Querying Collaboration: Extended Narrative
By: Angela Okune
October 2018

This orals document looks at how collaboration has been analyzed across a wide range of disciplines including informatics, organizational behavior, economics, and anthropology. Studying how collaboration is discussed across these diverse epistemic communities has helped me prepare for fieldwork encounters with researchers of diverse backgrounds to anticipate the wide range of discourse and positions related to research collaboration (at different stages of research).

The first half of this narrative includes explanation by James and I as to why we decided to use a heuristic of the "research life cycle" to categorize the works. The second half of the narrative is structured around insights gained from the process of reviewing and closely annotating a sub-set of the broader literature using an analytic structure.

"Research Lifecycle" as a heuristic for categorization
AO (July 2018): James and I decided to break the document across the research lifecycle as a way to put our different interests in collaboration in the same frame. I am particularly interested in collaborative formations in research fieldwork and investigating how other scholars have thought about ethics through the research process. Anthropologists have been long animated by how to get around long-standing issues related to the "othering" of research subjects and collaborative methods have been one way to do so. Development studies has also gone through cycles of participatory action research (and subsequent critiques of Participatory Action Research, etc.) so that literature also features here as well. James (as he explains below) was thinking about collaboration in a different vein so when we initially discussed working together (in April 2018), we were not sure how to put our distinct interests in the same frame. We realized that using the research cycle as our organizing framework enabled us to have a broad enough frame to fit both of our areas of interest but also specific enough to be productive.

(Update Aug 2018) AO: I came into this orals doc primarily interested in it from a question of ethics (as you can see in the initial way that I framed my descriptions above). However, over the course of developing the document together with James and doing the various readings from diverse disciplines (especially some of the organizational analyses of why groups work together), I realized that collaboration is much more of a core concept to my project than I originally envisioned. Given that my project entails working together with three research organizations to draw out their worries, desires, hopes, needs, requirements, and reservations regarding the sharing of qualitative research data, this is in fact not a project about studying data sharing practices, it is a study of collaboration! It just so happens that the specific kind of collaboration is related to data sharing (which as noted in this orals document is a largely understudied/ undertheorized area of collaboration in the research process). Therefore, I found that the literature spoke to me in unexpected ways that will be important for me to consider as I prepare for my fieldwork. For example, my role as a "convener" of the collaboration is one that I had not given enough thought to (Wood and Gray (1991) analyze the role of the convener in a collaborative formation extensively). Aellah et al. (2016) also provide me with helpful fictitious case study exercises that I may use as ice-breakers for the first focus group discussions as it could be easier to open up conversations about these issues using significantly different (and fictitious) examples instead of starting with real examples. Liboiron et al. (2018) also raised an important point about the key work of facilitation as well as taking ethnographic field notes,

leading me to mull over how I will realistically do both during the focus group discussion sessions that I have planned. This is something that I hope to discuss with my advisors both at UCI and in Nairobi.

JA (July 2018): As Angela discusses above, her interests in collaboration were initially primarily concerned with the ethical challenges of the researcher/researched relationship. My interests, by contrast, are more the epistemological challenges of facilitating collaborations across domains of expertise. What I am most interested in is how different ideas about the what constitutes proper data practices complicates efforts to co-produce a body of information that is intended to serve a common purpose. Ethics are, of course, still very much involved, as the ideas of propriety are shot through with both moral and epistemological concerns. However, this does mean that the literatures that I have been gathering have been primarily coming out of anthropology and STS, focusing on things like thought styles, boundary objects, and trading zones that get at the epistemological side of collaboration. Where I think there might be the most overlap is in our shared focus on data. It might be the case that I will be more concerned with the epistemologies of data ethics, and Angela will be more invested in the ethics of data epistemologies. Angela and I thought that organizing the sections on texts about collaboration at different stages of the research process would enable us enough flexibility to pursue our individual interests through the same structure.

BOTH (July 2018): We realized that a core shared area of our research projects is the double bind embedded in how our research interlocutors think about data. The double bind that Angela is concerned with is that on one hand the growing societal concerns about data privacy and protection (anonymity and lack of sharing is largely the default for IRB, for example) while on the other hand, there is growing concern (as expressed by open access advocates and other stakeholders like development practitioners) that scholarly data is a public good and we should all benefit from the data that is collected and produced by researchers (or at the very least, those that are studied should benefit). The double bind that James is concerned with is that some groups in Texas are concerned with the economic and technical complications of energy transitions. They have been hesitant and slow to move and largely focused on the economic data. Meanwhile, asthma rates are skyrocketing and other climate change damages and therefore other groups are urgently needing to move to ameliorate the energy transition in Texas. The groups are valuing different bodies of knowledge and practices and working with data and the question being tackled is which data models matter more or have more legitimacy/credibility. Do economic models matter more or do the health effects data matter more?

AO (August 2018): Given an unexpected turn of events towards the end of this collaborative endeavor, this extended narrative was written in the singular by Angela Okune. Thus "I" denotes Angela's thoughts. It is the hope that this narrative will be rewritten and added to at a future point by James Adams for re-submission as part of his own orals submission.

The next section is structured around insights gained from the process of reviewing and closely annotating a sub-set of the broader literature using an analytic structure. More details on collaboration as it pertains to each of the individual phases of the research life cycle can be found in the sub-essays (Research Design Analysis; Data Gathering and Production Analysis; Data Analysis Analysis; Artifact Production Analysis; Dissemination Analysis; Political Practice Analysis). It should be reiterated that the organizing into phases was not to suggest that the analysts were only dealing with research collaboration at that one particular phase but as a heuristic device to visualize more explicitly which part of the research processes analysts are

describing when they talk about "collaboration." Using this ordering has helped to more explicitly recognize gaps within processes of research where collaboration has not yet been heavily discussed.

Analysts within the annotated sub-set have thought of collaboration in a wide range of ways:
- Collaboration as cooperation (Handgraaf and van Raaji 2005)
- Collaboration as communication (Halabi et al. 2013)
- Collaboration as alliance between autonomous stakeholders using shared rules, norms, and structures to act or decide on issues related to a shared problem domain (Wood and Gray 1991)
- Collaboration as resource dependency (Wood and Gray 1991; Luukkonen et al. 1992)
- Collaboration for community self-determination of research harms (Liboiron et al. 2018)
- Collaboration as politics of difference (Fortun and Cherskaky 1998)
- Collaboration as productive tensions (between anthropologists) (Choy et al. (2009)
- Collaboration as co-working | Digital tools as collaborators (Griffin and Hayler 2018)
- Collaboration as a game of mutual deferral and appropriation (Holmes and Marcus 2008)
- Collaboration as synthesis (Cochrane and Cundill 2018)
- Collaboration as collective labor (Nayantara et al. 2018)
- Collaboration as shared conceptual labor (Holmes and Marcus 2008)
- Collaboration towards the production of multi-textual hermeneutics (Cervone 2015)
- Collaboration as co-authorship (Kaplan and Rose 1993; Luukkonen et al. 1992; Cerwonka and Malkki 2007)

Interpersonal relationships

Ethnographic relationships between individual researchers and interlocutors have highlighted potentials for collaboration (Marcus and Mascarenhas 2005), solidarity (Mohanty 2003), betrayal (Visweswaran 1994), deception (Bleek 1987), and intimacy (Blackwood 1995). Focusing on the concept of "collaboration," especially as it pertains to the research life cycle, James and I sought to better understand the underlying assumptions about sociality and what dynamics are said to be necessary for or emerge out of collaboration.

Communication and another more nuanced term, deliberation (Aellah et al. 2016), were mentioned and implied throughout many various works as important to thinking about successful interpersonal relationships and collaborations. However, little attention was paid to the underlying technical and data infrastructures that are needed to facilitate collaboration at many different stages (e.g. see here, here, here, here). Although I would have assumed that discussion of necessary sharing infrastructures (softwares, etc.) would have been particularly relevant at the data analysis stage, even the annotated piece looking at "collaborative synthesis" (Cochrane and Cundill 2018) focused on the need for strong inter-personal relationships. It could be that this analysis focused on inter-personal relationships to push back against optimistic technological solutionism that seek to use tech to build relationships (e.g. webinars rather than in-person conferences, etc.)

Choy et al. (2009) open up another way of thinking about collaboration as they work through the concept of collaboration with non-human actors such as the Matsutake mushroom. Griffin and Hayler (2018) also push the boundaries of normative understandings of collaboration as inter-personal, by arguing for the study of collaboration not only between humans and machines but also between machine to machine.

Requirements and reasons for collaboration

Within the annotated set, several works held that shared values are necessary to have meaningful collaboration (Nayantara et al. 2018; Halabi et al. 2013). Some analyses relied on dependency theory (Wood and Gray 1991; Luukkonen et al. 1992) and many believed that collaborations are entered into strategically (for each party's own benefit) (Liboiron et al. 2018; Wood and Gray 1991). Such theories also influenced analysts' thinking about why collaborators come together. Scholars like Luukonnen et al. (1992) held that cost-sharing was a key reason why international collaborators were sought, but the analysts then struggled to explain why more theoretical, non-lab-based scientific fields like mathematics had high levels of collaboration. The question of reciprocity and benefit sharing worried many analysts, especially those working in contexts with stark economic inequalities (Jentsch and Pilley 2003; Aellah et al. 2016).

Collaboration was thought to be important for several reasons including collective efficiency ("two minds are better than one"); diversity of perspectives and skills; and to address challenges given an increasingly complex world (Cervone 2005; Cochrane and Cundill 2018). There were differences across the analyses about whether shared values were a necessary pre-requisite (Halabi et al. 2013) or whether shared values emerge as a result of collaboration (e.g. Aellah et al. 2016).

Collaboration across the research life cycle

I find that much of the literature analyzing collaboration is missing discussion that outlines *how* the collaborative formation came to be. How and why the formation began and how it was solidified is one of the gaps in the discourse about collaboration. Halabi et al. (2013) begins to do this in their paper and is one of the few that talks explicitly about it. Other works such as Choy et al. (2009) are especially reflective on the nano level of analysis, focusing on why they pursued collaboration and what kinds of collaboration they believe to be most productive (that which sees "multivocality as a productive outcome of collaborations with each other"). However, interestingly, even such reflective works do not include any in-depth mention of the data sharing practices or technical infrastructures that undergird their collaborative experiments. Where did they store their collaborative writing drafts? Did they write in real-time or over bursts and spurts? How did they decided who would be listed as first author, versus second author, etc. Can all of them use the data for their own publication purposes or how did they decide who would use which data? Where did they store and analyze data from multiple fieldsites? Is any of it public? These questions continue to reverberate across most of the analyses of collaboration other than a handful that look explicitly at scholarly infrastructures (e.g. Kenner 2014).

Temporal rhythms of collaborations

As noted in reflections on our own collaborative endeavor, a key insight was that collaboration is not just a binary yes or no. Collaboration happens in spurts and starts, formally and informally, over time in more tightly coupled and loosely coupled instances. When does a "collaboration" begin? Wood and Gray (1991) study several different interorganizational collaborations which have more formal agreements. This brought me to ponder the transition from an informal to a formal "collaboration." In one of their case studies, Aellah et al. (2016) look at a collaboration that begun informally and initially worked well. After it became more formalized, one party felt that they were not getting as much benefit from the agreement as the other party.

Do strong collaborations emerge from formal MoUs and guidelines (tools advised by Jentsch and Pillay (2003), Aellah et al. (2016), and Wood and Gray (1991))? What do such formal mechanisms do to the relationship and dynamics of collaboration? Fortun and Cherkasky

(1998) were very interested in how collaborations shift over time and noted that organizations reposition themselves in response to new cultural forces and political-economic contexts. However, they do not explicitly discuss the moment of formalization and what that does to collaboration. Given Cochrane and Cundill (2018)'s main point that linking, whether formal or informal, does not automatically lend itself toward the desired outcome of collaboration, and examples given by Jentsch and Pilley (2003) and Crane (2010) who point to the inequalities perpetuated through research "partnerships," it seems important that, as recommended by Aellah et al. (2016), any formal documents or guidelines for collaboration be constantly revisited to ensure they reflect ongoing negotiations and changing dynamics of any collaboration.

Marginalization through collaboration

University incentive structures were mentioned severally by various pieces as structuring the terms of engagement, with Kaplan and Rose (1993) noting with shock that some universities assigned numerical values to scholarly publications of promotion and tenure candidates and then dividing the points by the number of co-authors. Unfortunately, this practice still appears to be used by several departments and Choy et al. (2009) and Holmes and Marcus (2008) write explicitly against thinking about collaboration as this kind of division of labor.

Importantly, Griffin and Hayler (2018) noted 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." Similarly, Fortun and Cherkasky (1998) find collectivity can not only be difficult to produce but also can be marginalizing and alienating. This highlighting of the differential stakes and gains from a collaborative formation is an important nuance and distinct from the point that Choy et al. (2009), Holmes and Marcus (2008); and Kaplan and Rose (1993) have made about the institutional disincentives to collaborate.

Information sharing

Jarvenpaa and Staples (2000) note that information sharing is a necessary part of knowledge management. While they are thinking information sharing in the context of organizations, I find the term interesting to play with when thinking about data collection, and recent feminist work on "informed refusal." Information sharing is an assumed part of research. But as the feminist work on refusal indicates (as well as older anthropological work on informants "telling lies" e.g. Bleek 1987), it may not be as straightforward or binary (i.e. information is shared or not). There is a lot of middle "wiggle" room where the power dynamics and assumed heirarchies of knowledge and expertise are nontheless played with. As Griffin and Hayler (2018) note through their examples of human-machine collaboration, the "human" usually thinks of themselves as acting on the machine but in fact, the machine also structures the way the human acts. Similarly, standard IRB protocols assume that the "good" research subject intrinsically desires to share all of their information. But in fact, my experience in Nairobi thus far demonstrates, research subjects appear to be shaping the interactions in intentional and explicit ways through what they ask for and what they tell the researchers. What do these dynamics in the "field" imply for information sharing amongst researchers? (How) are these dynamics captured and embedded or stored within the qualitative data? How can we probe the data and those who collected the data to better understand these dynamics?

Jarvenpaa and Staples (2000) noted that "as people's jobs and roles become defined by the unique information they hold, they may be less likely to share that information -- viewing it as a source of power and indispensability -- rather than more so." This is a phenomenon that I and

my interlocutors have also noted happening broadly in the Kenyan tech scene over time; with actors recognizing that their information, social capital and networks are becoming increasingly valuable to others (esp. many of the new arrivals to Nairobi who are trying to find their footing in Kenyan tech). As a result, those in Nairobi tech are less open with the sharing of their ideas and contacts. I anticipate this may also play out in how the various organizations decide what data to share or not share and how open they are with their data. If they perceive that there could be valuable data held by another party, then they might be more willing to share their own. It will be interesting to see how more economic model of "rational self-preservation" interacts with organizational commitments to sharing and openness and also moral claims to being "good" researchers and how those claims are influenced by ideas or perceptions about "international standards" and "local context."

My own project
Reading for this orals document coupled with presenting at the 2018 4S conference and explaining my interest in qualitative research data allowed me to also better distinguish the kind of "collaborative" formation that I will be studying. My assumptions are that the three different research groups I will do fieldwork with have very different epistemologies particularly because of their diverse disciplinary backgrounds (I expect mostly development economists and psychologists at Busara, social studies of tech and development practicioners at Map Kibera, and informatics and computer science at C4D Lab). I expect to work largely with Kenyan nationals, but those who are interacting with many non-Kenyan researchers (of different disciplines, nationalities, epistemologies as well). Therefore, the instance of collaboration I expect to look at is one where shared epistemological knowledge is not had before the attempted collaborative formation. The intended outcome of the collaborative formation is not necessarily shared values. I am interested in developing a better understanding of what it might take to foster collaborations across diverse disciplinary backgrounds and norms for dealing with data as well as academic structural incentives/legal infrastructures. Will this be a mutually shared outcome with the collaborators? If so, then the collaboration is formed around a shared problem space. But if that is not the factor that helps to enroll my interlocutors, perhaps the main "draw" or "value" will be the ability to say they are part of a project working on open data/open science (largely regarded by development funders as an intrinsic good); and/or to have an organizational archive (which can be private for their own team only if they so desire, or as "open" to the public as they would like).

Some of the researchers I interacted with at 4S did not appear to think qualitative research data sharing infrastructure is important to discuss. To them it appeared quite straightforward as to why qualitative researchers should not share data and was much harder for them to understand why we should share data. I am interested in observing if this is a similar challenge amongst the researchers I interact with in Nairobi or if alternatively they have received enough push-back already from those they work with to know why it is important to at least consider/think/test/experiment other ways to manage research data. Based on my initial fieldwork where nearly all that I spoke with about observations of sentiments of being "over-researched" resonated, I hypothesize that Nairobi-based researchers who do regular fieldwork with communities will be much more cogniscant of why current scholarly data infrastructures are inadequate.

I am also interested in further investigating how greater collaboration between researchers might affect dynamics "in the field." Given the scope of this project, it will not be possible to conduct extensive first-hand research of the field-based relationships between the

researchers and their subjects, but I plan to include probing questions about relationships on the ground as part of the focus group discussions that will be run over the duration of fieldwork.

Closing Thoughts

Like Choy et al. (2009), James and I did not take on the orals document collaboratively as a way to synthesize our separate, freestanding analyses into an organic, complementary whole or to divy up the labor of the orals document into two neat and distinct halves. Rather, we pursued a collaborative mode of "shared conceptual labor" (Holmes and Marcus 2008) towards producing the orals document in an attempt to gain greater insight into the processes involved in collaborative research and reflect on our own underlying assumptions about collaboration. I noted through the process of reading analyses of collaboration and doing it myself (continuing a trend of "maintaining an intertwined relationship between doing and researching" that Lindsay Poirer (2018) has observed about my own inclinations towards research practice) that I am more comfortable reaching agreement rather than maintaining productive tensions. This is an important lesson to be aware of, especially in preparation of the role I expect to soon be filling in the field as facilitator between at least three different research organizations with diverse epistemic cultures. The process of designing and executing this orals document has helped remind me that as ethnographer, I should aim not necessarily towards consensus, but rather seek to facilitate multivocality as a productive outcome of the collaborative exchange.

Invitation for Collaborative Elaborations

As I have discovered through the process of designing and implementing my orals documents on PECE, there is a wide range of possibilities that I have not yet had the time to explore. For example, I am only now beginning to take advantage of the tagging capabilities that allow for interesting new ways to bring diverse works together. The collaboration document began to use this by organizing the works not only by analytic questions but also by research life cycle (two intersecting lines of inquiry if you will). However, with greater work and time spent on refining the tagging categories, the works could also be pulled together around other emerging themes. For example, funding continued to be a topic that emerged severally in many of the papers. If I tagged some of the annotations with "funding" then doing so would pull those annotations together in one frame. I invite you to add any additional tags you see emerging throughout your reading; there is still much additional interesting scalar analysis that can be done on the work.