Privacy and the cost of *solutionism*
Miguel Gama Nunes / 2595557

Introduction

The right to privacy is a widely debated and valued principle across developed societies. It is a principle that spans disciplines and represents an individual freedom celebrated in the Universal Declaration of Human Rights. The advent of the internet, social media and widespread usage of Big Data to advance private corporate interests brought this concern to the forefront of public debate. It is now well understood to what extent large platforms like Google, Facebook and Twitter are able to harness data and build profiles of its users, as well as their capacity for monetizing it. The ubiquity of mobile phones — by 2017, there were an estimated seven billion mobile phones around the world [1] — made it possible for such companies to commoditize their users' profiling and attention in unprecedented ways. Social media platforms are the classical example for demonstrating the potential and profitability of data markets and advertisement micro-targeting, but they are far from being the only players utilizing user and customer data to gain further insights and participating in the global data market. We verify that by 2019 the Big Data industry's market size as a whole amounts to 49 billions, with projections of doubling that value by 2027[1].

Given that user data (personal information, actions, preferences) is the essential fuel for most of Big Data, it is then natural that concerns of privacy breaching, along with potential negative impacts that it might have on individuals as well as communities started arising alongside with data proliferation. Legal frameworks were developed to constrain how such data is harnessed, kept and most of all exchanged. GDPR in Europe is such an example of an ambitious legal framework that aims at protecting consumers from undiscriminated data collection, adding emphasis on transparency towards users and corporate accountability of wrong-doing. As digital literacy increases in the developed world and policies are designed as we collectively understand the implications behind our actions online, we witness however a whole different scenario in Low and Middle Income Countries (LMICs). While all but five of these countries had explicit mentions of privacy in their constitutions, by 2013 in Sub-Saharan Africa only 8 states out of 44 had any specific data protection laws [5].

Digitization landscape in LMICs

> *If the purpose of development is to empower those in developing countries to have access to the same rights and capabilities as those in the developed world, and if the transfer of knowledge and technology is essential to that purpose, then why diminish those very same people by granting them lesser human rights protections? If privacy and the protection of personal information are essential as constitutional and human rights protections in developed societies, this must also be true in developing countries.* [2]

The state of digitization in developing countries is very distinct from the state of widespread usage of technology that developed countries are accustomed to. We verify that the growth of technological infrastructure goes hand in hand with privately developed projects, and thus economically struggling countries logically lack the investment to make it possible. This is especially true for Big Data and data analytics projects, in which the majority of the technology and human resources are being employed by private corporations around the world.

LMICs are however a largely unexplored market for most technological companies and therefore represent an unprecedented opportunity for expansion. There are several projects from well-known corporations that embrace this opportunity:

IBM's *Lucy* project started in 2014 proposing to "solve Africa's grand challenges including healthcare, education, water and sanitation, human mobility and agriculture"[3]. Their method consists of creating a consortium comprised of universities, governments, business leaders and startup communities to collect an unprecedented amount of economic and social data, which can be then be fed into *Watson* for insights into the most pressing political, social, economic problems in the region.

---

[1] https://www.statista.com/statistics/254266/global-big-data-market-forecast/

Facebook inaugurated the internet.org project in 2013 with the goal of providing connectivity to "two thirds of the world's population that doesn't have it" by combining the use of "drones, satellites and lasers to deliver the internet to everyone", as stated by Mark Zuckerberg. The caveat being that users have free access to a collection of pages made available by Facebook and a number of selected partners.

On the other hand, there's a number of high profile IT projects being developed within public-private agreements. As mentioned before, the expertise and infrastructure required for large-scale data projects are not available for the majority of LMIC governments, and therefore it is common to turn to partner corporations. A high-profile example is the Aadhaar project, or UIDAI (Unique Identification Authority of India). This was deployed in 2008 as a large-scale effort to attribute an ID to every citizen of India. The ID is attributed based on biometric data of citizens — iris scan and fingerprint — generating a unique 12-digit unique number. The organization overseeing the project is managed and chaired by the co-founder of Infosys, one of India's largest tech consulting firms. He has made clear since inception that the project had a clear potential for targeted advertising and a commercial motivation at its core.

In 2013, USAID put forward a 53 million dollar project that, among others, helped fund the creation of national identity cards in Kenya, as well as set up the M-PESA mobile money project together with Vodafone. Kenya also received substantial funding to digitize and improve their electoral system.

Both Aadhaar and Kenya's projects brought around significant problems in that they are not socially inclusive. Iris scan properties erode with age, making it harder to register elderly citizens, and the use of fingerprints excludes a large part of the population who, due to a lifetime of manual labour no longer have marked finger traits. This is especially relevant since these people are more likely also the poorest, the ones the system is aiming to aid in the first place. Aadhaar is intended to identify people for a large spectrum of services provided by the state apparatus, including welfare and rationing for families in extreme poverty. However, the system is designed to deliver the family ration to one of its members, and several authors worry about the access to food when this person is not available. As far as research points, there are no systems in place to overcome this limitation, and more worryingly the UIDAI doesn't have any process in place that helps such families or holds departments and local agents (who are allocated a specific region) accountable. According to Taylor [1], "there are reports of growing malnourishment amongst families excluded by the database".

In the absence of nation-wide data regarding important economic and social factors in developing countries, such initiatives are most certainly laudable since information is power, and in order to tackle widespread problems such as disease spreading, persistent poverty, malnourishing and many others it is essential to have a clear representation and assessment of the state of society across dimensions.

Projects such as IBM's Lucy are often perceived as an appropriate replacement for, e.g. national surveys, but many authors have grave concerns about using forms of Big Data in place of national official surveys. These authors mention several dimensions that are at stake, and I enumerate and elaborate on a few of them.

Transparency

National surveys are volunteer based and present clear goals to the interviewees. When conducting a survey, citizens are asked explicit questions — thus being in control of what dimensions of their private lives is shared — and are informed of the purpose of the study and measures that will be taken to ensure the privacy and integrity of this data. LMIC countries and aid organizations such as the UN are increasingly resorting to data collected by private corporations as a proxy for such surveys, thus creating *data doubles* —"abstracted representations of people through their data, formed with the aim of monitoring or targeting people for intervention" [5] — for citizens that may or may not be representative of their status, but most importantly may include data or insights that he/she would not volunteer or would like to be aware of.

Inherent bias

Using privately-acquired data as a proxy for survey data inherently brings bias to the created models, since it samples the part of population which is digitally connected. As Shearmur states [4], those who use Big Data to study behavior or shape policy are not seeing society but 'users and markets'. Taylor and Broeders [5] describe a clear example of how this intrinsic bias can

hinder decision-making. They describe how an innovative African mobile calling dataset was used to guide Côte d'Ivoire's transport policy, and how they verified that the conclusions of what factors were impeding traffic flows were essentially incomplete.

Solving social problems as engineering ones

Several authors stress how smaller scale organizations, qualitative researchers and local activists are important for tackling problems in regions with specific social and cultural characteristics. These agents are in the field and serve as extensions of state bodies, and in a lot of cases have expert and domain knowledge on how to best tackle the problems. By focusing on larger-scale "engineering" solutions, "it may divert power, money and resources from these smaller-scale actors" **[5].** Furthermore, this kind of *solutionism*, a term coined by Morozov in **[6]** may ignore that certain problems are structural and long term in nature, such as problems indirectly caused by corruption, social stratification and discrimination.

Collateral effects of widespread data collection and sharing

Lastly, there are concerns that collecting, processing and sharing data indiscriminately can lead to a larger segmentation of the population and ultimately lead to discrimination. **[2]** provides an example of Rwanda's national identification department, in which the authors argue that although no ethnic information is stored at the moment, the constructed identities of "Tutsi" and "Hutu", which were used to promote discrimination, show an example of how possible would be to "imagine categories of identities relating to fingerprints being similarly constructed and used to the advantage of political or criminal groupings" **[2]**.

Conclusion

This essay presented some insights from important work done in the field of data justice and ICT4D privacy concerns. It presented three landmark *development* projects carried out by large private corporations and some of the concerns that span from there. Experts in the field have a myriad of other concerns such as how such development feeds surveillance or how the definition itself of *development* and *aid* may ultimately not be shaped in the best interest of the populations it intends to help.

In this essay, I mainly tackled some aspects of *solutionism* and collateral damage of applying Big Data solutions to large, structural and long term problems such as bias in the data due to reflecting connected people, solving problems with an engineering approach and discrimination that may arise from superficial collection of data.

Lastly, I refer to the work of Linnet Taylor in **[1]** and to the importance of supporting the creation of a framework that guides sustainable *development* projects. In this fast-paced field, it is of the utmost importance that companies and institutions are accountable and that public bodies can evaluate how socially and economically sustainable the projects being developed are.

References

[1] Taylor, Linnet. "What is data justice? The case for connecting digital rights and freedoms globally." Big Data & Society 4.2 (2017): 2053951717736335.

[2] Hosein, Gus, and Carly Nyst. "Aiding surveillance: an exploration of how development and humanitarian aid initiatives are enabling surveillance in developing countries." Available at SSRN 2326229 (2013).

[3] IBM, 2014. Project Lucy: How IBM Research is engaging in Africa. White paper, IBM, New York. Accessed 25.08.14 at <http://asmarterplanet.com/files/2014/03/Project-Lucy-High-Res.pdf>

[4] Shearmur, Richard. "Dazzled by data: Big Data, the census and urban geography." (2015): 965-968.

[5] Taylor, Linnet, and Dennis Broeders. "In the name of Development: Power, profit and the datafication of the global South." *Geoforum* 64 (2015): 229-237.

[6] Morozow, Evgeny, and Evgeny Morozov. To Save Everything, Click Here: Technology, Solutionism, and the Urge to Fix Problems that Don't Exist. Allen Lane, 2013.