pisani et al., 2016). Rather, data sharing should be treated as an integral part of every stage of the research cycle and part of what Castillo (Pels et al., 2018) has articulated as ongoing development of a researcher's capacity for continuous reflection on research ethics throughout the life of a research project. The reasons that researchers do not share data are becoming better understood (Borgman, 2012; Tenopir et al., 2011). There is a lack of training (Pels et al., 2018) and the labor required of the individual researcher to enable reuse and sharing is significant (Borgman, 2012) and still not yet acknowledged and rewarded by traditional academic structures. Many are worried about having their ideas and data stolen by others. I have been warned by various scholars and administrators that attempting to put in place the structures to enable data sharing as part of my project is additional labor, effort, and financial burden. Many have questioned why I am doing it. In this piece, I have sought to share the experiences that lead to my deep-seated desire for attempting fieldwork data practices differently.[28] I continue to consider ways in which information could be recursively doubled back into communities and allowed to be reintegrated and reused. For me, that importantly has begun to circle around the question of how we set up, organize, and manage the research data that we collect.

Archival Fever in/with/for Kenya

"I was proud to be a Kenyan today," one of the panelists confided to me as we munched on the marinated chicken and soft chapati roll provided for lunch. "Oh interesting, why?" I probed, curious how the day-long workshop on research data archiving that I had organized tied into her sense of national pride. "Because Kenyans are doing good things. It can feel really alone when you are abroad there. There are a lot of deficit narratives. But to hear all the interesting things that people are doing here makes me feel part of a community." She smiled, her Maasai beaded headpiece jingling slightly as she took a big sip of fresh mango juice. "Yes, I never would have considered myself a 'data' person, but coming together with all of these people to think about how we need to decolonize our knowledge and infrastructures has connected me with interesting people doing really great things here in our own Nairobi," another panelist chimed in.

The geopolitics of knowledge production have often been reduced to questions of national or racial representation alone. While not to be dismissed, such limited articulations of the problem risk reproducing the oversimplified fetishes and categories themselves. Like the difficulty of categorizing the "over-researched," Nairobi researchers also do not fit into neat categories and include a young tattooed Kikuyu woman studying for her CPA exams and translating research surveys into Swahili at night as a part-time hustle as well as a white French man married to a Luhya woman who is overextended in his role as the go-between for a decentralized management team in Chicago, London, Tokyo, and San Francisco, and the local Kenyan research team.

Beltran's (2020) contribution to this special issue describes how Mexican female hackathon participants consistently shuttle between conflicting markers of class, nation, ethnicity, and race constructed across the US/Mexico border. He points out that these markers intersect across a terrain of transnational labor that, despite valuing technical expertise, differentially recognizes and rewards such expertise. Similarly, the diverse researchers I studied are highly attuned to the multiple layers of local and global politics that structure their contributions to knowledge production. Most recognize their positions within transnational research assemblages and sense that existing structures marginalize them and their contributions. Building an ethnographic archive in/with/for Kenya then is not necessarily about enabling Kenyans to enter global conversations or showcasing their work so it becomes internationally recognized. It is as much about connecting "already global local"[29] players with each other in Nairobi to spur a collective imagining about what an ethnographic archive for Kenya's intellectual workers could be.

Open data emerged about ten years ago as a rough point of consensus for action among pro-democracy practitioners, internet entrepreneurs, open source advocates, civic technology developers, and open knowledge campaigners (Davies et al., 2019). In 2011, the Kenyan Open Data Portal was launched, making it only the second of its kind in Africa. While the open data label has shifted in its application around the world to genomic data, land registers, and parliamentary voting data, the concept of "open data" still largely refers to numerical data held by governments, NGOs, and private sector. In comparison, conversations about ethnographic archives have largely remained amongst libraries, archives, museums and their patrons (Silverman et al., 1992).

As part of my work, I have sought to break up some of the established discourse that has heavily focused on technical specifications, business models, and standards. On November 12, 2019, I and a team of volunteers and friends hosted an event titled "Archiving Kenya's Past and Futures" at McMillan Library in Nairobi's Central Business District. The event brought together fifty researchers, archival specialists, open data technologists, and government representatives to think about the intersection between open data technologies, digital humanities, and research data practices in Kenya. Figure 8 presents a photo artifact from the event and full proceedings including video footage, transcripts of all panels, and all distributed materials are available on RDS.[30] The aim of bringing such a diverse group together was, as Emily Yates-Doerr (2019) has described, to foster a space of exchange and learning where collaborators come from places of difference and practice "careful equivocation" (Yates-Doerr, 2019) to unsettle the binaries often drawn between one object, category, or term and another. Facilitating development of such connections across difference and emergent discursive communities—building up and spreading the "fever" of the "archive fever"—is also a critical component of developing archives as feminist decolonial practice.

Figure 8. This is a screenshot of an image in a photo essay in RDS. The photo essay includes images taken at the November 12, 2019, workshop. The full essay can be viewed here: https://www.researchdatashare.org/content/photos-archiving-kenya%E2%80%99s-past-and-future

Scholars have challenged the assumption that an archive is a neutral, immutable, historical repository of information, arguing instead that the archives are a place where important decisions about what documents—and therefore whose history—are made (Arondekar, 2009; Stoler, 2009). In my opening presentation on November 12, [31] I reminded attendees that it was not just the contents in the archive that needed careful attention as we sought to decolonize knowledge; the infrastructures themselves need to also be care-fully attended to. Decolonizing knowledge is not only about ensuring African voices are found in the historical archives and are represented at present-day academic conferences. If what is actively collected, recorded, counted, and audited is a form of memory, then the same can be said of socio-technical infrastructures, which are not just the built platforms of knowledge-making but also the non-innocent remainders and memories that shape contingent futures (Bowker, 2008). Decolonizing knowledge then is also about actively curating, building, and studying the archives we want in the future, including the socio-technical infrastructure on which materials sit. We can start by supporting and protecting more collective public knowledge commons and checking the power of Western corporate academic publishers, which are in the process of taking over much of the research infrastructure.[32]

From the standpoint of researchers based in Nairobi, it is apparent that, as Kenyan novelist and postcolonial theorist Ngugi wa Thiong'o has pointed out, "the languages and the literatures of the peoples of Africa, Asia and South America are not peripheral to the twentieth century. They are central to the mainstream of what has made the world what it is today. It is therefore not really a question of studying that which is removed from ourselves wherever we are located in the twentieth century but rather one of understanding all the voices coming from what is essentially a plurality of centres all over the world" (1993, p. 28). How to move the center of Kenyan research data back to Kenya while simultaneously de-nationalizing and building archives that are about discovery beyond the nation-state boundaries, precisely because we are in fact a community of "already global locals"? Following work by anthropologists and digital humanities scholars,[33] I see promise in archiving ethnographic data as feminist, decolonizing practice.

Calling for confounding our relationship to how and why we do archival work Anjali Arondekar (2009) reminds us that an archival approach should articulate against the guarantee of recovery. In line with this, I conceptualize my engagements with ethnographic archives using Derrida's conception of the archive as a kind of "techno-prostheses" (1998, p. 26) of a particular temporal and spatial experience of the world. The scholarly archiving of the nature I am undertaking is not simply to save or preserve but to provide grounds for further questions, working with people to take care of the data while also documenting to understand the processes, relations, and considerations at play. The point of this kind of an archive is to scaffold a deutero capacity to think about the world and support a rethinking of our habituated ways of understanding the world. Instead of worrying over the sharing of data for greater reproducibility or transparency, I echo Indigenous STS scholar Michelle Murphy who asks us to think about the infrastructures of relation that make some kinds of flourishing possible and other kinds of flourishing not possible (Kasdogan, 2018). An important first step is to squarely understand the exercise of power in decisions about where we store our research data and then try to do something more than that.

Acknowledgements

Thank you to the editors and anonymous peer reviewers who provided incredible assistance in improving this paper. A big thank you to Kim and Mike Fortun, Leah Horgan, and participants of the UC Irvine Anthropology Working Title session for your edits and feedback on earlier versions of this paper. This research was supported by a Wenner-Gren fieldwork grant; University of California Santa Barbara Blum Center on Poverty, Inequality and Democracy grant; and National Science Foundation Graduate Research Fellowship grant. A special thank you to Book Bunk, the Research Data KE Working Group, Quotidian Data and PECE

design teams, and interlocutors and collaborators of this project.

Notes

[1] Scholars from a wide range of disciplines have been working on similar questions critiquing existing scholarly practices and structures. See, for example, Pisani et al., 2016; Ferguson, 2012; Tuck and Yang, 2012; Rooks, 2006; Smith, 1999.

[2] I first found the concept of "over-research" described by Clark (2008). Upon entering the academy, I was exposed to a much longer standing set of work in Indigenous studies that describes the long-seated feelings of being exploited by research (see, for example, Smith, 1999; Deloria, 1988). A related concept is what Michelle Murphy (2017) has described as postcolonial thick data, where *thick* describes the dense enumeration practices in the twentieth century that are now entangled with the history of twenty-first-century big data to produce new forms of empire. The practices that produced such thick postcolonial data appear to have also produced feelings of research exploitation.

[3] Since 2010 when iHub opened, international headlines about Kenyan innovation have proliferated (Marchant, 2018). Along the way, the moniker "Silicon Savannah" has caught on and come to be used by some to describe Nairobi's technology community.

[4] This fatigue with the repetitiveness of the research questions echoes experiences recounted by Sukarieh and Tannock (2012), who quote a Palestinian refugee camp resident complaining that researchers keep coming into the camp and asking the same questions over and over again: "They do the same interviews with the same people all the time…Why don't they use the interviews that were done by researchers before?" (p. 501).

[5] I pluralize "Souths" here to flag what Aellah et al. (2016) have pointed out—that increasingly inequality is in fact appearing between a global middle class and the vast majority, rather than between countries. In pluralizing the term, I seek to point to the wide diversity of communities and experiences both within former colonized countries as well as within former European colonial powers and their North American successors that have spaces and peoples negatively impacted by contemporary capitalist globalization (Mahler, 2017). Comaroff and Comaroff (2012) have called out this suspect North-South dualism, suggesting scholars work the contradictions in order to move beyond such dualism and address the larger dialectic processes of which the term is a product. Thus, while this article

investigates computing research from the standpoint of the "South"—that is, inspired by, grounded in and based on lived realities in Nairobi, Kenya—the article also complicates that frame, demonstrating how ideas, data, and knowledge travel across time and space in diverse and unequal ways that do not necessarily align to such a binary dualism.

[6] Many of these ideas are studied deeply in the anthropology of development literature, which, for example, has critiqued ideologies of development (Cornwall, 2007), notions of expertise (Crane, 2013), and, more recently, the intersection of expertise and race (Pierre, 2013).

[7] I use the term "San" here, but would like to acknowledge and flag the ongoing debates over the terms of reference for the groups: San, Jun/oansi, "bushmen," "hunter-gatherers," BaSarwa, etc. For example, in Namibia, Jun/oansi call themselves "bushmen" when speaking Afrikaans, but otherwise call themselves Jun/oansi. Find more on this topic in my blog post "'Responsible Research': Reducing Risk or Improving Well-Being?" (Okune, 2017).

[8] See, for example, the First Nations Principles of OCAP® (http://fnigc.ca/ocap.html) and the Alaska Federation of Natives Guidelines for Research (http://www.ankn.uaf.edu/IKS/afnguide.html).

[9] See Lily Irani's (2018) description of design thinking as a form of technical expertise. Design thinking and related methods have gained increasingly wide-spread popularity in Nairobi and other parts of the continent, promoted and taught by private design companies, Western universities, and philanthropic organizations.

[10] I have anonymized organization names for this piece not because the organizations are not necessarily willing to be named but because of the ongoing nature of my writing and relationships.

[11] I use "qualitative" and "ethnographic data" interchangeably here to refer to materials produced as a result of ethnographic data collection methods. These are typically field notes, photographs, audio and video recordings, found artifacts, typed transcripts of interviews and/or focus group discussions, oral histories and may also include archival documents. In my field sites, the term "qualitative data" is much more recognizable.

[12] I came to learn about the Platform for Experimental and Collaborative Ethnography (PECE) in 2017. I developed skills in PECE in 2018 by participating in the STS Across Borders exhibit (http://stsinfrastructures.org/about) organized by Society for Social Studies of Science (4S). I also eventually developed my comprehensive exam "documents" in PECE. These documents are accessible online at http://stsinfrastructures.org/content/querying-science-and-technology-studies-africa. While I had begun dabbling in PECE simply for my own research infrastructure, over time I found its broader value in adding an analytic layer to ethnography at multiple points in the process.

[13] For more on this aspect of PECE, see M. Fortun et al., 2017; Poirier et al., 2014; Poirier, 2017.

[14] This is similar to many of the cases described in Estalella and Criado (2018). However, unlike Marrero-Guillamon (Chapter 8 in Estalella and Criado, 2018) who had the realization that he could not in fact "participate" at some kind of distance *while* he was already in the field, I intentionally built this kind of entangled participation and archival practice into the field research project at the pre-field project design stage, which gave me time to prepare textual devices, develop the technical infrastructure, and build up the skill sets that I believed might be necessary to execute such a project.

[15] The draft data circulation form (Okune, 2019b) emerged from pending questions I have regarding how to deal with anonymization as well as how to grapple with the ethical question of whether the release of co-created data from the researcher and research participant encounter is even a decision that can or should be made by the researcher alone (Mannheimer et al., 2019; Moore, 2007).

[16] The collaboration agreement draft document (Okune, 2019a) emerged for multiple reasons: first, to make clearer what exactly is entailed in my research project; second, to have a textual device with which to discuss expectations and desires from the research; and finally, given the expansive project scope, to protect myself from infringing unknowingly on existing copyrights.

[17] This data was produced collaboratively through conversations with various experts, including scholarly communications officers from the University of California, Berkeley and academic advisors. I decided to upload these documents to Zenodo rather than the UC institutional repository because I was unable to clearly determine the access policy under the institutional repository (i.e., access after graduation and abilities to continue to update or migrate materials).

[18] I circulated the collaboration agreement draft to all three organizations when I reached the stage of requesting access to their (private) data. The agreement, as predicted, allowed us to articulate expectations for each other. For one organization, there were several rounds of iterative feedback. Interestingly, none of the organizations actually signed it into force before sharing data with me, highlighting the document's value not necessarily as a legally enforceable document, but rather as a vehicle for the clarification of expectations and collaboration norms.

[19] While ethnographers may balk at the term "meta-data," perceiving it as computing jargon that signals a particular technical standard or practice, in fact, adding meta-data to ethnographic digital data can helps assuage some of ethnographers' biggest worries about their data being taken out of context. Through the application of meta-data, ethnographers can "situate[e] data collection and analysis in a particular time, setting, and cultural context… encouraging re-interpretation of archived data [and] perpetually generating new, interpretive data which itself can be contextualized, archived, and re-interpreted" (Poirier et al., 2020, p. 226). Figure 4 illustrates contextual meta-data added to a digital artifact in RDS.

[20] Work on care from the science and technology studies (STS) community highlights multiple definitions, finding care is often a site of "ambivalence, tension, and puzzlement" (Atkinson-Graham et al., 2015, p. 738). Murphy (2015) lays out four interlaced meanings of care, noting "first, it refers to the state of being emotionally attached to or fond of something; second, it means to provide for, look after, protect, sustain, and be responsible for something; third, it indicates attention and concern, to be careful, watchful, meticulous, and cautious; while its fourth meaning…is to be troubled, worried, sorrowed, uneasy, and unsettled" (p. 721). In thinking about how we might practice data care-fully as part of a decolonial feminist practice, I hold these multiple definitions in tandem—seeking not to reify data in and of itself but to simultaneously also not discount its importance in furthering collective knowledge and as an artifact of an ethnographic relationship. I seek to cultivate an affective stance towards and a practice that embodies being worried, troubled, and uneasy about the data and its margins while simultaneously also recognizing a responsibility to protect, look after, and be meticulous about ethnographic data.

[21] For example, in the case of PECE, it is important to consider the foundational work of the Brazil-based lead developer, who acts as the primary maintainer of

the codebase, as well as the volunteer system admin who is based in Upstate New York. Exploring ways to attend to such nested levels of often invisible support infrastructural work is also part of more care-fully attending to our data practices.
[22] Further analytic questions to guide a reading of qualitative data platforms can be found here: https://worldpece.org/structured-analytics-questions-set/reading-qual-data-repository-platforms

[23] Refer to Poirier et al. (2020) for more on how meta-data can enable a remixing of data and perpetually proliferating interpretations of data.
[24] I uploaded his response as a public annotation here (https://www.researchdatashare.org/annotations/user/15/artifact/157), with his permission. I further added to it after receiving his response to my email.

[25] These changes/edits are ideally things that he could have done himself on the platform, but which I did not ask him to learn to do as I did not want to dissuade him from engaging with the material. There is indeed a level of technical comfort with the platform that is needed to be able to log in and insert an annotation by oneself. Therefore, I uploaded the annotation myself on his behalf (as our conversation was transpiring via email). This is a potential barrier to flag—the investment of time and effort to get over the technical learning curve for full participation.
[26] The division between "technicians" and "designers" of platforms, for example, could also lead to various forms of "invisibilization."

[27] Find the full annotation here: https://www.researchdatashare.org/annotions/question/165/artifact/121. The full transcript is available here: https://www.researchdatashare.org/content/transcript-191029001-being-researched-kibera

[28] I am attempting less common ways of running an ethnographic project fully aware of intersectional feminist work that has highlighted there is no politics of purity and no way of acting inside existing systems of injustices non-innocently.

[29] In a forthcoming paper, Leah Horgan and I seek to complicate the intersections of expertise and subject positions packed into the terms "global" and "local" to carve out space for the existence of "already global local" expertise and knowledge that transcend essentialized categories of nation, race, gender, ethnicity, and class while simultaneously accounting for the situated nature of all knowledges (Boellstorff, 2003; Haraway, 1991).

[30] Proceedings from the event, including video footage, distributed materials, and photos as illustrated in Figure 8, are available here: https://www.researchdatashare.org/content/proceedings-archiving-kenyas-past-and-futures.

[31] Transcript available here: https://www.researchdatashare.org/content/transcript-opening-remarks-archiving-kenyas-past-and-futures; PowerPoint slides available here: https://www.researchdatashare.org/content/OKUNE-mutuku-2019-opening-remarks-archiving-kenyas-past-and-futures-november-12-2019; and video recording available here: https://www.researchdatashare.org/content/video-proceedings-opening-remarks.

[32] An important and underinvestigated aspect of decolonizing knowledge infrastructures must include a political economic perspective on the academic publishing industry. See, for example, Posada and Chen's (2018) work on the rent seeking and financialization strategies of academic publishers. Find more also in a 2019 blog post I authored (Okune, 2019d).

[33] See for example Lassiter's work making audio archives; Anita Chang's community web site called Root Tongue; Fabiola Hannah's work on oral history and Lebanon; and Elaine Gan's use of methods from art, science and digital/environmental humanities to study the temporalities of multispecies interactions.

References

Aellah, G., Chantler, T., & Geissler, P. W. (Eds.). (2016). *Global health research in an unequal world: Ethics case studies from Africa*. CABI. https://doi.org/10.1079/9781786390042.0000

Arondekar, A. (2009). *For the record: On sexuality and the colonial archive in India* (1st ed.). Duke University Press. http://gen.lib.rus.ec/book/index.php?md5=5ef44c2b9a7d756d5efeea0e35363110

Atkinson-Graham, M., Kenney, M., Ladd, K., Murray, C. M., & Simmonds, E. A.-J. (2015). Care in context: Becoming an STS researcher. *Social Studies of Science*, *45*(5), 738–748. https://doi.org/10.1177/0306312715600277

Behar, R., & Gordon, D. A. (1995). *Women writing culture*. University of California Press.

Beltrán, H. (2020). The first Latina hackathon: Re-coding infrastructures from México. *Catalyst: Feminism, Theory, Technoscience*, *6*(2).

Beresford, A. (2016). Africa rising? *Review of African Political Economy*, *43*(147), 1–7. https://doi.org/10.1080/03056244.2016.1149369

Biruk, C. (2018). *Cooking data: Culture and politics in an African research world*. Duke University Press.

Boellstorff, T. (2003). Dubbing culture: Indonesian "gay" and "lesbi" subjectivities and ethnography in an already globalized world. *American Ethnologist*, *30*(2), 225–242.

Borgman, C. L. (2012). The conundrum of sharing research data. *Journal of the American Society for Information Science and Technology*, *63*(6), 1059–1078. https://doi.org/10.1002/asi.22634

Bowker, G. C. (2008). *Memory practices in the sciences*. MIT Press.

Christen, K. (2011). Opening archives: Respectful repatriation. *The American Archivist*, *74*(1), 185–210. https://doi.org/10.17723/aarc.74.1.4233nv6nv6428521

Clark, T. (2008). "We're over-researched here!": Exploring accounts of research fatigue within qualitative research engagements. *Sociology*, *42*(5), 953–970. https://doi.org/10.1177/0038038508094573

Clifford, J., & Marcus, G. E. (1986). *Writing culture: The poetics and politics of ethnography*. University of California Press.

CNN. (2012, June 15). *Nairobi's new hub of creativity*. https://www.cnn.com/videos/bestoftv/2012/06/15/inside-africa-nairobi-ihub-b.cnn

Comaroff, J., & Comaroff, J. L. (2012). Theory from the South: Or, how Euro-America is evolving toward Africa. *Anthropological Forum*, *22*(2), 113–131. https://doi.org/10.1080/00664677.2012.694169

Cornwall, A. (2007). Buzzwords and fuzzwords: Deconstructing development discourse. *Development in Practice*, *17*(4/5), 471–484.

Crane, J. T. (2013). *Scrambling for Africa: AIDS, expertise, and the rise of American global health science*. Cornell University Press. http://gen.lib.rus.ec/book/index.php?md5=ba8ecf1c10908522f39e0712b010f6bd

Davies, T., Walker, S., Rubinstein, M., & Perini, F. (Eds.). (2019). *The state of open data: Histories and horizons*. African Minds and International Development Research Centre.

de Souza Leão, L., & Eyal, G. (2019). The rise of randomized controlled trials (RCTs) in international development in historical perspective. *Theory and Society*, *48*(3), 383–418. https://doi.org/10.1007/s11186-019-09352-6

Deloria, V. (1988). *Custer died for your sins: An Indian manifesto*. University of Oklahoma Press.

Derrida, J. (1998). *Archive fever: A Freudian impression* (E. Prenowitz, Trans.). University of Chicago Press.

Elinoff, E. (2017). Towards an ethnography commons. *Commoning Ethnography*, *1*(1), 72–81. https://doi.org/10.26686/ce.v1i1.4126

Estalella, A., & Criado, T. S. (Eds.). (2018). *Experimental collaborations: Ethnography through fieldwork devices*. Berghahn Books.

Ferguson, R. A. (2012). *The reorder of things: The university and its pedagogies of minority difference*. University of Minnesota Press.

Fortun, K. (2009). Scaling and visualizing multi-sited ethnography. In M. A. Falzon (Ed.), *Multi-sited ethnography: Theory, praxis and locality in contemporary research* (pp. 73–86). Ashgate.

Fortun, M., Fortun, K., & Marcus, G. E. (2017). Computers In/and Anthropology: The Poetics and Politics of Digitization. In L. Hjorth, H. Horst, A. Galloway, & G. Bell (Eds.), *The Routledge Companion to Digital Ethnography* (pp. 11–20). Taylor & Francis.

Griffin, G., & Hayler, M. S. (2018). Collaboration in digital humanities research – Persisting silences. *Digital Humanities Quarterly*, *12*(1).

Haraway, D. (1991). Situated knowledges: The science question in feminism and the privilege of partial perspective. In *Simians, cyborgs, and women: The reinvention of nature* (pp. 183–201). Routledge.

iHub. (2011, January 7). **iHub_ Nairobi* [Internet archive]. https://web.archive.org/web/20110107035034/http://www.ihub.co.ke:80/pages/home.php#close

Irani, L. (2018). "Design thinking": Defending Silicon Valley at the apex of global labor hierarchies. *Catalyst: Feminism, Theory, Technoscience*, *4*(1). https://doi.org/10.28968/cftt.v4i1.29638

Jentsch, B., & Pilley, C. (2003). Research relationships between the South and the North: Cinderella and the ugly sisters? *Social Science & Medicine (1982)*, *57*(10), 1957–1967.

Kasdogan, D. (2018, April 10). The economization of life – Interview with Michelle Murphy. *IstanbuLab*. https://stsistanbul.org/2018/04/11/the-economization-of-life-michelle-murphy/

Khandekar, A., Kenner, A., Poirier, L., Morgan, A., Fortun, K., & Fortun, M. (in press). Infrastructuring collaborative STS: A middleware methodography. *Science & Technology Studies*.

Kukutai, T., Taylor, J., Australian National University, & Centre for Aboriginal EconomicPolicy Research. (2016). *Indigenous data sovereignty: Toward an agenda*. http://proxy.cm.umoncton.ca/login?url=http://www.jstor.org/stable/10.2307/j.ctt1q1crgf

Lassiter, L. E. (2005). *The Chicago guide to collaborative ethnography*. University of Chicago Press.

Liboiron, M., Zahara, A., & Schoot, I. (2018). *Community peer review: A method to bring consent and self-determination into the sciences*. https://doi.org/10.20944/preprints201806.0104.v1

Lovett, R., Lee, V., Kukutai, T., Cormack, D., Carroll-Rainie, S., & Walker, J. (2019). Good data practices for Indigenous data sovereignty and governance. In A. Daly, K. Devitt, & M. Mann (Eds.), *Good data*. Institute of Network Cultures.

Mahler, A. G. (2017). *Global South* [Data set]. Oxford University Press. https://doi.org/10.1093/obo/9780190221911-0055

Mannheimer, S., Pienta, A., Kirilova, D., Elman, C., & Wutich, A. (2019). Qualitative data sharing: Data repositories and academic libraries as key partners in addressing challenges. *American Behavioral Scientist, 63*(5). https://doi.org/10.1177/0002764218784991

Marchant, E. (2018). *Inspiring, legitimizing, restricting & resisting the many roles of narratives and storytelling among Kenya's technology innovators*. University of Pennsylvania.

Moore, N. (2007). (Re)using qualitative data? *Sociological Research Online*, *12*(3), 1–13. https://doi.org/10.5153/sro.1496

Murillo, L. F. R. (2018). What does "open data" mean for ethnographic research?: Multimodal anthropologies. *American Anthropologist*, *120*(3), 577–582. https://doi.org/10.1111/aman.13088

Murphy, M. (2015). Unsettling care: Troubling transnational itineraries of care in feminist health practices. *Social Studies of Science*, *45*(5), 717–737. https://doi.org/10.1177/0306312715589136

Murphy, M. (2017). *The economization of life*. Duke University Press. www.dukeupress.edu/the-economization-of-life

Neal, S., Mohan, G., Cochrane, A., & Bennett, K. (2016). "You cant move in Hackney without bumping into an anthropologist": Why certain places attract research attention. *Qualitative Research*, *16*(5), 491–507. https://doi.org/10.1177/1468794115596217

Nothias, T. (2014). "Rising," "hopeful," "new": Visualizing Africa in the age of globalization. *Visual Communication*, *13*(3), 323–339. https://doi.org/10.1177/1470357214530063

Okune, A. (2017, September 25). "Responsible research": Reducing risk or improving well-being? [Blog post]. *Critical Investigations into Humanitarianism in Africa*. http://www.cihablog.com/responsible-research-reducing-risk-improving-well/

Okune, A. (2019a). *Collaboration agreement*. https://doi.org/10.5281/zenodo.2550977

Okune, A. (2019b). *Data circulation form*. https://doi.org/10.5281/zenodo.2551093

Okune, A. (2019c, May 27). Org data interview guide_May 2019. *Research Data Share*. https://www.researchdatashare.org/content/org-data-interview-guidemay-2019

Okune, A. (2019d, May 29). Decolonizing scholarly data and publishing infrastructures. *Africa at LSE*. https://blogs.lse.ac.uk/africaatlse/2019/05/29/decolonizing-scholarly-data-and-publishing-infrastructures/

Okune, A. (2019e, November 12). *Proceedings of Archiving Kenya's Past and Futures*. Archiving Kenya's Past and Futures, Nairobi, Kenya. https://www.researchdatashare.org/content/event-archiving-kenya%E2%80%99s-past-and-futures-nov-12-2019/

Okune, A., & Fortun, K. (2019, June). Analytics across scales and systems: Qualitative data capacity. *Research Data Share*. https://www.researchdatashare.org/structured-analytics-questions-set/analytics-across-scales-and-systems-civic-data-capacity

Okune, A., Ochola Mak'Anyengo, P., Ndaiga, S. W., Ochieng, S., Omenya, R., Sambuli, N., & Sika, V. (2018). IHub Research (2011—2017): A Critical Technology Action Research Group within Nairobi's Flagship Tech Innovation Hub. In A. Khandekar & K. Fortun, *STS across borders digital exhibit*. Society for Social Studies of Science. https://stsinfrastructures.org/content/ihub-research-2011-2017-critical-technology-action-research-group-within-nairobis-flagship-1/

Pels, P., Boog, I., Florusbosch, J. H., Kripe, Z., Minter, T., Postma, M., Sleeboom-Faulkner, M., Simpson, B., Dilger, H., Schönhuth, M., Poser, A. von, Castillo, R. C. A., Lederman, R., & Richards-Rissetto, H. (2018). Data management in anthropology: The next phase in ethics governance? *Social Anthropology*, *26*(3), 391–413. https://doi.org/10.1111/1469-8676.12526

Pierre, J. (2013). *The predicament of blackness: Postcolonial Ghana and the politics of race*. University of Chicago Press.

Pink, S. (2014). Collaboration and its contentions. In *Collaboration: Organising for design impact and Value, programme for the Transdiciplinary Design Research Symposium, 7-9 April at ELISAVA, Barcelona School of Design and Engineering* (pp. 48–

50).

Pink, S. (2018). Afterword. In A. Estalella & T. S. Criado (Eds.), *Experimental collaborations: Ethnography through fieldwork devices* (pp. 201-212). Berghahn Books.

Pisani, E., Aaby, P., Breugelmans, J. G., Carr, D., Groves, T., Helinski, M., Kamuya, D., Kern, S., Littler, K., Marsh, V., Mboup, S., Merson, L., Sankoh, O., Serafini, M., Schneider, M., Schoenenberger, V., & Guerin, P. J. (2016). Beyond open data: Realising the health benefits of sharing data. *BMJ*, *355*, i5295. https://doi.org/10.1136/bmj.i5295

Poggiali, L. (2016). Seeing (from) digital peripheries: Technology and transparency in Kenya's Silicon Savannah. *Cultural Anthropology*, *31*(3), 387–411. https://doi.org/10.14506/ca31.3.07

Poirier, L. (2017). Devious design: Digital infrastructure challenges for experimental ethnography. *Design Issues*, *33*(2), 70–83. https://doi.org/10.1162/DESI_a_00440

Poirier, L., & Costelloe-Kuehn, B. (2019). Data sharing at scale: A heuristic for affirming data cultures. *Data Science Journal*, *18*(1), 48. https://doi.org/10.5334/dsj-2019-048

Poirier, L., DiFranzo, D., & Gloria, M. J. K. (2014). Light structure in the Platform for Experimental Collaborative Ethnography. *Web Science 2014 Workshop Interdisciplinary Coups to Calamities*.

Poirier, L., Fortun, K., Costelloe-Kuehn, B., & Fortun, M. (2020). Metadata, digital infrastructure, and the data ideologies of cultural anthropology. In J. Crowder, M. Fortun, R. Besara, & L. Poirier (Eds.), *Anthropological data in the digital age* (pp. 209–237). Palgrave Macmillan.

Posada, A., & Chen, G. (2018). Inequality in knowledge production: The integration of academic infrastructure by big publishers. In L. Chan & P. Mounier (Eds.), *ELPUB 2018*. https://doi.org/10.4000/proceedings.elpub.2018.30

Rayzberg, M. S. (2019). Fairness in the field: The Ethics of resource allocation in randomized controlled field experiments. *Science, Technology, & Human Values*, *44*(3), 371–398. https://doi.org/10.1177/0162243918798471

Reyes, V. (2017). Three models of transparency in ethnographic research: Naming places, naming people, and sharing data. *Ethnography*, *19*(2),204-226.. https://doi.org/10.1177/1466138117733754

Rooks, N. M. (2006). *White money/Black power: The surprising history of African American studies and the crisis of race in higher education*. Beacon Press. http://public.eblib.com/choice/publicfullrecord.aspx?p=3117983

Scott, J. W. (1991). The evidence of experience. *Critical Inquiry*, *17*(4), 773–797.

Silverman, S., Parezo, N. J., & Wenner-Gren Foundation for Anthropological Research (Eds.). (1992). *Preserving the anthropological record*. Wenner-Gren Foundation for Anthropological Research.

Smith, L. T. (1999). *Decolonizing methodologies: Research and Indigenous peoples*. Zed Books.

Stoler, A. L. (2009). *Along the archival grain: Epistemic anxieties and colonial common sense*. Princeton University Press.

Sukarieh, M., & Tannock, S. (2012). On the problem of over-researched communities: The case of the Shatila Palestinian refugee camp in Lebanon. *Sociology*, *47*(3), 494–508.

Tenopir, C., Allard, S., Douglass, K., Aydinoglu, A. U., Wu, L., Read, E., Manoff, M., & Frame, M. (2011). Data sharing by scientists: Practices and perceptions. *PLoS ONE*, *6*(6), e21101. https://doi.org/10.1371/journal.pone.0021101

Tilley, H. (2011). *Africa as a living laboratory: Empire, development, and the problem of scientific knowledge, 1870-1950*. University of Chicago Press.

Traynor, C., & Foster, L. (2017, October 7). Principles and practice in open science: Addressing power and inequality through "situated openness." *Natural Justice*. https://naturaljustice.org/principles-practice-open-science-addressing-power-inequality-situated-openness/

Tuck, E., & Yang, K. W. (2012). Decolonization is not a metaphor. *Decolonization: Indigeneity, Education & Society*, *1*(1).

Visweswaran, K. (1994). *Fictions of feminist ethnography*. University of Minnesota Press.

wa Thiong'o, N. (1993). *Moving the centre: The struggle for cultural freedoms*. J. Currey; Heinemann.

Yates-Doerr, E. (2019). Whose global, which health? Unsettling collaboration with careful equivocation. *American Anthropologist*, *121*(2), 297–310. https://doi.org/10.1111/aman.13259

Author Bio

Angela Okune is a doctoral candidate in the Anthropology Department at the University of California, Irvine. She studies data practices and infrastructures of research groups working in and on Nairobi, Kenya, in order to explore broader questions of equity, knowledge production, and socioeconomic development in

Africa. From 2010 to 2015, as co-founder of the research department at iHub, Nairobi's innovation hub for the tech community, Angela provided strategic guidance for the growth of tech research in Kenya. She was a network coordinator for the Open and Collaborative Science in Development Network (OCSDNet) (2014–2018), and co-editor for open-access book *Contextualizing Openness* (University of Ottawa Press). She currently serves on the Society for Social Studies of Science (4S) Council and as a Design Team member of the Platform for Experimental, Collaborative Ethnography (PECE).