

**Advancing Data Justice Research and Practice  
CIPESA Report to The Alan Turing Institute  
March 2022**

## **1. Background**

As part of the Advancing Data Justice Research and Practice project, the Collaboration on International ICT Policy for East and Southern Africa (CIPESA) conducted stakeholder engagement exercises to elicit feedback and viewpoints on data justice in Africa. The stakeholder engagement was part of a mixed methods research approach combining desk review with key informant interviews and a brainstorming workshop.

In all, we engaged 25 respondents from across 11 countries - Botswana, Egypt, Ethiopia, Ghana, Kenya, Nigeria, Uganda, South Africa, Tanzania, Uganda and Zimbabwe. Sector representation included academia, the legal fraternity, data intermediaries, data protection authorities (DPA), Disability Rights Organisations (DPO), think-tanks, communications regulatory authorities, developers, development actors, digital rights and inclusion activists. See annex 1 for the respondents' affiliations and countries. The desk review involved literature review including relevant reports, news articles and commentaries on data governance and data justice in Africa.

The findings of the stakeholder engagement and desk review are outlined in section 2 below, from which conclusions and recommendations are drawn under section 3.

## **2. Findings**

### ***2.1 Overview of Data Governance in Africa***

The exponential growth in data generation by governments, corporations and civic organisations, as well as improved skills in storing, processing, visualising and sharing this data, hold much promise for enabling inclusive and sustainable development. However, while the “open data” movement to improve the availability, openness, use and reuse of data to power development and improved governance has been active for about a decade, results are mixed, often undocumented, sometimes negative. Today, Big Data and Artificial Intelligence (AI) products and services and non-traditional sources of data are optimised for extraction and attention, which has further increased concerns on data security and safety, and on human rights, which could undermine the efficacy of data initiatives.

In Africa, as elsewhere, challenges remain to harnessing data to inform policy, development and governance practice. Major studies have cited common challenges applicable to developing countries and to Africa, including [data protection and privacy concerns](#), [poor quality, completeness and reliability of data](#), [limited access to and usability of data](#). Others include [data illiteracy](#), limited capacity of civil society actors to [re-use data and to publish own data](#), limited qualitative evidence on the [performance of data-based initiatives](#), [limited focus on the “development sector”](#), and [deficient](#) interdisciplinary and cross-continental research collaborations. Conversely, African governments need multi-stakeholder input from non-traditional data entities such as banking, telecommunications and fintech, as well as development agencies, to help them understand what legislation, policies and oversight and enforcement mechanisms are necessary to strengthen the protection of human rights in the rapidly changing Big Data age.

Meanwhile, digital technologies have become pivotal to the enjoyment of rights and improvement of livelihoods in Africa. However, many governments in the region are taking [steps that weaken](#) the potential of digital technologies to catalyse inclusion, free expression and civic participation or to drive innovation.

Digital rights issues of concern include increased surveillance, personal data breaches, internet disruptions, [insufficiently resourced data protection authorities](#) and a [proliferation](#) of laws and regulations that undermine technology's potential to drive socio-economic and political development.

A [2017 study](#) noted that while more sources of data can make the development sector more agile, this may open opportunities for surveillance and threats to individual privacy. Further, the study noted that the protection of personal privacy may be used as an excuse to withhold public sector data that could be made open for citizens to advocate for better public services, hold governments accountable and tackle public sector corruption. Still, as privacy regulations continue to strengthen safeguards against data misuse, there is [a need](#) for widespread public support for regulatory reform to address the risks of 'missed use' of this technology to benefit the whole of society.

Many African governments are undertaking rapid data collection and digitization initiatives – e-government services, digital identity (digital ID), biometric voters' cards, drivers' licences and SIM card registration. These exacerbate the gamut of challenges to data governance as noted above. Where digital rights are under threat, the appetite for public participation has been dulled for many citizens because there is a high perception of surveillance, and mistrust of government-initiated (digital) programs. Without addressing the digital rights issues inherent in data initiatives, they are bound to meet limited success. A key area that is not well articulated in mainstream discourse about data rights is data justice, yet it is central to ensuring that the digital society and the data revolution deliver dividends to all citizens, most notably those that are most vulnerable.

## **2.2 Regional Instruments and Data Justice**

The African Union (AU) Data Policy Framework that is under development, [seeks](#) to “provide a shared vision of a data realm that is achieved in a just and fair manner, whilst creating the safe and trusted digital environment necessary for the development of a sustainable and inclusive African digital economy and society.” The [African Union Convention on Cyber Security and Personal Data Protection](#) also has some provisions that may be utilised to advance data justice. For instance, section III provides for obligations relating to conditions governing personal data processing; article 13 has the basic principles governing the processing of personal data; and article 14 has specific principles for processing sensitive data.

The [Digital Transformation Strategy for Africa \(2020-2030\)](#) provides that special attention should be given to women, people living in remote areas, persons with disabilities, disadvantaged and marginalised communities through the establishment of a platform for dialogue and social cohesion that involves these target groups. The strategy adds that the promotion of online cultural diversity must be supported to ensure that every person participates fully in society. Further, the strategy urges states to “design policies based on a human-centred and holistic approach that takes into account the local context and cross-cutting issues relevant to all stages of policy design and implementation.”

However, although many countries have legal frameworks concerning personal data protection, a [UNESCO study in 2021](#) suggested that these legal provisions may need to be updated to the new uses and applications of data engendered by AI to offset biases and discrimination, including on the basis of race and gender, or loss of personal privacy through predictive analysis, among others. According to the UNESCO report, beyond data governance and personal data protection, there is also a need for legal protection against algorithmic bias and discrimination, an area in which most countries have not developed any measures.

### **2.3 Understanding of Data Justice**

There were mixed levels of understanding of the meaning and import of data justice. This was reflected in respondents' choice of words representing the essence of data justice. They included fairness; inclusion; accountability; transparency; data classification; data as a resource; discrimination and bias; data equity; presumption of innocence; addressing structural inequalities; data feminism; data collection, analysis and use in society. Others were empowering data subjects to become data owners; social justice; availability and accessibility of data; opportunity; data privacy and personal data protection (or lack thereof); ethical collection and use of personal data; data privacy ecosystem; and representation.

One respondent described data justice as "fairness in the way people are made visible, represented and treated" in digital data. Another stated: "Data justice to me would be equated to the concept of Human Rights Based Approach to data, whose principles include participation, data disaggregation, self-identification, transparency, privacy and accountability." Some of those who were uncertain of the term data justice considered it a "Western concept".

The key intersections of data justice and social justice outlined included discrimination and absence of redress mechanisms especially in cases of automated decision making. The resultant data misrepresentation and under-representation was particularly exacerbated amongst marginalised communities such as women, ethnic minorities, persons with disabilities and migrants who faced challenges in equal and equitable access to public services and participation in society and governance. Through the facilitation of surveillance and interception of communications, data-driven technologies were also said to have curtailed the work of activists, critics and opposition groups in pushing for good governance, transparency and accountability.

#### ***Testimony of intersections of data driven technologies and social justice.***

*I am a black African man working in the United Kingdom and I believe the issue of race impacts on my life and the wellbeing of people like myself. Discrimination along racial lines is a reality, as is the fact that I am a migrant. These two factors are able to affect the number of opportunities I have when it comes to issues around the type of job I can do, the amount of interest I pay when I apply for loans or vehicle and home insurance, as well as the quality of the service I get when it comes to public healthcare and the way I am handled when I interact with law enforcement officials. I am more likely to be treated more harshly or suspected to be a criminal compared to how white people would be treated. When it comes to calculating my insurance premiums, or access to bank loans, my postcode may be used to discriminate against me because I likely live in a lower-to-middle class part of the city that is characterised by immigrants and black or non-white people. Police facial recognition technology may mistake me for another suspect of colour because that technology was not properly calibrated to accurately differentiate between non-white faces. My enjoyment of voice command-based technologies is hampered by my non-European accent, and non-English name which services like Alexa and Siri are not able to capture or process. Overcoming these challenges requires an admission that these biases do exist and that they are particularly targeted against non-white communities.*

At continent-wide level, data-driven technologies were exacerbating social injustice through data colonialism. As one respondent noted, "Powerful, data-rich countries and corporations in the global north undercutting capacity development in the global south is at risk of becoming a reality in most of Africa." He added that most of the data on the continent was not collected or owned locally and was stored/hosted abroad. "Tech giants, using sometimes biased and prejudiced algorithms, have monopoly over most of Africa's digital data," he said. These sentiments were echoed by another respondent who

stated that, “when you look at the people who develop these technologies, when you look at where they were developed, who the primary users are, who the target market is, you realise the technologies are developed usually in Northern America, Europe, and maybe Asia then exported to Africa in most instances.” As a result, the data within these technologies entrenched the “biased views and perceptions [that] Europeans, Americans, and Asians have of Africa and its social struggles.” The respondent added that the social injustice entrenchment was worsened by the fact that platforms “were not bothered” about representation of demographics outside of their home regions.

Data colonialism was said to present challenges in developing and implementing regulatory and governance frameworks “where generators and users of data are not actually regulated entities within a specific jurisdiction”. This was in reference to the fact that companies such as Meta, Google and Twitter, which collect, store and process a lot of data about users of digital technologies in Africa, are not locally licensed or regulated. Further, digital tools such as search engines were said to have the potential to feed biases and prejudices “by feeding powerful digital infrastructure with biased back-end inputs that produce biased search results that can lead to real world harm.”

Similar to the harmful automated decision-making at individual level as highlighted above, the data rich-countries and corporations in the global north purportedly influenced development financing, resource allocation and intervention in crisis as demonstrated during the COVID-19 pandemic. Many (mostly western-based) epidemiologists anticipated that the pandemic would kill millions of Africans. Researchers at the Imperial College London [put the number of estimated deaths at three million](#) as the worst case scenario if nothing was done. As the pandemic ravaged the western world, however, many of these scientists and analysts were [mystified](#) by the comparably fewer deaths in Africa. While many credited Africa’s young population, less travel infrastructure, and a stroke of luck for this success, John Nkengasong, Director of the Africa Center for Disease Control and Prevention, [attributed](#) this early success to African countries taking radical preventative steps very early on.

#### ***Digital Technologies: Democratising or Deepening Social Injustice?***

*Digital technologies have taken centre-stage in African citizens’ lives. However, these digital technologies can either amplify the data justice concerns, or they may deliver digital dividends that come with increased digitalisation and the use of technology. Unmistakably, the effects of technology on different communities are not the same and it is possible that the groups that are historically most marginalised in society stand to experience deeper data justice harms than the rest of society.*

*A number of positive areas were noted in which technology can promote social justice. It empowers individuals with means to speak out and participate in community affairs with relative ease and mobilise regardless of location. Inclusion in a wider community is also a key positive attribute of social justice. Financial inclusion and the ease of conducting financial transactions is yet another notable way in which digital technologies positively affect social justice. On the other hand, these benefits are not universally enjoyed. Those who lack access to technology are left out, as are those with low usage skills and those that lack [meaningful connectivity](#) (i.e. “the ability to use the internet every day using an appropriate device with enough data and a fast connection”). Moreover, specific groups such as persons with disabilities, such as those with hearing and sight impairment, are often unable to enjoy the benefits of technology due to inaccessible formats or incompatibility with assistive devices.*

The potential ways of technology overcoming social injustice were pinned to overcoming biases and discriminatory policy and practice within system design and data collection. Further, through

representation in large data sets that are disaggregated by impacted communities “which will show the extent of marginalisation, hence making it the basis of effective interventions.” This, in addition to the development of context-specific algorithms that account for unique differences in geography, ethnicity and political thoughts.

*“Technology can not overcome social justice issues if those affected by social injustice are not leading any of the processes. The technologies exacerbate social injustice issues because they are designed and built by systems and structures that already oppress.”*

To overcome data colonialism, respondents put forward the idea of “democratisation of the data value chain” by practising equality at every stage - collection, analysis, use and dissemination; in hiring data collection staff and algorithm designers as well as independent fact-checkers.

Another potential way for data-driven technologies to overcome social injustice was through the development of applications that track trends and aid advocacy by organisations working to advance the rights of marginalised and minority communities. However, caution was raised about the use of digital platforms for social justice advocacy. The perception of causing change via technology was said to pose a “danger to the real hard work of promoting social justice which is more physical and grounded.” This perception was said to have also contributed to paranoia among governments on the continent, which has led to a crackdown on digital rights through [network disruptions](#), [increased surveillance](#), [digital taxation](#), and [electoral malpractice](#).

#### **2.4 Concerns around data collection and use**

Whereas some of the respondents acknowledged the existence of data protection and privacy laws and policies, as well as enforcement authorities in some African countries, these were deemed insufficient as data processors were still able to employ discriminatory algorithms for processing. Respondents pointed to COVID-19 as the most live example, with numerous governments on the continent employing disease surveillance measures to curb the spread of the virus, with little regard to privacy and data protection. Indeed, it has been found that in efforts to learn more about how the virus was spreading, many countries on the continent were sharing medical records, tracking people’s movements and tracing contacts, which had created “[new norms](#)” for data governance, whilst [hurting digital rights](#) and [deepening democratic regression](#).

Meanwhile, weak oversight and safeguard mechanisms in data collection and processing often facilitated unlawful surveillance and interception of communications. Here, a respondent cited Uganda’s Data Protection Act, 2019, which “came too late to be effective” given that many pre-existing laws which remain on the statute books, such as the Computer Misuse Act, 2011, the Official Secrets Act, 1964, and the Anti Terrorism Act, 2010, all of which undermine privacy and data protection and grant state authorities sweeping powers over data.

Other concerns highlighted included framing of questions during data collection, language, culture and multiple collection centres. There were also concerns around limited characteristics/ attributes in data collection. For instance, the exclusion of disability characteristics which ultimately affects their inclusion in vital statistics, resulting in “scanty data on many aspects of disability.” This lack of disaggregated data, as demonstrated by CIPESA and the World Benchmarking Alliance, had exacerbated [digital exclusion](#) and [failure to meet digital accessibility obligations](#).

Within the governance context, ‘data politics’ was also raised as a concern, with “some data conveniently never collected or made public”. An example cited here was the absence of data on torture and human rights abuses against opposition groups by the state as part of the [Annual Crime Report by Police in Uganda](#), despite widespread reports of the [prevalence](#) of torture in the country. Similar ‘data politics’ tactics were witnessed in countries such as [Burundi and Tanzania](#) at the outset of the COVID-19 pandemic.

### **CSOs working to advance data justice on the continent**

Notable efforts to positively influence data collection and use include by [Data4Change](#) and [Development Initiatives](#). Data4Change works to strengthen data capacity of communities, empowering them to collect their own data using robust and ethical methods, and to use [data insights to influence policy makers](#), create evidence-based programming and tell their own stories. Data4Change’s work focuses on gender-based violence, the rights of sexual minorities, disability rights, racial injustice, poverty, asylum and citizenship rights.

As part of their programming activities, Data4Change supports affected communities to engage in conversations with those in power. However, they point to difficulty in accessing the “right people in power” to question them - with bureaucracy in systems and processes being the main obstruction. Data4Change also facilitates conversations about the relevance of principles such as data equity, data justice and data feminism amongst affected communities. The organisation notes that terminology in the field has been created within systems, structures and institutions “that already have a lot of power” and it is important to understand how affected communities describe what is happening to them.

In Kenya, Data4change, in partnership with the Social Justice Centre, initiated the [Life Under Curfew](#) project which through a survey “repatriated” data that had been extracted from local communities using physical murals installed within the community. The project also collected additional data using accessible and collaborative activities. All the data is on display for anyone to see, access and engage with. Although small scale, the Life Under Curfew example is relevant to access within data justice. According to the Data4Change respondent, “whereas the project doesn’t disrupt power structures, it hopefully works towards it.”

The limitations to Data4Change’s work include funding. They also note that the recent trend to compensate individuals with phone credit or lunch for providing data through responding to surveys “creates a very complicated power dynamic.”

For its part, Development Initiatives has been active in advocacy for fairer national data ecosystems. In Uganda, it has registered some success in promoting adoption of [attributes of good data and use of technology](#) in monitoring data production activities by the official data producing agency - [Uganda Bureau of Statistics \(UBOS\)](#). The entity has also pushed for the enactment of open data legislation, with a [draft law](#) in the pipeline. Further, through engagement at hackathons and other events such as the [Kampala DataFest](#) and the [World Data Forum](#), as well as [commemoration of OpenData Day](#) and [Statistics Week](#), Development Initiatives continues to raise awareness about best practice in the data value chain towards adoption of a stronger and fairer data ecosystem in Africa and beyond. Development Initiatives also works to promote women in science, technology, engineering and mathematics (STEM) through a girls-only data science and software development internship programme.

## **2.5 Feedback on the six pillars of data justice**

### 2.5.1 Power

Across the stakeholder categories engaged, there was consensus that mapping out power relations and imbalances as shown in the table below was necessary. “One can’t achieve data justice without interrogating these and more dynamics,” said a respondent.

#### **Power dynamics within the data life cycle/value chain**

1. Production stage: Investment Power or Power of the Purse
  - a. Who allocates the money for data collection?
  - b. Which data to collect and what level of disaggregation?
2. Collection state: Analytical Power or Power of the Collection Tools
  - a. Which indicators, facets and population dynamics will the data reflect
3. Dissemination power
  - a. How far certain data will reach

Beyond the geopolitical, infrastructural, cultural and political sources of power in data innovation ecosystems, religious, gender, literacy, accessibility, regulatory and economic power sources were the additional levels suggested.

*“A lot of vulnerable populations such as women and girls, persons with disabilities, sexual minorities are still largely invisible in data in Uganda because they have no [power] advocates at the start of the data life cycle.”*

Factors affecting power awareness among impacted communities included illiteracy and limited knowledge stemming from data collectors not declaring the intended purpose of data collection and seeking consent. “There isn’t a lot of knowledge beyond the consent forms that we are made to fill in to access services,” noted a respondent. Using an example of loan applications, the respondent added that outcomes/decisions of data processing are not discussed and there is often no explanation or disclosure of third party access or targeted advertising.

Due to this lack of access to adequate information on data rights, impacted communities were said to be unable to “link the unjust exercises of power causing harm or marginalisation to the powers that have control or influence over their data” to, for instance, discriminatory pricing based on age, or restrictions on access to services based on location. Whereas there may be interest amongst impacted communities to mobilise against unequal power dynamics, this has mostly been in the offline context and less so in the digital sphere.

Where there has been some history of mobilisation against unequal power dynamics and social injustices offline, strategies have included human rights advocacy, participation in policy making processes, strategic litigation, community awareness raising, and engagement with stakeholders on obligations and compliance with international human rights instruments. Some respondents indicated that offline movements have in some cases been reflected in online/digital contexts. However, for some communities, such as persons with disabilities, the actual means to mobilise online were limited “if not non-existent” due to extreme marginalisation offline. Empowerment for such communities would entail improved access and use of data-driven technologies “in ways that enhance their control over personal data.” It would also entail “active participation” in data collection processes by, for instance, “contributing relevant questions pertaining to disability in surveys and management information systems.”

For developers, civil society organisations and data intermediaries, the indication was that of a high level of understanding of the power dynamics surrounding data processes. This awareness informed activities and programming aimed at addressing the imbalance. “I am personally well informed due to a long career as a researcher of the political economy of data. However, many stakeholders still fail to see the role of politics in data and vice versa. Data, like all things, is not bias-free.” For data and social justice advocates to overcome power imbalances and adverse impacts on communities, “creative engagement” with policy makers as opposed to confrontational approaches were recommended. “Study the pros and cons or existing legal and policy frameworks, weigh each and present alternatives respectively. Many policy makers also want to change hence they are likely to be responsive,” said a respondent. This respondent cited the need for identifying champions and strengthening of bargaining voices through data collaboratives or associations such as the [Uganda Data Revolution Initiative](#) and [Civil Society Budget Advocacy Group \(CSBAG\)](#) which advocate for more resource investment in fairer data ecosystems in the country.

An area of concern is the capacity of lawmakers, the executive and the judiciary to formulate, implement and enforce policies and laws concerning [AI](#) and data justice. However, there were mixed findings with regards to awareness of the power dynamics surrounding data collection and use among respondents from the policy maker and regulatory authority community, with some deeming it to be inadequate, while others indicated full awareness. Some policy makers indicated that the data justice field was emerging, with the private sector “in the driving seat” and regulators as well as policy makers had inadequate resources (financial, human, skills and exposure) to question power and raise critical awareness. Those who indicated awareness of power dynamics in data alluded to compliance rules and requirements which govern their day-to-day operations.

### 2.5.2 Equity

Awareness of the concept of equity among policy makers would be useful so as to put policy makers squarely at the intersection of data justice and socio-economic-political development. However, given the challenges around resources among many policy makers, reflection on the decision-making process in data ecosystems was deemed unfeasible/impractical. Within this stakeholder group, the idea of measurement justice was welcome, as focusing on weaknesses rather than the strengths of marginalised and vulnerable communities “aggravates discrimination”.

**Exclusion in digital ID:** *The digital identity card has the potential to promote the inclusion of marginalised groups and enable them to access services or even credit. However, many of the social justice challenges in the region mean that groups that are vulnerable and marginalised also tend to have difficulties in accessing and owning digital identities. That hampers their access to services, since the digital ID has become a needed passport to access a range of public and private services. Increased digitalisation, of which the digital ID is an integral part in a number of countries, could thus exacerbate the digital divide for some communities that lack IDs, lack ample access to digital services, or who do not have the skills to use digital devices and services. Indeed, the issue of capacity and skills is not confined to digital ID and Services - it permeates other sectors. Those who lack the requisite skills have less power, less avenue for participation, inequitable access to and no power over how their data is collected, stored and processed. In 2020, global consultations on the World Bank’s Principles on Identification for Sustainable Development recognised that digital ID systems across the world were being deployed in a “manner that enables repression through enhancement of censorship, exclusion and surveillance”. As part of the consultations, civil society [called](#) for “centering transparent and democratic processes” in ID systems and processes, and*

*for United Nations and regional human rights mechanisms, including the High Commissioner on Human Rights, treaty bodies, and Special Procedures, to take up the human rights risks in digital identification systems as an urgent agenda item under their respective mandates.*

Similarly, for impacted communities, awareness of the concept of equity was deemed useful. Amongst communities such as persons with disabilities, knowledge about representation in data may be low. Nonetheless, a small number were said to “know how exclusion, discrimination and vulnerability have been used as sources of inequality in data representation.” As for the idea of measurement justice and advancing social justice within the community, “focusing on strengths rather than perceived weaknesses of persons with disabilities gives them the opportunity to be active participants in data-collection processes.” Broadly, however, algorithmic bias would potentially continue to undermine the positives and strengths of measurement justice. As an example, one respondent stated: “current applications for housing and bank finance do contain positive data points such as an above average salary, or good credit record. But such positives are still not enough to overcome racial or gender bias.” While applauding the strengths-based approach to analytics that involves communities, another respondent emphasised the need for “good balance to be retained” to ensure that the overall findings of the data remain unduly influenced. “Sometimes it is important to ‘let the data speak for itself’ regardless of the unpleasant findings,” said the respondent.

***Case study of discriminatory focus on disadvantage and negative characterisation***

*The [Uganda Demographic and Health Survey](#) which is conducted every five years by the Uganda Bureau of Statistics (UBOS) and Ministry of Health contains a chapter on “attitude toward wife beating”. Some of the questions include reasons for which beating is justified. Over the years, the survey responses to the chapter have returned alarming findings - at some point over half of respondents (including majority women) indicated at least one reason to justify domestic abuse.*

### **2.5.3 Access**

Approaches to data justice that are mindful of real world inequality would be relevant including to “empower marginalised communities and address their rights and concerns” towards improved livelihoods. However, again, limitations of resources, skills and knowledge posed a barrier to implementation of the principle on access as described in the six pillars, especially in light of wider socio-economic challenges on the continent, making access as defined by the pillar “the least of policy makers urgent concerns.” Nonetheless, the perception was that policies related to data protection and privacy audits could be integrated in the airing and sharing of unjust data processes, and form the basis for policy making in the context of the access pillar.

The approach proposed under the access pillar was said to be much needed to address bias and discrimination among impacted communities. However, it also had limitations in terms of mobilisation of stakeholders to understand the importance and use of data justice. To a certain extent, the pillar pre-supposes that the beneficiaries of biased and discriminatory practices would be able and willing to initiate the required change. This was because the beneficiaries of bias and discrimination may not of themselves act to improve the situation.

Another limitation was the “failure to take into consideration all the different facets of vulnerability and marginalisation”. As with the principles of the power and equity pillars, awareness among impacted communities of the benefits or risks associated with data processes was limited, largely due to historical marginalisation. However, amongst some communities such as persons with disabilities, there was some

awareness of existing inequalities relating to distribution of resources and access to public services albeit “without linking them to data processes”. Advancing the benefits of data work, responsible data sharing, research and innovation among communities would thus enable communities to use the principle of access as the “basis for demanding for better service delivery and better representation in decision-making centres”.

Meanwhile, advocating for transparency in data systems and for the exposure and sharing of potentially unjust data processes and outcomes would nurture cautiousness among communities about sharing personal data and loss of socio-economic opportunities. Transparency in data systems would also enable communities and advocacy groups to challenge and push for reforms against discrimination and marginalisation. According to one respondent, for effective transparency, there must be “a participatory and consultative national statistical system where a diverse set of stakeholders are part of the design of any data collection, data analysis and data use activities.”.

For developers, civil society organisations and data intermediaries, the indication was that “data is not neutral” and transparency would help with understanding biases or limitations in data collection. Further, it would help with creating replicable/scalable data systems as “proprietary systems tend to be exclusive”.

#### **2.5.4 Identity**

Respondents deemed it feasible for policy makers and developers to interrogate how impacted communities are represented in data. However, engagement in this regard would require funding, skilling and education campaigns in digital rights, minority rights and the data harms of invisibility. It would also be beneficial to have guidelines, policies and toolkits covering aspects such as the intersection of identity and social justice, big data and Artificial Intelligence.

Erasure, exclusion, grouping or non-acknowledgement of some identities in Africa is reportedly prevalent and well-known to stakeholders. The invisibility of some groups is exemplified by intersex persons. For instance, while Kenya in its [last census](#) introduced a category for intersex persons, in most countries in the region such data is not collected.

Exposure, skilling and training on the principles of equality were identified as necessary for policy makers and developers to recognise and tackle harmful categorisation and erasure which can occur during the development of data-driven technologies. Crucially, any such interventions should be informed by affected communities.

Without the knowledge or access to platforms as is the case currently and highlighted in the sections above, impacted communities are unlikely to interrogate and contest representation in data. Requisite resources and requirements for more engagement include an enabling legal and policy framework, access to information, training in data collection, use and contestation. Developers were called upon to engage in diversity studies and design systems that empower communities to do the interrogation directly. Among the stated tools for identifying erasure, exclusion, grouping or non-acknowledgment of some identities were reports by developers of data systems and social-economic development interventions.

There was no doubt that more accurate representation in data would be an appropriate way to establish equal dignity and autonomy. However, for some communities, such as persons with disabilities, there would be “difficulty in ensuring accurate representation in data processes because persons with disabilities are a heterogeneous group with various and overlapping categories.”

### **2.5.5 Participation**

The participation of impacted communities in agenda setting and decision making around data governance was deemed beneficial as “communities need to be empowered and centred in the processes”. Prevailing factors impeding participation which need to be addressed included lack of awareness, illiteracy, poverty, cultural and societal bias. Meanwhile, with bureaucracy possibly limiting community participation in data processes, agenda setting and decision making, value-addition needs to be demonstrated given that a lot of resources are being spent on data processes that do not include community members.

*“I have seen a number of occasions where iNGOs (international non-government organisations) have designed a data collection methodology in-house with limited consultation with communities and have then recruited and trained community members to collect the data to show ‘participatory inclusion’. But they didn’t include (centre) them when it mattered - even before the design stage.”*

Community involvement in the development and design of data-driven technologies would only be productive in cases where the knowledge gaps in data governance have been addressed. Without it, their participation would be uninformed and “of little value”. It would also be counterproductive if the technologies and processes for involvement do not cater for the usability and accessibility, particularly for persons with disabilities.

Numerous interpretations of power-preserving inclusion emerged. Among them: “practices intended to maintain a regime of existing data injustices”; “token representation which is not conducted to secure meaningful input from the community but for the purposes of ticking boxes or complying with formalities”; and “a box-checking exercise so that those with power can say ‘this was a consultative/participatory’ process, but they still hold the microphone”. From a policy perspective, power-preserving inclusion was described to entail the idea of ‘copy and paste policy making’. “The government policy environment is skewed in a way that excludes utilisation of valuable opinion and feedback from the public, even though it seems like the policy making process is participatory”, noted a respondent.

Among the existing development practices that were said to reinforce hierarchies despite outward attempts at participatory inclusion included:

- The tendency to stratify society into classes;
- Practising affirmative action instead of “equality of opportunities”.
- Basing access to opportunities on favouritism (nepotism) rather than merit.
- Racism and gender-based discrimination
- Failure to appreciate the intersectionality of the pillars of oppression

The questions developers, civil society, data intermediaries and policy makers might be failing to ask themselves included:

- Why are some groups of people perpetually marginalised or under-represented in data?
- Are there intersectionalities that still cause some people not to be included in data processes?
- Has inclusion catered for everybody that must be targeted by a specific intervention?
- Has the input from community members been taken on board?
- Why are people still speaking for others when they have their own voice?

### **2.5.6 Knowledge**

Not only integration but “centering” of cultural understanding and lived experiences of impacted communities in knowledge that informs data practices was indicated as important. As stated by one

respondent, “culture greatly influences community data practices. An example is the African Ubuntu principle, where individuals feel the need to share as much information as possible or tend to be easily trusting,” often at the expense of privacy. Such context-specific aspects need to be “documented and easily accessible”, if they are to inform data justice. Indeed, there have been a number of [calls for cultural and linguistic diversity](#) within the digital sphere to ensure equal and equitable access to the benefits and opportunities of the information society.

At policy level, it was deemed unfeasible to ask members of impacted communities to engage with and rationally scrutinise expert knowledge about data innovation practices and their governance which emanates from developer and policymaker communities. This was attributed to the majority of impacted communities having no knowledge and or skills to scrutinise expert knowledge. “[They] will not see a direct impact of data innovation practices to their livelihoods,” noted a respondent.

From an impacted communities perspective, inclusion of African cultural understanding and lived experiences in current data practices was deemed unrealistic and impossible given that many data governance practices and frameworks were currently developed in Europe and North America. Once exported to the rest of the world, these frameworks were often not good fits for local contexts, especially in low to middle income countries.

Nonetheless, integration would “help in interrogating and/or clarify certain issues that were hitherto not considered important” with regard to particular groups of people. As for the feasibility of engagement and scrutiny of expert knowledge about data innovation practices and governance by impacted communities, this was subject to training in data rights and governance. As emphasised by one respondent, “It is feasible if the knowledge in question can be broken down and explained to members of my community in the required details.” Another respondent stated that feedback integration from impacted communities was not only feasible but “should be expected”, adding that developer communities should seek out feedback and “ensure they act on it appropriately - bearing in mind that impacted communities are not peripheral stakeholders but actually active team members.”

But obstacles to confront and challenge knowledge include high levels of illiteracy and poor experience in data processes. The external obstacles highlighted were exclusion from positions of power and other forms of marginalisation that have perpetuated inequality in society. For their part, civil society and developers urged for funding structures to support community-led approaches to data processes i.e “focus less on pre-defined outputs and appreciate the value of participatory, agile, responsive processes.”

On strong objectivity addressing community needs and relevance, respondents from impacted communities were sceptical. “Strong objectivity is a welcome addition, but I am not sure how much more difference it will create for community members,” said one. Indeed, the concept of strong objectivity whose origins are feminist was found not to differ from other marginalised communities pre-existing conceptions of objectivity such as the “nothing about us without us” and “leave no one behind” mantras. As such, emphasis should be placed on recognition, reflection, analysis and explanation of what effect bias has on outcomes.

The proposed topics for knowledge exchange with other impacted communities to advance data justice were:

- Intersectionality and how it affects data practices
- Inclusive development
- Real life consequences of data processing

## **2.6 Data Justice and the Sustainable Development Goals (SDGs)**

According to one respondent, “those most powerful are collecting data on those most-vulnerable and using it to inform programming and policy, and to try to report on progress towards[SDG] goals.” Therefore, data practices that are “not fair, not representative and do not portray impacted communities in the best possible way” present missed opportunities for advancing the SDGs. For instance, inability to access affordable financing and lines of credit may perpetuate “the cycle of poverty” and threaten SDG 1 on eradicating poverty.

Overall, the data justice pillars were said to be applicable to the SDGs, with participation and equity receiving most mention. “The pillar of equity is particularly applicable to SDG 5 because it centres on disadvantage and discrimination” said one respondent.

“Data must be seen to promote justice, equity and be used to improve livelihoods and not take from them.” Monitoring of implementation of SDGs pertaining to specific communities and country reporting on progress is impeded by the absence of relevant data.

### **Shortage of Disaggregated Data**

*Whereas the national census exercises in the region collect data on the number of persons with disabilities (mostly numbers by type of disability), there is a stark lack of disaggregated data on persons with disabilities beyond this level, which makes it difficult to plan for the community. For instance, while the technology sector is key to improved livelihoods and inclusion for persons with disabilities, countries in the region [do not collect data](#) that shows how the nature of disability affects access and usage of technology. Persons with disabilities are not a homogeneous group and the nature of disability influences how they may perceive, be able to access and to use ICT. The lack of comprehensive disaggregated data, including the specific challenges that persons with different types of disabilities face in accessing information and using technology, also undermines the design and implementation of interventions that would improve their access. This lack of disaggregated data in the technology sector is largely replicated in other development sectors, such as education and health services delivery, and employment.*

*Notably, Goal 17 of the SDGs stresses the need to increase the availability of high-quality, timely and reliable data that is disaggregated by disability, in order to strengthen the means of implementing social development programmes. As such, all progress made by the SDGs [must be monitored](#) through disability disaggregated data, a measure that is consistent with article 31 of the [Convention on the Rights of Persons with Disabilities \(CRPD\)](#). This article provides that states shall undertake to collect appropriate information, including statistical and research data, to enable them to formulate and implement policies to give effect to the Convention. As per article 31(2), the information collected shall be disaggregated, as appropriate, and used to assess the implementation of countries’ obligations under the Convention and to identify and address the barriers faced by persons with disabilities in exercising their rights. The [Africa Disability Protocol](#) also requires countries to ensure the systematic collection, analysis, storage and dissemination of national statistics and data covering disability to facilitate the protection and promotion of the rights of persons with disabilities. It enjoins states to disaggregate statistics and data on the basis of disability, gender, age and other relevant variables, including by ensuring that the national population census captures data on disability.*

### 3. Conclusion and Recommendations

The concept of data justice was not well known among the stakeholders that were consulted. However, the stakeholders were conversant with the constitutive elements of data justice and the principles. Indeed, in describing what data justice implied to them, the prominent terms given by several respondents included words such as participation, empowerment, fairness, inclusion, transparency, equity, empowerment, and representation, which are core to the principles. Moreover, even without being able to define data justice, there was widespread ability to state examples of how data injustice plays out in reality in the different African contexts. Further, all stakeholder groups (policy makers, affected groups, intermediaries, innovators) were able to articulate the groups that stand to suffer most from data injustice, such as persons with disabilities, women, ethnic minorities, and migrants.

The consultation also pinpointed some of the drivers of data injustice in the region, with the increasing use of technology seen as a double-edged sword that could alleviate or exacerbate injustice, similar to the digital IDs that many countries are rolling out but which many vulnerable citizens are unable to access. Poor or missing legal protections for personal data, and abuse of existing laws by state agencies and private companies further [exacerbate](#) the erosion of citizens' data rights, and could deepen data injustice. Emerging prominently among the needed actions are the deliberate and consistent collection of disaggregated data, progressive data protection protocols at country level, engagement with platforms so they adhere to local laws including on cross-border transfer and storage of personal data.

Capacity challenges among policymakers were apparent, and it is crucial to grow this capacity for them to meaningfully promote and superintend data justice. Data protection authorities should be a primary target of any such efforts, and could be engaged as champions to push the data justice agenda throughout government ministries, departments and agencies. That effort needs to go hand-in-hand with awareness raising and capacity development for impacted communities so they appreciate their data rights and have the agency to claim them, exercise them, and advocate for them. The research found that data intermediaries, innovators and civil society (notably digital rights groups) have better awareness about data justice than public officials or affected groups. Yet this group itself needs more grounding in the data justice principles and their applicability at national and regional levels.

Some regional instruments speak to some of the principles in the ADJRP Six Pillars of Data Justice , notably Access and to a lesser extent Equity. However, they are largely on Power, Knowledge, Participation and Identity. It is noteworthy that even on the principles which these instruments touch on, the framing does not clearly engender rights respecting and participatory data justice the same way it is encapsulated in the pillars.

Accordingly, the following recommendations are made:

- Engage with regional processes, such as the African Union Data policy framework to embrace and promote the data justice principles, since actions at that level can influence policy and practice at national level.
- The principles should be framed in simple language that is easily understandable to the average stakeholder as opposed to the current wording that is characterised by use of several technical and academic-sounding terminology and phrasing.
- The data justice principles should be aligned with initiatives that promote data rights and digital rights in order to place them on the agenda of a wider range of actors that advance them.
- To popularise the principles at national level, it may help to consider making popular versions in

local languages.

- The pillar of power should add an element of influence to it as there can be power without influence.
- Each of the proposed pillars should apply to every stage of the [Data Value Chain](#) from collection to analysis, dissemination, and the final impact of data on decision making.
- The pillars should include specific mentions of visibility and equality.
- To ensure application of the pillars, there is a need for consistent advocacy and engagement on the collection of disaggregated data that represents diverse groups such as those most affected by data injustice.
- The role of incentivisation should be critically explored within the data ecosystem and should inform the evolution of the pillars.
- Another area for exploration is inefficiency within data ecosystems as an injustice - multiple data collection points and little harmonisation especially across government data-drive initiatives/programs. This disadvantages individuals outside of marginalised groups in the form of lost time, additional economic costs and lost opportunities.
- As part of the push for application of the pillars, avenues for access to legal aid for data injustices/impacted communities should also be explored.

**Annex 1: Respondents**

<b>Affiliation</b>	<b>Country</b>
Personal Data Protection Office	Uganda
Kyambogo University	Uganda
Community Rehabilitation Network	Uganda
Consumers International	Worldwide (in Africa, Zimbabwe)
Data4Change	Worldwide (in Africa, Ethiopia, Uganda, Tanzania and Egypt)
Civic Tech Innovation Network	South Africa
Development Initiatives	Worldwide (in Africa, Kenya and Uganda)
Data Protection Commission	Ghana
Communications Commission	Nigeria
Ministry of ICT, Postal and Courier Services	Zimbabwe
Investment Fund for Electoral Communications/ Universal Service Access Fund (USAF)	Ghana
KTA Advocates	Uganda
BudgIT	Nigeria
Pollicy	Uganda
Paradigm Initiative	Nigeria
Centre for AI and Digital Policy	Botswana
Sahara Ventures	Tanzania
Lagos Business School	Nigeria
Research ICT Africa	South Africa
Financial Inclusion Forum Africa	Ghana
Collaboration on International ICT Policy for East and Southern Africa (CIPESA)	Continent-wide