Beyond Good Intentions

Data Sharing in and for Africa

Rediet Abebe, Harvard University
Kehinde Aruleba, University of Witwatersrand
Abeba Birhane, University College Dublin
Sara Kingsley, Carnegie Mellon University
George Obaido, University of Witwatersrand
Sekou L Remy, IBM Research - Africa
Swathi Sadagopan, Deloitte Canada

Introduction

➔ Data sharing is double-edged sword:

◆ With the rise of ML systems comes the need for quality data

◆ Improves productivity and efficiency in the broader economy
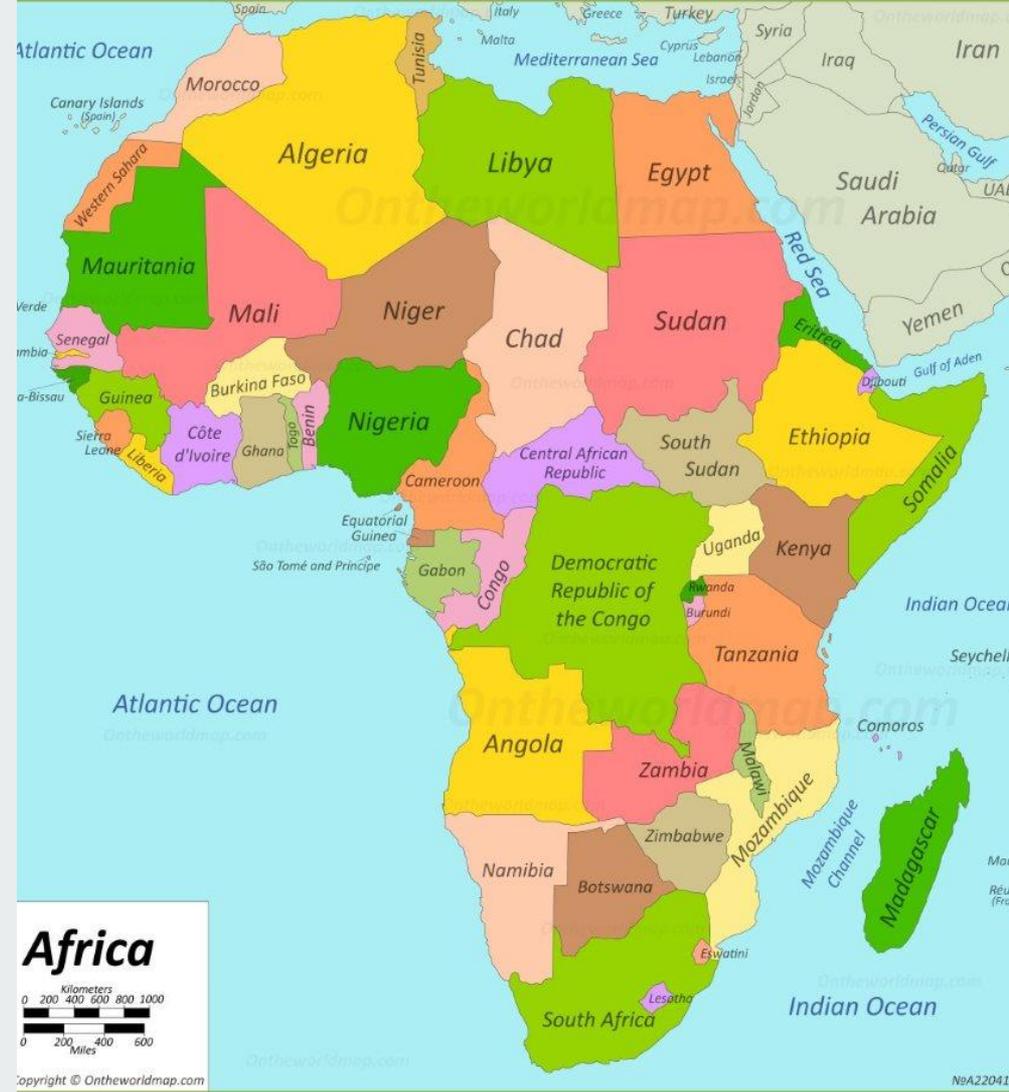
However, data sharing in a responsible and respectful manner is difficult

- Who benefits from data sharing?

- Who is exploited?

- What kind of data is shared?

- Who has access?

- Who is telling the story?

- What are the economic returns to communities?

Research Questions

1. What data sharing barriers exist in the data ecosystem, and for whom?
2. How could we share data responsibly and respectfully?

Data Tropes

- Solutions to all problems rests on data
- AI can "leapfrog" the continent into development
- "Backward" and uneducated

The AI Invasion is Coming to Africa (and It's a Good Thing)

Before Africa can reap the benefits of artificial intelligence, African governments, investors, and NGOs must train workers for complex tasks, and reform laws and education to meet the demands of tomorrow's economy.

SHARE   COMMENT   PRINT   ORDER REPRINTS

By Lexi Novitske | Feb. 12, 2018

For many countries, the prospects of artificial intelligence (AI) are thrilling. They conjure up the kinds of innovations we see in science fiction. In Africa, however, the dawn of AI carries with it a fear of falling further behind more-developed economies, rather than the eager anticipation of new technology—the World Economic Forum predicts a net loss of five million jobs to AI worldwide by 2020.

Financial issue. Their backward economic situation has always been the biggest problem for developing countries, especially in rural areas. Dalaba et al. studied the costs associated with implementation of a computer-assisted clinical decision support system for antenatal and delivery care in Northern Ghana. They observed a

B Brookings Institution

The Fourth Industrial Revolution and digitization will transform Africa into a global powerhouse

The spread of digital technologies can empower the poor with access to ... (For more on the promise of artificial intelligence in Africa, see the ...

Jan 8, 2020

E Entrepreneur

Artificial Intelligence Is Filling The Gaps In Developing Africa

Pockets of development. In widening the equality gap, the report also stated that the incomes of the poorest 10% of the population increased by ...

Jul 25, 2019

CIO

How AI/ML is helping to eradicate poverty

How AI/ML is helping to eradicate poverty ... and day-light images of many areas of Africa (Rwanda, Nigeria, Uganda, Malawi and Tanzania) ...

Jun 18, 2019

Methodology

➔ We employ *storytelling* to address issues, such as:

   ◆ Economic inequality

   ◆ Colonialism

   ◆ Uneven bargaining power

➔ Stories allow us to tell narratives and express challenges and protect the identities of interview/data subjects.

Why storytelling

- Protects data and research subjects

- Allows us to engage in sensitive/critical conversations with broader

  audience

- Focuses on local narratives from perspective of the African

Power of storytelling

I always think about the data as story, and each person who contributed to that data as storytellers. What is our responsibility to the story and our responsibility to the storyteller? Those are all indigenous concepts, that we always care for our storytellers, and we always have a responsibility to our stories.

Abigail Echo-Hawk

Challenges

- Infrastructure (lack thereof)
- Funding
- **Trust**
- Awareness

- Data Regulation (lack thereof)
- Power imbalances
- **Data colonialism**

Data Colonialism

Decolonizing data means that the community itself is the one determining what is the information they want us to gather. Why are we gathering it? Who's interpreting it? And are we interpreting it in a way that truly serves our communities? Decolonizing data is about controlling our own story, and making decisions based on what is best for our people. That hasn't been done in data before, and that's what's shifting and changing.

Abigail Echo-Hawk

Drink US West LLC

Discard for local norms

Lack of regulation

While DrinkUS's research was well-intentioned, the NGO failed to comprehend the consequences of sharing their research findings publicly, including their predictions about which geographic locations in the yetet'ebek'e community were at risk for water-related disease. Such findings, even shared in aggregate, can reveal information on the community as a whole. While notions of privacy often focus on the individual [7, 10], there is growing awareness that collective identity is also important within many African communities, and that sharing aggregate information about communities can also be regarded as a privacy violation [12]. Global North agents and organizations, whether governments or multinational firms, exert an out-sized influence in the global data economy. For the many African data economies, this can translate to issues around lacking access to data collected by Global North stakeholders, not reaping the same rewards and benefits as non-locals, and having to trade autonomy in decision-making for the marginal financial incentives offered by Global North agents.

Benefited Global North stakeholders

Soil and apartheid

Trust

Soil and apartheid. A PhD candidate from Bozatta conducting research in Nova Africa wants to get soil samples across a few Nova Africa provinces for her research. In her efforts to do so, she learns that farmers are not enthusiastic about sharing this information, but she is unable to understand why. After months of attempts, she eventually learns from her neighbors that farmers may not divulge this information to researchers due to fear that the Government might want to claim ownership of such lands. Even for local researchers, this issue might not seem to be the case. This issue was exacerbated by the apartheid regime, when land grabs were widespread, and black Nova Africans were mostly affected. Understandably when sharing data has led to the forfeiture of rights, access becomes the price of histories, as in most locations it is.

Future work

Continue to conduct interviews and construct personas

Offer guidelines for context - and - human-centered data practices

Research gaps between policy + practices

Thank you!