Exploring the Future of Data Governance in Africa
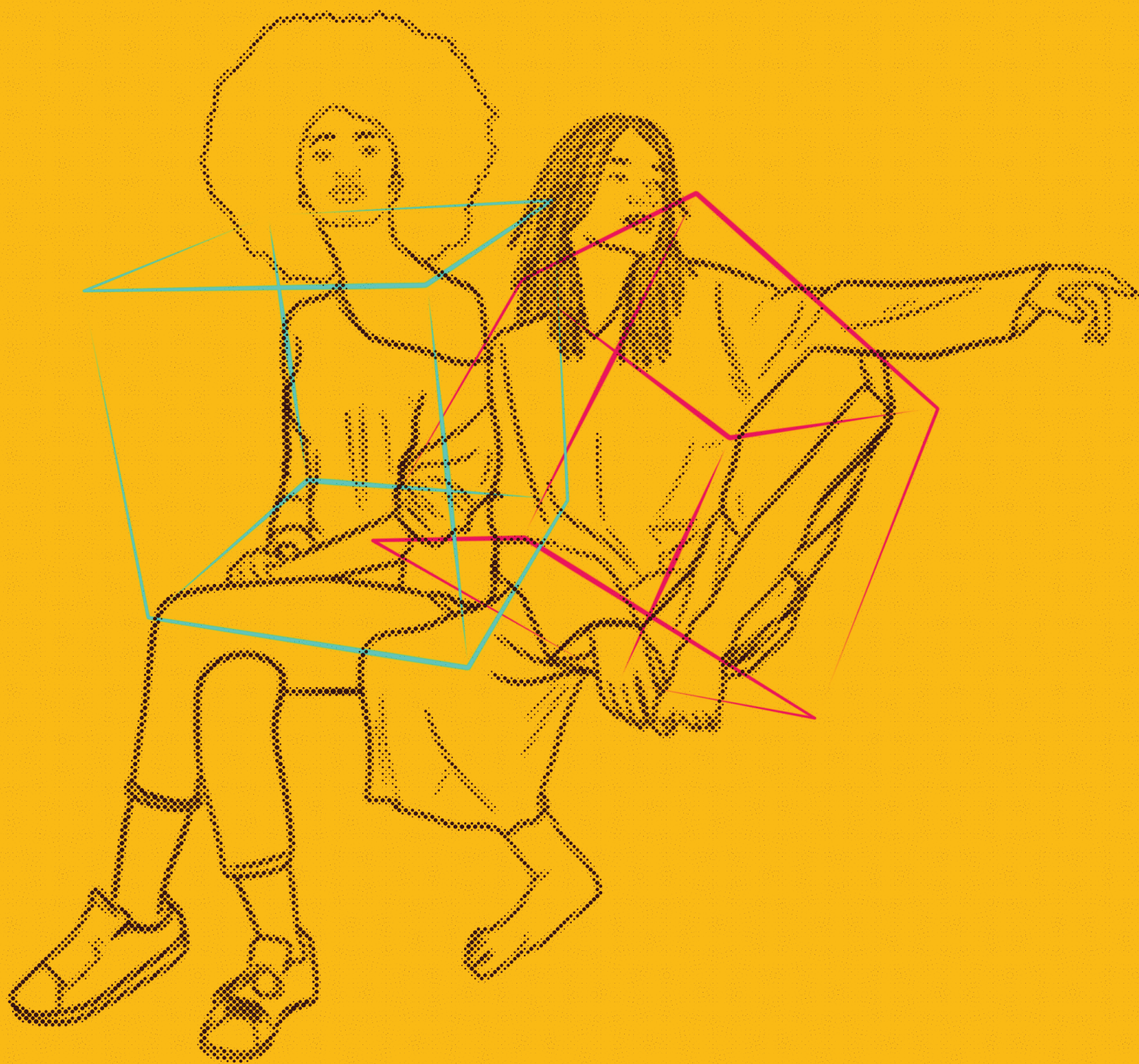
Data Stewardship, Collaboratives, Trusts and More

Favour Borokini
Bonaventure Saturday

POLLICY

Contents

Introduction

The significance of data to today's economy has stimulated ongoing conversations about the regulation and management of data. This is evidenced in the massive amounts of value constantly derived and extracted from it and the severe impacts of its misuse in today's economy. The inadequacy of traditional, mainstream regulation methods, when the wide-ranging, limitless global impacts of data on equality, worldwide prosperity and justice are considered, has given room for the reimagining of these methods. It has also led to the development of new proposed regulatory approaches.

These approaches or methods, sometimes referred to as data stewardship, involve the using and sharing of such data through avenues like data collaboratives and data trusts. Data stewardship refers to the ways through which, at every stage of its generation, creation and use, data is to be carefully managed. Peng (2018) noted that the concept of data stewardship has roots deep in the science and practice of data collection, sharing, and analysis. Thus, data stewardship incorporates all activities which preserve and improve the information content, accessibility, and usability of metadata and data. Notably, a fiduciary (or trust) level of responsibility is conveyed through the concept of a data steward.

This paper investigates broadly data stewardship methods, particularly data collaboratives and data trusts efforts and aims to identify the existing challenges, gaps, opportunities and potential recommendations on how to drive this discourse forward.

01

Where is the Current Discourse on Data Stewardship?

Globally

Data plays an increasingly important role across all sectors, industries and organisations. However, the impact of the abuse and misuse of data have come to light in recent years, sparking debate on alternate methods of safeguarding, regulation, monitoring and benefiting from one's data.
Current approaches to data stewardship guide decisions regarding who has access to data, under what conditions and to whose benefit and prioritising different ownership, partnership and use models.

Data collaboratives are one such method. Bhardwaj et al. (2014) define data collaboratives as a new form of collaboration, whereby participants in particular institutions from different sectors exchange their data for value creation to the public. Data collaboratives prioritise public and private partnerships and the exchange of data between these organisations whenever necessary. A prominent example of the use of a data collaborative was during the 2015 Nepalese earthquake. Flowminder, a Swedish data analytics nonprofit, and NCell, a mobile operator in Nepal, formed a data collaborative that entailed using telecoms data collected by NCell to track the movement of displaced people to assist humanitarian efforts (Winowatan, et. al., 2020).

On the other hand, much like legal trusts, data trusts are "legal structures that provide independent stewardship of data" (Open Data Institute, n.d.). Ruhaak (2020) defines a data trust as an independent arrangement with a fiduciary responsibility for controlling and holding data for prospective data users or beneficiaries. Unlike with data collaboratives, where partners more or less retain control over the data under their control or possession, under data trusts, data belonging to a group of people or organisations are pooled together and thereby attain a separate identity. This distinct organisation then is tasked with protecting and managing the data (rights) under its control under the guidance of trustees. The first institutional proponent of data trusts has been said to be the United Kingdom in 2017 through its independent commissioned report, Growing the Artificial Intelligence Industry in the UK (Artyushina, 2021). However, the popularity of data trusts owes to an extent to Sidewalk Labs' proposal to govern urban data in Toronto with "civic data trusts" in 2018 (McDonald, 2019).

At the moment, data trusts exist in various forms with varying structures, features and liabilities. These variations are also attributable to the goals and objectives of the proponents and trustees.

02

For instance, In March 2021, the Mozilla Data Futures Lab was launched in partnership with Luminate and Siegel, as a way of giving support to the ever-growing field of people working on data cooperatives, data trusts, and other efforts to build trust and shift the dynamics power with regards to data. This will actively steward and create collective value from data over time (Surman, 2021). Earlier on, Sidewalk Labs, a sister company to Google, had also launched a civic data trust approach that exists as an independent, not-for-profit data trust. However, there is less information on Sidewalk's civic data trust approach and it was also criticized for a "failure to incorporate feedback from community organizations and residents, and for not including fiduciary obligations for the proposed trustee organization" (McFarland, 2019), among others. The Open Data Institute (2019) notes that the Open Data Institute's approach to data trusts, in addition to civic data trusts, focuses on the initiation and use of data trusts by one or more organizational data holders and this may or may not include personal data depending on the purposes and incentives. Therefore, there is a need for a bottom-up data trust approach pioneered by Silvie Delacroix and Neil Lawrence (Delacroix and Lawrence, 2018) under which data subjects would be empowered to pool their data into a trust that would champion a social or economic benefit of their choosing. AI (2021) agrees that it is important to allow the participation of community

organizations and residents in a way that their perspectives form part of the data governance design and are responded to. This builds the beneficiaries' capacity and confidence to continue participating in the data governance processes of the data trust.

Globally, approaches to data stewardship are dictated by the existing data protection laws, in addition to the constitution, consumer rights laws and the framing of the law of trusts.
Organisations, governments, research institutions, feminist and women's rights organisations (WROs), and other initiatives can use such data to tackle major societal and economic problems by collecting, managing, sharing, and using data relevant to gender programs.

03

Africa

A data revolution is slowly underway in Africa, with large volumes of data produced daily by many research collaborations such as Africa Open Data Collaboratives. Although data is emerging as a powerful force in the digital economy, the digital footprint of Africans, and African women, in particular, remains minimal.

African research institutions that produce data, such as universities, are scrambling to avail necessary data services to researchers. Managing and sharing huge volumes of data sourced in Africa has created several operational or technical considerations and requirements. Institutions like National Statistics Offices (NSO) have collected and published data across sub-Saharan Africa however, these datasets are plagued with completeness, disaggregation, and timeliness gaps. These inadequacies have prompted most stakeholders to manage their data or turn to other commercial sources. As such, questions have

arisen on data management and methodologies and gaps in data sharing and accessibility (Juma, 2021).

It is therefore imperative for Africa to rethink models of data collaboration and access. There is an urgent need to develop its data stewardship and trust capacity to ensure that data is more open, accessible, and usable. This will address the numerous bottlenecks organisations, and movements face across the African continent by providing sustainable data access and use. This data can inform and track tangible changes in policy responses and mitigation measures to address developmental challenges in Africa, particularly with a focus on gender data, which remains a significant challenge on the continent. Therefore, as African governments race to respond to *digitalisation* and *datafication*, great effort and critical discussions should be made regarding innovative data classification methods.

04

Case Study

Discourse on data stewardship and collaboratives are still lacking on the continent. The South African government's Draft National Data and Cloud Policy (2021) made mention of the need to ensure the free flow of data for economic growth, referring to data silos being held by corporate interests, which the government often funds.

Some Use Cases in Africa

1. *Illegal Wildlife Poaching Monitoring*
2. *Migration and planning studies*
3. *Health research*

These coincide with the increase in open data sharing frameworks being advocated for, especially by foreign organisations. However, reservations surround the general conversation. A history of exploitation and extraction stemming from colonial and neocolonial activities in Africa has fostered African researchers' suspicion towards open data sharing. According to Abebe et al. (2021), these reservations are influenced by power asymmetries, trust and context and local knowledge.

Challenges and Gaps

Practical Challenges

Data trusts remain largely theoretical concepts that are yet to be fully deployed. This means that their legitimacy and structure are not fully understood, explainable or supported by law. There is limited guidance as to how they ought to be set up and what the duties and liabilities of parties to the data trust are not fully spelt out.

Ethical Challenges

While there have been some impressive results from efforts to further the data stewardship discussion to date, Hoffman (2019) noted that there are still few global success stories where large scale commercially controlled data is used for the common good. The world is struggling in reducing data abuses of all kinds, enhancing data accountability and improving ethical standards while also ensuring that data value can be beneficial to both the public and private stakeholders. There are various gaps in how data is being shared, sold, and misused, with little ethical awareness and involvement of the affected persons by the organisations in charge.

There is also the underlying concern that rather than utilising data for public good, data stewardship might succeed in transferring even more data into the hands of private sector interests.

Lack of Trust

Hidalgo-Sanchis and Stefaan (2019) pointed out that uncertainty has been created by the lack of trust among individuals and institutions, especially women rights and feminist organisations. These organisations report a lack of meaningful and accessible data and a lack of internal financing and capacity to improve gender data publication, stalling incentives for innovation in data stewardship

Lack of Transparency and Consumer Rights

Lack of control and transparency regarding the use of personal data is yet another challenge since persons feel that they cannot control data about their own lives.
This risks abuses of privacy, leading to a decrease in trust. Trust is vital in determining the possibility of users sharing their personal data. This conforms with the findings of Consumers International and Internet Society Joint Report (2019), which revealed that 75% of the people distrust the way data is shared with other organisations without their permission. This means that individuals, organisations and movements find it challenging to derive value from their data and effectively use data to solve their pressing problems. Some institutions may also have trust issues, especially how data is collected and managed, thereby declining to share or use data, leading to data hoarding.

Weak Regulatory Frameworks and Political Manipulation

There are trust and transparency issues on the effectiveness of the regulatory framework in securing the infrastructure used for exchanging and storing data against cybercrime intrusion. There is also a potential for governments' abuse of citizens' data for political gains through politically motivated discrimination or surveillance.
In addition, not all countries prioritise the transparent documentation of data, thus undermining public trust and leading to a deficit of public value.

Lack of Technical Skills

Many bodies, especially in Africa, such as feminist movements and women rights organisations, often lack the capacity in data analytics to process, analyse, establish controls, etc., limiting the positive impact of data to advance their work. This is in addition to the high burdensome transaction costs incurred primarily by smaller entities at numerous points in the data lifecycle.

Opportunities

Some of the challenges faced by data collaboratives and data trusts have been highlighted above. However, this remains a blossoming field. For African countries to succeed in building data infrastructure to achieve economic growth, the above concerns must be taken seriously.

Opportunities exist in reshaping the existing models on how data is collected, managed, shared and used to enhance evidence-informed development and growth programs. Data collaboratives and stewardship can improve the way data is collected, maintained and transmitted to create value to all the stakeholders, including organisations, governments, research institutions etc., to enhance service delivery and crucial to guaranteeing success is the facilitation and coordination of data collaboratives or trusts.

It is also of great importance to dismantle structural inequalities by exploring and prioritising various segments of data stewardship through a feminist lens. This includes issues to do with data sovereignty, data localisation and data systems and safeguards to manage the vulnerabilities associated with the use of data. Moreover, strengthening gender data production, quality, availability, and use also builds up the overall data ecosystem.

There is also a need to pass supportive legal and ethical frameworks and other mechanisms which empower persons to exercise their data rights, share data, prevent digital harms as well as data

collaboratives that are independent, non-partisan and intersectional, thereby filling the existing gaps in data governance and accountability to the citizens. This will lead to the collection of data that is more credible for intersectional programming and use creating data infrastructures and systems that are more resilient, thereby improving the status, rights, and opportunities for all stakeholders involved at global, regional, country and local levels.

There are numerous opportunities for growth evidenced by the growth of digital companies, data architects, technology startups, university programmes, emerging feminist organizations and an overall presence of technology in businesses. The importance of complying with regulatory frameworks and improving the existing data access policies and structures should be given attention. This is because they are crucial activities for seizing the opportunities in data governance. While the unfounded notion that regulation stifles innovation is prominent, it must be noted that far from this, regulation exists to guide and support safe innovation practices.

The discussion surrounding data stewardship, therefore, gives room for more discussion and interdependence between public, private and third sector organisations. By lending and converting our energies towards these conversations, it is possible to develop alternative ways through which we can better imagine data futures beneficial to all.

08

Next Steps
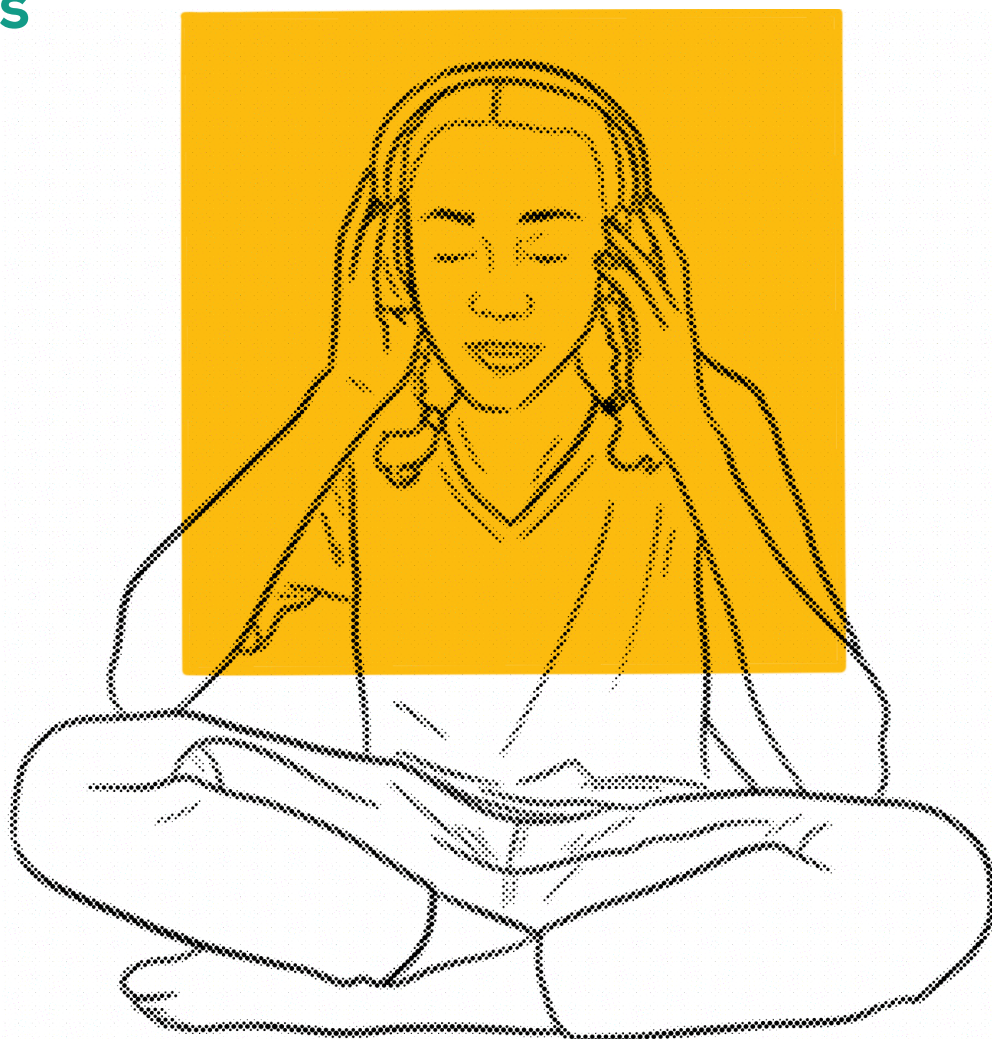
Pollicy (2021) has researched the Afro-feminist ways and methods of data sharing and data collection. In our research, we prioritise African and feminist ways of knowing, which centre women's experiences and place the data futures of African women in their own hands. In particular, we view an Afrofeminist data future as one where African women have the right to privacy and complete control over personal data and information online at all levels - a form of data justice (Pollicy, 2021).

Data stewardship systems such as Data collaboratives and Data trusts are a glimpse into what this future might eventually turn out to be. Through our roadmap for strengthening the feminist data ecosystem, Pollicy has begun the process of driving this discourse forward. We invite all interested parties to join in this discussion on collectively imagining the future of data stewardship, collaboratives and trust within Africa and on the global stage.

09

References

Abebe, R., Aruleba, K., Birhane, A., Kingsley, S., Obaido, G., Remy, S. L. and Sadagopan, S. (2021). Narratives and Counternarratives on Data Sharing in Africa. In Conference on Fairness, Accountability, and Transparency (FAccT '21). https://doi.org/10.1145/3442188.3445897

Affordability Report, 2017, Accessed on Oct 26, 2021,
from http://a4ai.org/affordability-report/report/2017/

AI, J. (2021). Participatory data stewardship. Reading time.
https://www.adalovelaceinstitute.org/report/participatory-data-stewardship/

Artyushina, A. The future of data trusts and the global race to dominate AI. (2021).
https://www.bennettinstitute.cam.ac.uk/blog/data-trusts1/

Bhardwaj, A., Bhattacherjee, S., Chavan, A., Deshpande, A., Elmore, A. J., Madden, S., and Parameswaran, A. G. (2014). Datahub: Collaborative data science and dataset version management at scale. arXiv preprint arXiv:1409.0798.

Consumers International and Internet Society Joint Report (2019) accessed on Oct 26, 2021, from https://www.internetsociety.org/wp-content/uploads/2019/05/CI_IS_Joint_Report-EN.pdf

Delacroix, S. and Lawrence, N. (2018). Bottom-Up Data Trusts: Disturbing the 'One Size Fits All' Approach to Data Governance. International Data Privacy Law. https://dx.doi.org/10.2139/ssrn.3265315

Department of Communications and Digital Technologies. (2021). Draft National Policy On Data and Cloud. The Republic of South Africa.
https://www.gov.za/sites/default/files/gcis_document/202104/44389gon206.pdf

Hidalgo-Sanchis, P. and Stefaan G. V. (2019) Opinion: The promises — and challenges — of data collaboratives for the SDGs Accessed on Oct 26, 2021, from
https://www.devex.com/news/opinion-the-promises-and-challenges-of-data-collaboratives-for-the-sdgs-94082

Hoffman, W. (2019). Collaborating for the common good: Navigating public-private data partnerships.

Iyer, N., Chair, C., and Achieng, G. (2021). Afrofeminist Feminist Data Futures. Pollicy. https://archive.pollicy.org/wp-content/uploads/2021/03/Afrofeminist-Data-Futures-Report-ENGLISH.pdf

Juma, L. (2021). Building Sustainable Gender Data Systems accessed on Oct 26, 2021, from https://www.developlocal.org/building-sustainable-gender-data-systems/

Mcdonald, S. (2019). Reclaiming Data Trusts. Centre for International Governance Innovation. https://www.cigionline.org/articles/reclaiming-data-trusts/

McFarland, M. (2019). Alphabet's plans to track people in its 'smart city' ring alarm bells. https://edition.cnn.com/2019/07/09/tech/toronto-sidewalk-labs-google-data-trust/index.html

Open Data Institute. (2019). Data trusts: lessons from three pilots. https://docs.google.com/document/d/118RqyUAWP3WIyyCO4iLUT3oOobnYJGibEhspr2v87jg/edit#heading=h.3fngvdcfo2cs

Peng, G., (2018). The State of Assessing Data Stewardship Maturity – An Overview. Data Science Journal, 17, p.7. DOI: Accessed on Oct 30, 2021, from, http://doi.org/10.5334/dsj-2018-007

Ruhaak. A., (2020). Data trusts: what are they and how do they work? Accessed on Oct 26, 2021, from, https://www.thersa.org/blog/2020/06/data-trusts-protection

Winowatan, M., Zahuranec, A. J., Young, A., and Verhulst S. (2020). Data Collaborative Case Study: Leveraging Telecom Data to Aid Humanitarian Efforts. Medium. https://medium.com/data-stewards-network/data-collaborative-case-study-leveraging-telecom-data-to-aid-humanitarian-efforts-9b5c2fc893a4