DATA LOCALISATION IN SOUTH AFRICA: MISSTEPS IN THE VALUING OF DATA

Gabriella Razzano
CONTENTS

1. Introduction 1
2. Backstory 1
3. The political question 1
4. The value(s) of data 3
5. From value to control (objectives in governance) 4
6. Understanding localisation through control 5
7. The law and data ownership 5
8. Roles of law and policy 7
9. Back to the Policy 8
10. Conclusion 9

Endnotes 10
1. INTRODUCTION

Those who work with data are perennially calling for 'nuance' and the abandonment of 'unhelpful metaphors': Data is the new oil versus data is not the new oil becomes a war of similes over substance. This policy brief will explore where generalisations are unhelpful, but will also seek to identify some that can nevertheless aid policy and law-making.

Central to this pursuit of policy clarity is a discussion of how ownership concepts merely further muddy the waters. In considering the South African policy context, the policy brief examines the National Data and Cloud Policy as an asset for analysing not only political patterns in respect of data governance, but also shortcomings in notions of data ownership for recognising the full range (and realities) of benefits (and risks) of data in South African life and in the South African economy.

Data governance is not just about protection, but also access and interoperability ... and security and trust ... and, this paper argues quite fundamentally, about control. The policy brief will explore how, instead of being a random association, the merging of ideas of data ownership and localisation – when explored together – help outline some of the clear political and technical challenges in the South African approach to data governance.

2. BACKSTORY

In 2021, the South African Department of Communications and Digital Technologies (the Department) opened its draft National Data and Cloud Policy (the Policy) for comment. The document could be noteworthy in different ways – for one, it demonstrated the readiness of the Department (and national government) to intervene in the data economy, a readiness which has often been shown in the regulation of the digital economy (through information and communications technology (ICT) regulations), and in emerging regulatory intersections like the Competition Commission's recent draft document on Competition in the Digital Economy. Economic regulation is not surprising in the South African context, though it is sometimes useful to remember the foundations of what a policy is seeking to build. Across digital and data services, and content issues, the call for regulation has followed not just international concerns, but also responds to local challenges in participating in this globalised space. When South Africa's Parliament recently attempted to call 'social media companies' before it to answer questions on the role of platforms in relation to misinformation, Google presented, but Facebook chose not to without other companies answering too, with (the now former Member of Parliament) Phumzile van Damme noting: 'I will admit some idealistic naivety in my belief that self-regulation [of social media] can still work if improved.'

More noteworthy in relation to the Policy, however, was its content: the document demonstrated a strong protectionist agenda at different points, even while trying to provide nods to ideas from around the African Continental Free Trade Area (AfCFTA). Those nods were perhaps strongest in the narrative, though not consistently driven in the recommendations, with an emphasis on the use of data infrastructure development as a means of deriving local digital benefits accompanied by support for direct public investment in broadband infrastructure and Cloud capabilities, all balanced with steady acknowledgements of open data and open systems as mechanisms for furthering these benefits. However, within the policy issues specified on Localisation and Cross Border Data Transfers, it makes a relatively novel proposition, proposing a form of state ownership of data:

1) Data generated in South Africa shall be the property of South Africa, regardless of where the technology company is domiciled.

2) Government shall act as a trustee for all government data generated within the borders of South Africa.

While there is mention in the narrative of the government acting as a form of 'trustee' of data, the property assignment in the recommendation is very clear. Before unpacking both the implications for South Africa's policy position and challenges to this position, the political ascendancies should first briefly be considered.

3. THE POLITICAL QUESTION

South Africa's first governing party, like many of the parties that led independence movements in Africa, was nationalist in character – and the African National Congress (ANC) has democratically held power in the country since the first democratic elections in 1994. Academics have noted how these nationalist tendencies have frequently overridden social priorities in domestic policy and law (though not expressed in concurrent asset nationalisation). And this nationalism has frequently manifested in policy and authority centrism, with academics noting: 'Fairly early in the ANC’s first term of office, Mbeki, then deputy president, began to control policy development through the Office of the Deputy President. The venues designed to facilitate civil
society and parliamentary participation exercised little clout. That centrist has often been at the exclusion of multi-stakeholderism in policy development, with the limited role of the National Economic Development and Labour Council’s influence in economic and labour policy in practice standing as a good example.

These nationalist and centrist tendencies present a challenge to regional policy priorities in the area of data and digital governance, and other forms of economic governance too. There is no doubt that many in the ANC, and certainly President Cyril Ramaphosa, have been vocally supportive of the ambitions of the AfCFTA). And recent years have seen concerted efforts to rehabilitate our regional image. But the patterns in policymaking mentioned here are not deterministic, instead seeking to highlight the conflicting political agendas that can influence data policy decisions.

Global dimensions impacting politics must also be considered. The dominance of global platforms and companies in the digital sphere has been identified as a significant inhibitor to digital economic progress on the continent. Yet these diminutions of African influence are not only economic, but also political. Geopolitical relations that govern data and data policy at a global level tend to exclude African perspectives in their discussion and design. Nevertheless, for local dividends to be reaped from a data economy underscored by facilitating data governance, participation in global governance mechanisms must be prioritised for African member states. National policies must be contextualised within a global data ecosystem. Yet national political agendas have sometimes impeded African participation in global trade negotiations and multilaterals, and this has extended to critiques of levels of engagement in Internet and data governance structures as well.

Data governance is not just about protection, but also access and interoperability … and security and trust … and, this paper argues quite fundamentally, about control.

These broader reflections certainly have resonance in the structure of the Policy in question. Visions for a grand, central-government data infrastructure echo those of the grand ‘panopticon’-style centralisation of digitalised, national identity ambitions like the Department of Home Affairs National Identity System (HANIS) (and the subsequent development to the Automated Biometric Identification System (ABIS) system), and much of its technocentric emerging visions in the area. As Klaaren highlights in his policy brief (which also examines the Policy), the ‘over-reliance on state capability in South Africa’ is a position that does not augur well for success.

However, this centralised technocentrism is not just a challenge of capacities and capabilities but is also a challenge of trust and governance. The move to the ABIS system was another example in the South African public-sector technology ecosystem of where reliance on external technological capacity was tainted in its tender and procurement challenges, derailing the national identity project in the past at several points, including the contract disputes that arose in the development of the HANIS system, and painfully reminiscent of a number of other such collaborations.

Certainly, though, the Policy is not emerging in a data-governance vacuum – the Protection of Personal Information Act (POPIA) of 2013 is finally fully effective as of 2021, and Regulations and Guidance Notes are being issued with relatively frequency. Yet, even after the passage of POPIA, and prior to the drafting of the Policy, the Department of Trade and Industry-funded Industrial Development Think Tank had noted that a central dynamic of data in the digital economy was acknowledging the need for ‘a clearly defined set of policies on data ownership, data quality, data categorisation and anonymity’. Data governance (given the profound connection to data’s value in its use and in the practice of its management) and the increasing availability of digital data are driving a shift in policymaking worldwide from being data-informed to being data-driven.

Still, the Policy seems unable to reconcile the need for first establishing broader governance trust prior to the prioritisation of protectionism and centrist. Consider, for instance, that the Department’s acknowledgement of POPIA, and exceptional focus on data governance in seeking to lay foundations for data digital infrastructure, nevertheless only mentions South Africa’s Information Regulator (established in terms of POPIA to oversee both personal data protection and access to information) only twice, with recommendations to accompany the Policy with ‘a review of the existing regulatory authorities’ mandates with a view to establishing a single data regulator, reporting to the Minister of Communications and Digital Technologies’.

The approach is the result of a political position that sees centralised consolidations of state power in relation to data as foundational to good data governance. There are strong historical and regional dynamics that have led to this position, yet the reminder is simply this: particular political ideologies form the foundations of this Policy that are particular, and not incidental.
4. **THE VALUE(S) OF DATA**

It is worth turning to look at data as the subject of policy (and law), and its values and nature, before considering some of the positions on data that the Policy reveals. The economic dimensions of data are increasingly of interest to policy commentators as conversations on data economies drive the impetus and focus of study. And economic dimensions are well framed with respect to the consideration of its nature. Data, generally, is non-rivalrous (at the technical level, it is infinitely usable without detracting from another person’s ability to use it). Its excludability is partial: excluding others from access is a question of continuing financial investment in practice, bringing it slightly outside the definitions of a ‘public good’ to a place where the law (through both intellectual copyright and contract) is often used to drive the economic value of data precisely by seeking to make data proprietary and ‘inaccessible’ to those who do not contribute to the maintenance of its value. Although there are attempts to render data excludable through technological and sometimes legal means, these are not inherently features of data.

Expenses on data’s generation and control are investments because of its value for firms and controllers. It is the collective nature of data collection (and the ability to store it and analyse it collectively) which actually drives economic value for firms – big data is the source of the dominance of the GAFAM (Google, Apple, Facebook, Amazon and Microsoft) and BATX (Baidu, Alibaba, Tencent and Xiaomi) companies, alongside the capacity to interpret that data. Though some firms get direct economic value from selling access to data, such as through data brokerage and data-as-a-service, globally that value appears only secondary to the economic value that the dominating firms can extract from interpreting that data and feeding it into their own product and service design, which relates to economies of scale. The other dominant model for extracting value is through targeting marketing and advertising, which heavily focuses on the collection and aggregation of personal data for extracting models for segregating audiences, such as in Facebook’s business model. This is why, while it is non-rivalrous technically, it is still treated proprietarily.

Yet the idea that simply gathering more and more data creates economic benefits for companies does not recognise the microeconomic realities of data, in spite of significantly influencing data practices (there are, however, economies of scale in terms of the data-collection costs for companies). This is because network effects related to data are not the same as data network effects, and, in fact, data is more often connected to scale effects for business, which can be overstated. So value at the level of the firm requires more consideration of context. Still, while economists can better articulate and debate the finesses of these dimensions (especially considering introductions like this joyful gloss of the heterogeneous nature of data), it is worth highlighting simply that there are complexities born of trying to determine the value of data outside of the technologies they drive, and are driven by, with data’s economic value sitting as a factor (input) of production. Deriving economic value requires infrastructure to control, and create, but it must always be remembered that there is inherent complexity to data value chains.

While that speaks to data more broadly, what about the economic value of personal data more specifically? Recent calls for models which allow individuals to monetise and sell their personal data, or benefit from a ‘digital dividend,’ have been met with scepticism, as the individualised value of that data would practically be paltry. This is not least of all; as has been stressed here, it is collections which create benefits (for scale effects, for learnings, and so on.). And, economically, devising a price system for goods that have zero marginal costs is a difficult (and useless?) endeavour. In seeking to determine personal data’s value in the form of a licensable asset, there are challenges in determining its provenance and traceability. Yet the personal costs of data protection seem to be a necessary part of the calculations for governance given that they are born by the subject, while the economic value goes elsewhere. This has a regional, macroeconomic dimension too: the Global South is being ‘computationally appropriated and siphoned’ to power technologies from outside of the continent, with little safeguarding or economic exchange. This tendency towards ‘extraction’ is facilitated by the nature of the economic value of data, which creates tendencies toward concentration.

There are complexities born of trying to determine the value of data outside of the technologies they drive, and are driven by, with data’s economic value sitting as a factor (input) of production.

Yet the emergence of regulatory conversations in the data-governance space must be preceded by understandings of the social and public values of data, not merely its economics. Data is the foundation not just for information, but for knowledge. Information and knowledge are instrumental to distributions of political power, for individuals and groups. The social value of data connects it as well to its influence in sustainable development and service delivery; with the Sustainable Development Goals reiterating the facilitative role of open data within development goals, whilst centring...
the role of statistical agencies in the measurement of progress. As noted in emerging policy instruments, this vision of data sees it as allowing governments to deliver more effective public services, facilitating effective environmental stewardship, and letting citizens live healthier and economically secure lives. These social development visions frequently centre, of course, on forms of statistical and biographical data for policymaking, but extend to issues like reliable, ‘Good ID’ as an expression of personal identity for the provision of services (though these digital identity components should be met with caution in the regional context). With the significant infrastructural shortcomings that still mar the lived experience of African citizens, considering the opportunities of data (and its foundations for digital services) for ‘leapfrogging’ development challenges remains tantalisingly pertinent. When we frame data value as data benefits to people, the question then becomes how to best derive these benefits, whilst mitigating against associated data risks. In other words, as insightful as economic valuations are to understanding the incentives at play, data governance might be better understood through the aim of regulating data for maximising social, rather than economic, benefit. Of course, social and economic benefits are not mutually exclusive. After all, these mutual ambitions are seen as the foundations for achieving data justice holistically. Social and structural dimensions have seen expression in development economics as well – with the capabilities approach noting that an individual’s rights and freedoms (which include rights to privacy) are insufficient without the actual capability to achieve them. This is not simply ‘an access to justice and process’ capability (i.e. having the place to act on your freedoms), but also entails having the resources (material and otherwise) to take opportunities to enact those freedoms. These visions of data justice centre broad ambitions for a datafied world, which better understands that data governance is not just about regulating an economy, but is a metaphysical component of the expression of control for one of the objectives of governance (mitigation of risks and harms). Facilitating different kinds of access and interoperability becomes a necessary expression for the deriving of benefit. Klaaren, in his policy brief, notes how data portability has begun to dominate competition instruments in advanced digital societies, to the preference of consumer protection as an overriding policy concern.

So how does data ownership intersect with these objectives? It is worth reflecting on the point already made that most companies are not making money from owning data (whether an individual’s or a machine’s); they are making money from using it. This is not to say that individuals, companies and governments do not seek to lay claims to forms of ownership over data, but it is merely a reflection that the value derived is not frequently from a direct trade or exchange, which has much to do with data’s non-rivalrous nature.

5. FROM VALUE TO CONTROL (OBJECTIVES IN GOVERNANCE)

When discussions of data governance start from an appreciation of the varied values of data, it becomes clear that deriving benefits at the level of the individual, firm and state requires different forms of data control, whilst necessitating mechanisms to facilitate the flow and sharing of data. And control is an important concept for contextualising protection, and in many ways is more instructive than ownership in this context too. An example of this with regard to data types is that in relation to personal data. Personal data protection has frequently been realised in data protection through notice-and-consent processes – and these processes and ‘opt-ins’ tend to be the main method for individuals exerting their agency in relation to their data, with much of that language having emerged from consumer protection discourse. In South Africa, POPIA defines consent as ‘any voluntary, specific and informed expression of will in terms of which permission is given for the processing of personal information’. Recent writings have sought to provide more specificity to the notion of consent in a digital context: ‘…consent is most valid when we are asked to choose infrequently, when the potential harms that result are easy to imagine, and when we have the correct incentives to consent consciously and seriously’. Yet, it suffers through a chief legal fallacy: that of equal bargaining power for fair exchange. A focus on individualised consent as the mechanism for the exercise of a freedom, has one net result: ‘Consent without power leads to inequality’. And besides the many contradictions and challenges with regard to notice-and-consent as a mechanism, what needs to
be recognised is what is meant by the very concept of privacy itself, which ‘implies that individuals have control not only over who communicates with them but also who has access to the flow of information about them’. It is the pursuit of control of personal data that has been a key driver in many practical approaches to the exercise of data protection for individuals. An additional challenge in this frame is that the ‘notice-and-consent solution’ places all the responsibility on the individual actor in a manner which is neither fair, nor practical. This is why data-governance frames are broader in terms of the mechanisms with which they seek to exert protection; yet it is important to note how these mechanisms may be better framed in seeking control as a priority, rather than ownership as a priority.

6. UNDERSTANDING LOCALISATION THROUGH CONTROL

The pursuit of data sovereignty is essentially state-level efforts to exert control over data and its flow. The realities that underscore data flows and extraction have often raised questions in the region about the potential for data localisation, as a form of extremity in a data sovereignty ‘spectrum’. Recent regional data policy conversations have frequently centred on localisation as a governance mechanism.

Though there are benefits associated with the possibility of organisations (and states) being able to easily share data across borders, a number of countries – across every stage of development – have erected barriers to cross-border data flows, such as data-residency requirements that confine data within a country’s borders, or data localisation. Data localisation can be explicitly required by law or can result from other restrictive policies that make it onerous to transfer data legally, such as by requiring companies to store a copy of the data locally, to process data locally, and mandating individual or government consent for data transfers. It can also be sectoral, that is, only related to data within a particular sector or function.

It is unsurprising, perhaps, that the Policy, whilst meaning to be focused on data infrastructure, includes localisation provisions. These localisation provisions are an attempt to assert sovereign data control (with the data infrastructure providing the means to do it), underscored by an assumption that state centrism is a mechanism both for ensuring protection and for preserving value retention. Yet the Policy is both acknowledging a shortcoming in public (and even private, local) digital infrastructure, whilst mandating its use. In a region where a lack of physical and digital infrastructure remains a feature of its capitalist reality, the development of public infrastructure will not be sufficient to meet demands the Policy is acknowledging already exists, and effective public–private partnerships remain a necessity. Nevertheless, any duplicated cost for mandated domestic data storage will certainly be borne chiefly by local companies and firms.

It is the pursuit of control of personal data that has informed many practical approaches to the exercise of data protection for individuals.

But the aim here is not to assess the mechanics and economics of the specific data localisation proposal, which other papers like that of Van der Berg have addressed before, but, instead, it is to explore some of the relationships between localisation and ownership, and underlying problems in those areas, for considering a more data-centred perspective in the further development of law and policy.

7. THE LAW AND DATA OWNERSHIP

Whilst earlier I addressed the economic nature of data, it is worth considering the legal nature of data in the context of ownership. The non-rivalrous, but also incorporeal, nature of data precludes the categorisation of data as a type of physical property as understood in most legal systems in Africa. This is partially due to challenges of provenance (if you can copy it without detracting from its value, how easily can you define its ‘original’ owner?). In the digital content arena (as an expression of data), the use of non-fungible tokens is emerging as an exciting salve for this legal challenge, but that is not of utility to data more broadly outside of content products. As incorporeal property, it could be managed within intellectual property and copyright law, but the fit is not necessarily an easy one.

In South Africa, the law related to ownership of data exists in two main fields currently: (a) the ‘in principle’ vesting of ownership of one’s personal data to the data subject through POPIA; and (b) the allocation of proprietary rights to datasets through existing law on copyright. The proposal of a state form of data ownership expressed in the Policy is a completely novel one, and the Policy is unable to articulate well the legal basis of such vesting might be, instead acknowledging expressly that ‘it remains unclear how data generated
through intellectual activities would be correctly categorised in terms of Intellectual Property Rights’. And, of course, as a factor of production, that ‘intellectual activity’ may be a few steps removed – in the case, for instance, of data accumulated from robot or mechanical activity. Certainly, there are utilities in considering data as labour for the consideration of foundations for rights, rather than as capital or assets.

How data is coalesced into a database is often a significant question, given that the database is what allows the information to be sorted, traded and used.

Whitting down from ideas of data ownership to personal data ownership can provide more legal nuance. From the outset, of course, the provenance of personal data is a somewhat simpler question to answer than of other data (my personal identifiers may not have been generated by me, but I am an easily identified owner). The POPIA framework has sought to be asserting the vesting of ownership of personal data in the subjects themselves, and in litigation pursued under Black Sash Trust v Minister of Social Development and Others (Freedom Under Law Intervening) [2017] ZACC 18 the Information Regulator (established under POPIA) sought an amendment into a proposed basis that ‘personal information of the beneficiar[ies] belongs to the beneficiar[ies] themselves. There is no basis in law to divest them of that ownership and vest ownership upon SASSA. This thus appears to be a data context where vested legal ownership in a data type is statutorily founded, and clear. Yet, even amongst the mechanics of the Black Sash matter, the realities of that ‘ownership’ bring incongruencies. Even if personal data, individually, is ‘owned’ by the data subject, the databases through which the data is coalesced can be subject to proprietary regimes through South Africa’s copyright regime. How data is coalesced into a database is often a significant question, given that the database is what allows the information to be sorted, traded and used.

In The Philanthropic Collection (Pty) Ltd v Girls & Boys South Africa 2017 ZAGPJHC 302, for example, the Respondent, a non-governmental organisation (NGO), was ‘interdict[ed] and restrain[ed] … from reproducing, adapting, accessing or using’ a database containing the names of potential donors. It had received access to the database because the Applicant – which assisted in fundraising – had collated the database (with a little assistance from the Respondent) as part of another fundraising event. After the event, it learned that the NGO was using that database to source new donations. In describing the award, the court held:

The definition of author has to be considered in the context of the whole event, the role of the parties during the event, the innovation and its purpose (Philanthropic para 23).

Notably, of course, the personal nature of the data is absent from the enquiry. And, of course, Philanthropic was decided prior to POPIA, and data protection was unrelated to the pleadings. Yet this is, in some ways, the very point: ownership seldom equates to an effective mechanism for data protection. As has been noted:

There is some limited protection afforded to data sets by way of database rights, but this does not protect the individual subject of the data from the collection of the data nor from its aggregation, but instead protects the holder of the aggregated database from its use by competing third parties who might want access to the compilation either in whole or in part.

Might there be other ways that data ownership forwards protection? Ownership rights are real rights, in the sense that they are universal; but, typically, we can understand the ownership of data best under the idea of personality rights, given data’s nature as non-rivalrous and incorporeal (it is an area of debate in South African law whether intellectual property rights are real or universal, so, like many legal areas we have seen, we will have to apply a legal fiction). Property ownership principles are, fundamentally, about the sharing or exchange of assets (and not just the holding of assets):

(Property rights) enable legal owners to share the benefits of their assets with third parties by way of different types of derivative interests.

This sharing notion is of interest when we consider co-production as well: so, while I might have a fingerprint, it is the technology owned by another party that turns that into data. It might be a mechanism for co-benefit, if architecturally possible. However, ‘personal data ownership’ really seeks to engage two personal (or individual) benefits: increased personal control (which might be met by processing requirements like those seen in POPIA), and the desire to derive economic benefit for individuals from their data. Yet, as addressed earlier, the ability to individually and financially benefit from personal data is mechanistically challenging, and feasibly minimal. When we consider vulnerable populations that may be at risk of exposure to surveillance, or exclusion from social benefits, through inappropriately collected or maintained personal data, and begin to consider the context of the real harms experienced given inequality, it is the notion
of increased control that becomes most relevant in data governance. In other words, personal data ownership does not forward data protection, nor does it equate to control, and its relationship to questions of flow and access are in fact largely antithetical: the ability to control one’s data is facilitated by concepts like data portability, whereas as ownership of personal data by a subject may in fact be contradicted by corporate assertions in databases or datasets. Quite simply, ownership does not seem to take us to where we need to go. Even in discussions on the facilitation of data protection, with the notion of ‘trust’ being traditionally largely in relation to ownership of corporeal property, those positing trusts as protection are seeking to circumnavigate the issues around ‘data property’ by focusing on ‘data rights’. It is undoubtable that different data would require different rights, from full portability, to access and erasure. Data trusts acknowledge rights, alongside acknowledgements of data existing not just for individual benefit, but also as a public good. The Policy itself acknowledges possibilities of trusts and stewardships as mechanisms for data governance. Trusts could be an instrument for ensuring processing compliance, but could also be used to negotiate benefits (which includes specific licence limitations) on behalf of data subjects, assisting in the ‘challenge’ of negotiation. The establishment of rights does not require ownership.

8. ROLES OF LAW AND POLICY

Briefly, some of the roles of law and policy will be considered as a final frame for reconsidering the substance of the Policy. Policies should reflect law – as law creates the standards and prescripts by which social life is engaged, policy creates the method and principles for the achievement of political goals. Not only should the methods in policy be constrained by law, but the political goals in a constitutional order should be constrained (and dictated) by constitutional norms. The law’s normative function, from a rights perspective, is a powerful creator of constraint and ambition. The law creates (or reflects) rights and obligations across actors. Regulation, as a form of law, seeks to regulate economic activity – it is this aspect of it which has often considered it largely as a form of ‘constraint’ on activity. Yet regulation is perhaps most manifest in its ambition: ‘…a binding legal norm created by a state organ that intends to shape the conduct of individuals and firms’.

It may seek not only to constrain, but also to promote, or engineer to a specific policy goal; it is a creature of construction.

Yet, the law is frequently challenged practically by data and digital contexts. The Collingridge Paradox explains, for instance, that efforts to influence or control the further development of technology face a ‘double bind’: an information problem (because, in emerging technologies, the real impacts cannot yet be wholly predicted) and a power problem – because control or change is difficult when the technology has become entrenched before regulation or law is in place. This is associated, too, with the law’s ‘pacing problem’, which refers to the notion that technological innovation is increasingly outpacing the ability of laws and regulations to keep up. As Larry Downes noted: ‘[T]echnology changes exponentially, but social, economic, and legal systems change incrementally’. This is a challenge of course, not a prescription.

Another challenge is the normative function of law in the data economy space. One of the many values underscoring the normative capacity of human rights concepts is its ability to keep our focus on human-centred and social considerations when addressing economic questions. This is, of course, predicated on an understanding of human rights that views them not primarily as a juridical matter, but as related to the dignity of individuals and for ensuring the respect and protection necessary for vulnerable individuals. This normative ‘yardstick’ also brings synergy for our constitutional order from the courts, to policy, to the streets. Yet the human rights normative lens can also contribute to obfuscation in its failure to acknowledge the need for structural changes to prevent data and AI-induced inequalities that are fundamentally linked to the neo-liberal economic superstructure that law is often insufficient in addressing because of its limited redistributive capacity. A purely rights perspective may problematically only consider risks born of AI inequalities (in particular) as issues of ‘representation and recognition’, disregarding the ‘ever-present inequality in opportunity structures’. This is why data governance should consider redistribution as a component of data justice. In South Africa, given our entrenched socio-economic rights, the nuance to these limitations can be better explored than in many other jurisdictions: this reiterates a call for economic inclusion as a necessary component of social justice, and is confirmed by South Africa’s rich academic discourse and jurisprudence on substantive equality. Economic justice and redistributive ambitions may not feature strongly in other constitutional contexts, but South Africa’s constitutional order does not theoretically shy away from notions of transformative
justice (this is reflected on later when considering socio-economic rights).\textsuperscript{86}

This exposition is perhaps simpler than the wording may suggest: it is to say that the law’s articulation of rights is a powerful instrument for entrenching normative justice concepts that are both social and economic, and that a normative foundation for policy can help mitgate the law’s inherent weaknesses in providing clarity for data’s heterogeneity (and its inability to define prescripts for ownership productively) – that clarity can be instrumented through policy (and supported by regulation).

9. BACK TO THE POLICY

The Policy seeks to assert state ownership of all South African data, under the auspices of supporting data localisation efforts. Both initiatives are seen as acts of data sovereignty, which facilitate a particular vision of state-centred development. Within the Policy itself, some foundational narratives are instructive:

*The possible implication … is that data generated in Africa and South Africa is mostly stored in foreign lands and, where stored locally, is owned by international technology giant companies.*

*It is essential to recognise that data is a tradable commodity which is a central productive force for the development of the digital economy. Technology companies primarily make use of ‘freemium’ business models, where customers access services free of charge in exchange for their data being collected. This data is then sold globally for advertising purposes. In this regard, countries have developed policies and legislation that limit unrestricted flow of data outside of their borders.*

…

*South Africa must also derive socio-economy benefits from its data. Essentially the data must be of a common good for all residing in South Africa.*

If the value of data is understood as largely economic, and it is believed that that value is extracted through direct exchange, one can see how trying to create generalised state ownership (which as of now has not existed) may be an ambition. Nevertheless, this would need to be an ambition of law and not policy, the legal foundations for which could not be assumed (particularly given the position on personal data ownership). Yet, the economic value of data is not chiefly acquired through exchange, as we have seen (even the idea that data ’is sold’ for advertising purposes contradicts the models of market segmentation driven by data which are in fact the basis of advertising business models for social media companies such as Facebook). Where data is understood as a factor of production, control becomes a priority for economic benefit – and control implicates capacities and infrastructure as the focus for policy, an ambition the Policy recognised in its narrative, yet contradicted in parts in its recommendations. Ideas of control better inform understandings of the data-governance ambitions of both protection and access, as they draw attention to the need to (a) create realisable mechanisms for the exercise of personal data protection, and (b) highlight interoperability and data portability as a means for ensuring access to data within a sound data-governance regime.

If states hope to themselves extract economic value from South African data (as opposed to facilitating other economic benefits from data), the question is not one of ownership, but of taxation.\textsuperscript{87} This taxation question is not necessarily unrelated to questions of data localisation insofar as specified sectoral forms of data are concerned. The financial-data localisation rules of Nigeria may have practical challenges, but they are proving to be a powerful legal instrument in its tax dispute with Multichoice.\textsuperscript{88} However, those motivations should be express, and the legal underpinnings clear.

Where data is understood as a factor of production, control becomes a priority for economic benefit – and control implicates capacities and infrastructure as the focus for policy.

Economic benefits for data subjects are not met through ownership, and, although firms may try to secure ownership through bilateral arrangements and forms of intellectual property rights, the reality of economic extraction is both more complicated and also more associated with the exertion of control. Having a data policy that focuses on providing secure, trustworthy and quality data infrastructure must recognise the realities of infrastructure and capacity shortcomings, and recognise that the political calls to state-centred development are not fully acknowledging of these realities – particularly when there is a dearth within the Policy of recommendations on the particular skills needed to create a trusted infrastructure, nor satisfactory recognition of public investment into research and development to create the knowledge base for these ambitions. Instead, obligations to use particular infrastructure are created before the capacity to provide that trusted infrastructure is secured.
And the ambitions of a South African data-governance framework should not be only economic, but social and transformative as well. Data protection, from a human rights perspective, is also not forwarded by notions of ownership – nor are broader data security ambitions; again, the mechanisms for control are key. Similarly, access to data, portability, and flows of data for public benefit and learning are not facilitated through forms of ownership in any recognisable way. As noted by Klaaren, the Policy, in prioritising ownership over control (and protection over access), does not adequately reflect the recommendations developed previously by the Industrial Development ThinkTank that:

has made two key proposals in respect of data localisation and data portability: (a) that localisation of data should only be enforced on a case-by-case basis for strategic sectors; and (b) that South Africa should develop a data-governance regime, which must prioritise interoperability and portability of data, and privacy protections (and, further, prioritise data-governance regulations for consumer data in healthcare; telecommunications; online search and location data; and financial and transactions data). 89

Seeking to exert data sovereignty does not require ownership (it merely requires jurisdiction), but seeking to create blanket forms of ownership and rights also suffers from creating blanket approaches to data governance that contradict the specificity required for considered solutions for control (levels of specificity which may in fact be better facilitated by considered regulation, rather than centralised, national-level policy). 90

10. CONCLUSION

The Policy takes an approach of ‘exert ownership first, and ask questions later’. Legal precedents do not provide the base for this approach, though there are political forces in South African policy which tend toward centris that would align to ambitions of centralised state ownership regardless. Yet ownership of data does not create a mechanism for extracting economic or social value for states, firms or data subjects. In short, in a world of complexity where data is heterogeneous and its benefits and risks dependent considerably on context, ownership takes us no further – particularly in relation to policy, which should be seeking to ascertain the means for attaining broad benefits (and reduced risks) in a considered response to values and natures. Ideas of control help to emphasise the need for capacities, infrastructure, demand and investment, and part of the reason the Policy is not able to fully realise these needs is not just because of complexity, but also because of its approach to value. Data sovereignty solutions, both localisation and otherwise, will also need to be developed more directly in response to these realities, rather than political imperatives alone. Most of all, it demands specificity in policy – especially given the limitations of the law in providing that specificity.
ENDNOTES


22 Savona (see note 35 below) notes how it might better be classified as a form of ‘club good’.
26 As Martinez notes, though largely because of the costs brought by increased data regulation, ‘the third-party data ecosystem is already imploding’.
28 It is definitely worthwhile highlighting that the economic value in proprietary access to data will grow with the emergence of Artificial Intelligence (AI) and AI-as-a-service models, given the dynamics of existing data monopolies, but it is important to remain focused on data economy priorities. Evgeny Morozov, ‘Will Tech Giants Move on from the Internet, Now We’ve All Been Harvested?’, The Guardian, 28 January 2018, http://www.theguardian.com/technology/2018/jan/28/morozov-artificial-intelligence-data-technology-online.
31 Casado and Lauten, ‘The Empty Promise of Data Moats’.
33 De la Chapelle and Porciuncula, ‘We Need to Talk About Data’.
34 Martinez, ‘No, Data Is Not the New Oil’.
36 Savona.
45 Taylor, ‘What Is Data Justice?’
48 Klaaren, ‘Competition Policy and Data Protection in Africa’.
49 Martinez, ‘No, Data Is Not the New Oil’.
Missteps in the valuing of data


Savona, ‘The Value of Data’.

Klaaren, ‘Competition Policy and Data Protection in Africa’.

Discussions at a British Academy, 2018.


Worthington, ‘Legal Notions of “Property” and “Ownership”’.


Klaaren, ‘Competition Policy and Data Protection in Africa’.

Savona, ‘The Value of Data’.

Savona, ‘The Value of Data’.


Orbach, ‘What Is Regulation?’,


Gurumurthy and Chami, The Wicked Problem of AI Governance.


Albertyn, ‘Contested Substantive Equality in the South African Constitution’.


89 Klaaren, ‘Competition Policy and Data Protection in Africa.’
91 Nissenbaum, Privacy in Context.
ABOUT THE MANDELA INSTITUTE

The Mandela Institute is a centre in the School of Law of the University of the Witwatersrand. The Mandela Institute conducts research, develops policy and offers basic and advanced teaching in different areas of law. Further, the Institute conducts executive teaching, training and capacity-building through offering short-course certificate programmes, conferences, and public seminars in areas of law and policy which are domestic in operation but are impacted by global developments.

ABOUT THIS POLICY BRIEF

This Brief is part of a series of publications under the Mandela Institute’s 2021 research project on The Economic Impact of Data Localisation in Africa. This project is funded by Facebook.

ABOUT THE AUTHOR

Gabriella is the Executive Director of OpenUp, a civic technology hub in Cape Town focused on empowering people and government through data, technology and innovation. She holds a BA LLB from the University of Cape Town.