CCRED
CENTRE FOR COMPETITION, REGULATION AND ECONOMIC DEVELOPMENT

the dtic
Department:
Trade, Industry and Competition
REPUBLIC OF SOUTH AFRICA

SARChI Chair:
Industrial Development

POLICY BRIEF 9: Data Governance: Towards A Policy Framework[1]

Rory Macmillan[2]

Expert Panel II on Regulating Digital Platforms for Economic Development

Introduction

The Sustainable Development Goals (SDG) depend on the effective exploitation of data across numerous sectors. The wide host of issues relating to how data is to be governed in society today, whether globally, regionally or nationally is referred to here as data governance – a framework of policies, laws, regulations and processes that enable, guide, sometimes limit, and hold market participants accountable for, the collection, use and sharing of data.

Information and trust

Numerous systems gather and organize particular data to increase confidence in its reliability, rendering it useful for economic decision-making. Identification systems, consumer and credit reporting agencies, financial markets and securities exchanges, healthcare systems, education institutions, media organizations, judicial and other dispute resolution systems, all rely to some degree on certain rules about the organization and sharing of information. Some of these involve making data available while others restrict the flow of data. Trust is central to many dimensions of data governance – and not trust in the accuracy of data, but in the systems that collect, use and share it.

Data, opportunities and risks

Data is non-rivalrous (it may be used by multiple persons for multiple purposes without depleting it), and is replicable and transferable at relatively low incremental cost. Technologies using it have the potential to achieve radical improvements in transport, health, financial services, energy, education and other key areas of economic and social life. Access to data can also reduce information asymmetries in business and between citizens and government, and can empower individuals.

The nature of data also introduces important risks. It is vulnerable to unauthorized access, theft and manipulation. Insecure processing, storage and transfer of data may weaken its value and usefulness, and even expose critically important infrastructure and services to serious risk. Access to data about individuals may increase a government's, a corporation's or others' power over them. Disclosure of personal data can leave individuals vulnerable to an array of potential harms, including discrimination, identity theft and violation of privacy. The opportunity and incentives to use big data for private profit or public welfare may put strain on public and private goods such as competition and privacy.

From protection to production

An effective data governance policy should address the risks inherent in its nature while ensuring that a wide variety of data will be used to its greatest economic and social potential.

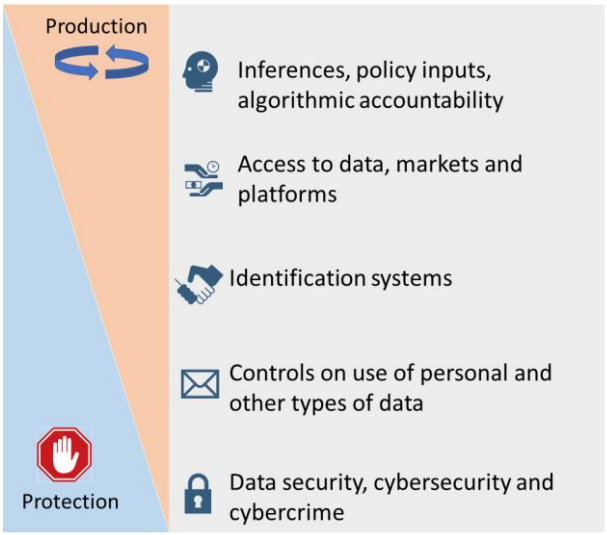
---

[1] This paper is contributed to the second Expert Panel on Regulating Digital Platforms for Economic Development under the auspices of the Department of Trade & Industry and the Centre for Competition, Regulation and Economic Development of the University of Johannesburg.

[2] Macmillan Keck Attorneys & Solicitors, rory@macmillankeck.pro

Good data governance recognizes data's economic and social utility and aims to deliver welfare benefits across the economy, doing so securely where security matters, all while respecting increasingly widely recognized norms of consumer protection and privacy.

Figure 1: Data governance seeks productivity founded on protection



Source: Rory Macmillan

The range of data governance issues is illustrated in Figure 1. At one end of the spectrum are foundational protective measures designed to prevent harm to systems and persons. These are prerequisites for the use of data in business, government and society, where data governance seeks to facilitate the economic productive use of data by establishing and maintaining incentives and minimizing barriers to such use. The key areas explored in this paper include:

- the security of the computing systems and telecommunications networks on which data is processed, stored and transferred to protect against unauthorized access to, or modification or theft of data (**data security, cybersecurity and cybercrime**);
- the conditions on which data about individual persons and other data may be collected, processed, retained and shared with a view to ensuring a digital environment in which populations are confident to participate (**controls on use of personal and other types of data**);
- the organization and processing of data on select attributes of individuals, legal entities and objects to enable their identification for the purpose of transacting and otherwise interacting with one another on a trusted basis (**identification systems**);
- access to data and its free flow across national and organizational boundaries, and the effective functioning of markets and business processes in which data is a central feature, with the goal of ensuring that data supports innovation, competition and trade (**access to data, markets and platforms**); and
- organization of and accountability for the use of outputs from data processing that produce insights for commercial and policy making, automated processes and decisions (**inferences, policy inputs and algorithmic accountability**).

Data security, cybersecurity and cybercrime

Many public and private infrastructure systems of vital importance to the economy depend increasingly on digital systems that host, process and transmit their data. These include

payment systems, banking networks, defence systems, electricity networks, hospitals, data centres and telecommunications networks for example. These are threatened by intentional breaches of data systems to extort money, disrupt government or business, influence political processes, and cause personal harm. Other threats include natural disasters and human error.

Data security involves sets of practices and techniques to limit the risk from these threats and allow for recovery of any lost or altered data. A high priority for a country like South Africa, which can host data processing and export data-driven software and infrastructure services that use its data processing capabilities, must be to ensure a stellar reputational level for data security.

This requires ensuring effective cyber security readiness procedures and expertise that traverse public and private sectors. It also depends on building the human capacity of policy-makers, legislators, judges, lawyers, prosecutors, investigators and civil society with regard to legal issues relating to cybercrime. This requires a multidisciplinary, multi-stakeholder, public-private approach and assessment to prioritize how it allocates scarce, capacity-building resources.

Controls on use of personal and other data

Personal data protection and privacy concern limits on who should be authorized to have access to or to alter or share personal data (which typically relates to attributes of identified or identifiable living individuals), and the conditions on which they may do so.

Privacy and personal data protection may concern competing claims to information, and may influence relationships of power, in both commercial and political contexts. Controversy over these ideas often concerns political and economic ideologies. Notwithstanding these debates, privacy and personal data protection are vital to inclusive growth of the digital economy, which depends fundamentally on achieving and sustaining widespread trust in access to and use of personal data.

South Africa's POPI Act is still not yet fully in force (and will have a 12 month grace period after it takes force before applying in full), but will position the country in the mainstream of relatively well-developed data protection laws, restricting the collection, processing and sharing of personal information. Increasing engagement to introduce measures that implement privacy protections in effect in organizations will be necessary.

The following might be considered to bolster the data protection regime being established in the POPI Act in South Africa:

- requiring privacy by design and default;
- introducing required data protection impact assessments;
- taking the reasonable expectation of the data subject into account with regard to privacy and controls on personal data;
- permitting data to be used only for legitimate purpose that are compatible, consistent, and beneficial to consumers;
- introducing information fiduciary responsibilities for certain data controllers;
- encouraging the use of data intermediaries or personal data management service providers to act as an agent or guardian on behalf of the consumer;
- introducing personal data management tools conferring on the consumer greater control over data about him or her;
- treating notions of property ownership right in personal data with caution;
- forcing a change in the underlying business models that rely so extensively on personal data; and
- improving the consumer's control over use of data about him or her.

Digital identification

Privately-operated and state-operated identification systems all present data governance issues. In addition to data security and privacy and personal data protection measures required to ensure trust in the system, a digital identification system depends on an effective 'trust framework.' This combines generally applicable laws of contract and liability, data-specific laws and standards, and identification scheme rules and protocols. A trust framework that works ensures that rights and duties of participants of the scheme are clear and ensures that there are sufficient economic incentives for them to play their respective roles, such as scheme operator, enrolment agent, consumer or authentication agent.

Other issues important to the design and operation of such schemes include non-discriminatory inclusion of the population at large, interoperability with other public agencies' and private firms' systems to leverage digital identification for multiple functional purposes, the use of open standards, and ensuring competition among scheme vendors (avoiding lock-in). Recognition of digital identification systems across borders is increasingly important.

Development of such mutual recognition standards in the SADC area and across the continent may be a valuable part of the development of regional markets in digital services in which South Africa could be a leader.

Access to data, markets and platforms

The use of data for social and economic good depends on access to it across national and organizational boundaries.

Opening up data

a. Proprietary data

Access to proprietary data of private organizations may have immense economic and social opportunity. A variety of institutional forms may be used to exploit the opportunity of proprietary data developed by organizations.

Interoperability, protocols and standards are required for much data sharing to be useful at all. This involves labour intensive IT work to establish, and so costs on the organization. Concerns about expropriation of a commercial asset and weakening of competitive incentives must be weighed against the benefit to society at large. Data trusts, data cooperatives and other models are mechanisms that could be deployed to allow collaboration and sharing of data for public good in a trusted manner, whether among private entities, between public sector and private sector, and across-borders.

This will require city authorities or vertical ministries to take a lead, alongside information regulators, competition authorities and private entities involved. Consumer and other civil society organisations may be able to contribute support that also builds trust among potential participants where these are individuals. They might monitor performance, and formal auditors may be required to provide reports assessing conduct against pre-agreed criteria.

b. Open public data

Making data held by Government and other public institutions freely available and redistributable offers important opportunities for medical, climate change and other scientific research, improved organization and regulation of public and private services, and development of new digital applications and services.

Concentrated data, competition, data portability and access

a. Data concentration and competition

Regulatory policy makers are today vigorously re-examining their competition laws and enforcement practices as they relate to big data, the platform economy and artificial intelligence. Data may increase information asymmetry between consumers and firms able to extract systemic information from large datasets. It may also increase information asymmetry between successful platforms and other firms. Much of the concern about market power arises from aggregation of data through vertical and horizontal consolidation, and leverage of market power from one market to another. South Africa will have to engage with these issues as its digital economy develops.

b. Data portability and access

Data portability and access to data have been proposed as a solution for competition problems arising from concentrated data. They can reduce switching costs by enabling the consumer to make relevant data available to an alternative service provider. However, they involve challenges as data is often unstructured, or structured in different ways in different organizations, and its organisation is often sector specific. Such remedies need to be deployed only where the benefits are likely to exceed the costs, both financial and administrative. They may be more feasible in the case of some vertical sectors, such as open banking. South Africa will need to be ready to deal with these sorts of issues as platform economy grows, and is already confronting them in the areas of healthcare, financial services and ecommerce.

c. Cross-border data flows – data sovereignty, ownership and localization

Access to services across borders offers huge opportunities, particularly for export to countries whose domestic tech industries may take time to develop such services. Cross-border trade in goods in the physical economy depends on cross-border flows of information to communicate demand and ability to supply, and to manage logistics and process of transport and delivery.

The cross-border dimension presents questions for policy makers and regulators. It will be vital in particular to examine whether excessive requirements to keep data within the country may undermine the efficiency and innovation opportunities of big data, the cross-border provision of cloud services, customer relationship management and regional and global value chains. Data localization may become a tool for protectionism if it effectively prevents foreign providers from offering services in a country, and so is today a key component of trade policy. At the same time, bilateral trade negotiations with leading economies that seek to minimise data localization are resulting in a patchwork approach.

South Africa may have an opportunity for export of digital services and data processing, where the rest of the continent badly lags behind. It appears likely to benefit from a relatively liberal regime for cross-border transfers. This would be supported by greater multilateral efforts to find common ground among countries on data regulation and common rules for trade in e-commerce and digital services.

Inferences, policy inputs and algorithmic accountability

Governments and businesses can use vast data troves to build a detailed personal profile of an individual and their behaviour (preferences, activities and movements) which may be used for commercial offers, State and private surveillance. They may be used to identify an individual and to determine their eligibility for a service or product. They may bring benefits to healthcare provision, medical research, transport, education, advertising, policing and the justice system.

However, use of algorithms to make decisions based on these datasets presents a new set of risks. Big datasets drawn from structured and unstructured data gathered from multiple direct and indirect sources over time risk being inaccurate or out of date. Inaccuracies may lead to

erroneous inferences and decisions. Algorithms trained on data from past experience may reflect and perpetuate the biases embedded in historical treatment of ethnic, religious or gender groups even where efforts are made to avoid using special or protected categories of data about a person (such as ethnicity, religion or gender).

Initiatives are underway in many countries from several angles to address such problems. Various ideas are being developed and could be considered for introduction in South Africa's legal and regulatory framework. These include legislating for a right to receive an explanation for automated decisions, the right to appeal to a human, efforts to address algorithmic bias, and development of ethical frameworks for the use of artificial intelligence.

Conclusions

Data governance is not only important to protect the population but is strategically central to economic success and social cohesion in the future – near and far. South Africa has some strong measures in place, including a relatively well-developed data protection law in the POPI Act, even if it is not yet properly in force. But there are numerous steps that it should be considering taking in order to capture the opportunity presented by data. The protections are prerequisites to building the trust necessary to assure the growth potential, and such trust will only enhance the South African brand as the country pursues the realistic opportunity to be a regional and even international hub for data-centric services.