

Beyond Good Intentions: Data Sharing in and for Africa

REDIET ABEBE, HARVARD UNIVERSITY

KEHINDE ARULEBA, UNIVERSITY OF WITWATERSRAND

SARA KINGSLEY, CARNEGIE MELLON UNIVERSITY

GEORGE OBAIDO, UNIVERSITY OF WITWATERSRAND

SEKOU L. REMY, IBM RESEARCH - AFRICA

SWATHI SADAGOPAN, DELOITTE CANADA

1. INTRODUCTION

Data sharing is a complex issue involving multiple stakeholders and models of governance. The mannerisms and mechanisms by which data becomes accessible via data processes and data collectors (both *technical* and *human*) impacts the distribution of economic benefits and resources in societies (Okunoye and Sesan, 2018). In economies that are heavily reliant on (information) services or financial markets, the ability to harvest and accumulate data is an important source of economic growth. Firms capitalize on the data they collect to allocate *services* or *commodities* based on data inferences (Okunoye and Sesan, 2018). Those who are able to monopolize and accumulate vast amounts of data are at an advantage, whether in the trade of data about things and people for monetary value; or, financial transactions in stock exchanges; or, in the sale of goods and services that rely on information about demand and supply to optimize profits.

Economies that have fewer data sharing capacity (or power over) data infrastructures are at a sizable disadvantage. Outside data extractors may own and determine the data that is available in the economy (Coleman, 2019), possibly making data relatively scarce. When data is scarce or not public, data access is typically limited (i.e., the data is collected or collect-able, but privately kept), and therefore, affordability is problematic (i.e., because data is private, access is costly, and only available to those who can buy it).

Africa, as a continent, is home to diverse and rich communities; histories; traditions; cultures; languages; and values. In recent years, the African continent, as a whole, has been considered a “frontier” opportunity (Cisse, 2018; Hao, 2019). This, some argue, alludes to a colonialism for powerful multinationals and data brokers who seek a stake in data economies in Africa, where data markets are expected to have the largest *worldwide* economic growth, and the most significant role to play in the global data economy in the near future (Birhane, 2019; Coleman, 2019). Multinationals have therefore begun (or expanded long-term) projects to capitalize data markets in the region. In the meanwhile, globalized standards, norms of practice, and regulations for data access and sharing (including privacy and security) are not agreed upon (Coleman, 2019). Hence, the practices of data economy stakeholders are sometimes divergent, ad hoc, contradictory, and/or violate community desires or hopes, let alone values and ethics for data-use.

While these issues prevail in places and communities across the globe, communities across the African continent face dynamics worth evaluating and learning from. However, a notable trend is that studies in the literature tend to primarily focus on data economies of the Global North. This fact precludes the dynamic we seek to address in our work; namely, how are data practices -- the collecting, sharing, and ownership of data rights -- determinant of where power is centered (and for whom) in the global data economy? How do the histories of uneven power entrench legacies of disparity, colonialisms, ethnocentrism, slavery, national and racial superiorities -- as well as patriarchal values -- into data transactions in and for Africa at this time? Histories (re)generate across centuries, through the dynamics of economies, and determine who has power, and who reaps the most rewards in terms of wealth. For this reason, we argue it is our ethical responsibility to interrogate patterns in how power is decided, and to make sure our idea of which things deserve addressing do not replicate historical oppressions that we wish to eliminate from our futures. Part of this power is what we name in our studies.

The 54 distinct African countries offer a wealth of previously untapped and potentially valuable data sources. But the history of power in these countries is mirrored by the power dynamics between data collectors and local stakeholders (Taylor, 2015; Andrejevic, 2014). This fact challenges us to imagine a present, and future, characterized by *data colonialism* (Birhane, 2019; Coleman, 2019); where African individuals, communities, and entire heterogeneous geographies of people, have their data accessed, harvested and shared, but do not reap the same rewards from data transactions as the extractors/beneficiaries (Benjamin, 2019).

Hence, in our work, we investigate what power stakeholders have. We consider sources of inequality among data collectors and human subjects and communities (i.e., data targets) to learn why misalignment occurs between practices and norms. Instead, we argue that data extraction efforts that wish to do good in the world can intentionally or unintentionally cause harm, especially when community values are not familiar or incorporated into the design of practitioners' work. We believe this is why it is important to document what community norms are, and design guidelines through the research process. Our motivating research questions are: *How do data sharing practices align with the norms and values of communities in places across the African continent? What values and norms should govern data practices in the future?* In relation, we consider, and ask, how do different African communities protest, and take actions to control the rights of local data? We examine these questions through stories from interviews that were conducted to learn about the tactics people employ to resist, insist on, and protect community standards for data governance. From the interviews, we intend to learn what factors shape participation and outcomes in local/global economies. We evaluate how power structures affect whether a community benefits from transactions involving the (mis)use of their data. Ultimately, we seek to illuminate and contribute stories of resistance against multinational stakeholders who exert undue influence on data infrastructures in Africa. Through these stories, we wish to inform the practices of stakeholders about the norms and values of the communities that are most impacted by their practices. Through our research, we argue for the design of open, consent-based systems, in which the voices of local stakeholders are empowered, have precedence, and take a central role in the development of data processes and standards.

To achieve our goals for this work, we build on our previous research (Aruleba et al., 2019), and investigate these questions:

RQ1: What data sharing barriers exist in the data ecosystem, and for whom?

RQ2: How might infrastructures lose credibility among intended beneficiaries?

RQ3: What creative forms of resistance arise when implied trust is, or appears, compromised in data economies?

In this paper, we explore our primary research questions from the lens of “humanism” in human-computer interaction (Bardzell and Bardzell, 2016). As it is called, humanistic HCI leverages theories and concepts from the humanities for the purpose of building or designing systems and processes (Bardzell and Bardzell, 2016). This type of work may take methodologies from the humanities for the purpose and in the service of design research (Bardzell and Bardzell, 2016). Accordingly, we leverage interviews and storytelling to diagnose the issues (“*economic inequality*”, “*colonialism*”, “*uneven bargaining power*”) germane to our research question (“*how could we collect and share data respectfully?*”).

In the tradition of critical design, we offer an alternative narrative (“theory of power”) to the normative assumption that promotes the idea that data is (equally for everyone) the “new oil” of economic growth and development (Hummel et al., 2018). We follow the ethos of work in critical design that seeks to “make consumers more critical about their everyday lives, and in particular, how their lives are mediated by assumptions, values, ideologies, and behavioral norms inscribed in designs” (Bardzell and Bardzell, 2013). Finally, we rebut the trend reflected by what is named in the literature, that the data economy is exclusively a priority of and domain for stakeholders of the Global North. Our central thesis is that historical and contemporary differences in power held among individuals, communities, and nations affect outcomes and who benefits from transactions in the data economy. We argue that historical patterns of oppression, inequality and injustice are reflected by these power differentials, and hence, in data transactions.

Similar acts along the data creation and use pipeline have been studied among other geographic contexts; e.g., Myanmar (Goldstein and Faxon, 2018) and in other contexts, including e.g., deploying data to preserve privacy (Brunton and Hissenbaum, 2015). Others including Li et al. (2019) explore protest uses of technology, in which users actively resist technological exploitation through deleting or intentionally opting not to use software. Other forms of technological protest included installing ad blocking technology to subvert or disrupt business models of companies. While the authors focus on users in the Global North, they also highlight the need to understand protest technology usage in other regions of the world, and why different communities engage in these practices. In our paper, we respond to this call by collecting and analyzing data about data practices (and resistances) in and for communities across the African continent.

The full version of this work includes discussions of methodologies employed including storytelling and semi-structured interviews. We present 6 fictional personas as well as a discussion around common themes emerging from the interviews: lack of infrastructure for data sharing and procurement, lack of education on the value of data, lack of trust among different stakeholders, and coordination and policy limitations especially in cases where data needs to be shared across companies and countries.

REFERENCES

- Jeffrey Bardzell and Shaowen Bardzell. What is “critical” about critical design? In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '13)*. Association for Computing Machinery, New York, NY, USA, 3297–3306. 2013.
- Jeffrey Bardzell and Shaowen Bardzell. Humanistic HCI. In *ACM interactions* 23, 2. February 2016, pp. 20–29.
- Nicola Bidwell, Thomas Reitmaier, Gary Marsden, and Susan Hansen. “Designing with mobile digital storytelling in rural Africa.” In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 1593-1602. 2010.
- Abeba Birhane. “The Algorithmic Colonization of Africa: Startups are importing and imposing AI systems founded on individualistic and capitalist drives.” *Real Life Magazine*, July 18, 2019.
- James Bòdé Agbájé. Proverbs: A Strategy for Resolving Conflict in Yorùbá Society. *Journal of African Cultural Studies*, Vol. 15. No. 2, pp. 237-243. December 2002.
- Finn Brunton and Helen Nissenbaum. *Obfuscation: A user’s guide for privacy and protest*. MIT Press, 2015.
- Susan Bull, Phaik Yeong Cheah, Spencer Denny, Irene Jao, Vicki Marsh, Laura Merson, Neena Shah More, Le Nguyen Thanh Nhan, David Osrin, Decha Tangseefa, et al. Best practices for ethical sharing of individual-level health research data from low- and middle-income settings. *Journal of Empirical Research on Human Research Ethics*, 10(3):302–313, 2015.
- Moustapha Cisse. Look to Africa to advance artificial intelligence. *Nature* 562, 2018.
- Danielle Coleman. “Digital Colonialism: The 21st Century Scramble for Africa through the Extraction and Control of User Data and the Limitations of Data Protection Laws.” *24 Michigan Journal of Race & Law* 417 (2018-2019).
- Asebe Regassa Debelo, Abiyot Legesse, Tema Milstein and Ongaye Oda Orkado. “‘Tree is Life’: The Rising of Dualism and the Declining of Mutualism among the Gedeo of Southern Ethiopia.” *Frontiers in Communication*. 21 August 2017.
- Spencer G Denny, Blessing Silaigwana, Douglas Wassenaar, Susan Bull, and Michael Parker. Developing ethical practices for public health research data sharing in south africa: The views and experiences from a diverse sample of research stakeholders. *Journal of Empirical Research on Human Research Ethics*, 10(3):290–301, 2015.
- Google, Policy Division. “From Big Data to Big Social and Economic Opportunities: which policies will lead to leveraging data driven innovation’s potential?” In *The Global Information Technology Report* of the World Economic Forum: 2014.
- Karen Hao. The future of AI research is in Africa. *MIT Technology Review*, 2019.
- Eric Horvitz and Deirdre Mulligan. Data, privacy, and the greater good. *Science*, 349(6245):253–255, 2015.
- Patrik Hummel, Matthias Braum, Steffen Augsburg and Peter Dabrock. “Sovereignty and Data Sharing.” In *International Telecommunications Journal of ICT Discoveries*, Special Issue 1(2), 23 November 2018.
- Bartha M Knoppers, Jennifer R Harris, Isabelle Budin-Ljøsne, and Edward S Dove. A human rights approach to an international code of conduct for genomic and clinical data sharing. *Human Genetics*, 133(7):895–903, 2014.
- Patty Kostkova, Helen Brewer, Simon de Lusignan, Edward Fottrell, Ben Goldacre, Graham Hart, Phil Koczan, Peter Knight, Corinne Marsolier, Rachel A McKendry, et al. Who owns the data? open data for healthcare. *Frontiers in public health*, 2016.
- Trudie Lang. Advancing global health research through digital technology and sharing data. *Science*, 331(6018):714–717, 2011.
- Oitshpile Mmab Modise, Rebecca Nthogo Lekoko, Olekae Tsompi Thakadu and Masego Ayo Mpotokwane. Toward sustainable conservation and management of human-wildlife interactions in the Mmadinare Region of Botswana: villagers’ perceptions on challenges and prospects. In *Human-Wildlife Interactions*, 12(2): 220-232. Fall 2018.
- Hanlin Li, Nicholas Vincent, Janice Tsai, Jofish Kaye and Brent Hecht (November 2019). “How Do People Change Their Technology Use in Protest?: Understanding ‘Protest Users.’” *Proceedings of the ACM on Human-Computer Interaction*.
- Moses Namara, Daricia Wilkinson, Byron M Lowens, Bart P Knijnenburg, Rita Orji, and Remy L Sekou. Cross-cultural perspectives on ehealth privacy in africa. In *Proceedings of the Second African Conference for Human Computer Interaction: Thriving Communities*, page 7. ACM, 2018.
- Ndana Ndana. Subiya “Traditional Literature”: A Preliminary Survey. *Botswana Notes and Records*, Vol. 38, pp. 99-114: 2006.
- Hagos Nigussie. Indigenous Communication Forms and Their Potential to Convey Food Security Messages in Rural Ethiopia. *Indian Journal of Human Development*. 10(3): 414-427.
- Lexi Novitske. *Stanford Social Innovation Review*, (2018).
- Babatunde Okunoye and Gbenga Sesan. “Data for Good in Africa.” In *International Telecommunications Union (ITU) Journal of ICT Discoveries*, Special Issue 1(2), November 16, 2018. ISSN: 2616-8375. <https://www.itu.int/en/journal/002/Pages/05.aspx>
- Francois Van Schalkwyk, Sumandro Chattapadhyay, Michael Cañares, and Alexander Andrason. Open data intermediaries in developing countries. 2015.
- Jathan Sadowski. “When data is capital: Datafication, accumulation, and extraction.” *Big Data & Society* January–June 2019.