EMBEDDING OPEN DATA PRACTICE

DEVELOPING INDICATORS ON THE INSTITUTIONALISATION OF OPEN DATA PRACTICE IN TWO AFRICAN GOVERNMENTS

François van Schalkwyk, Michelle Willmers & Tobias Schonwetter

Organisations are … more likely to decouple structure from practice when there are high symbolic gains from adoption but equally high costs associated with implementation.

Peter Scott, *Institutions and Organisations*

Currently, there is little evidence-based research available to help implementers understand what precisely might be required for the success of open government data initiatives.

UN Public Administration Programme

Published in June 2015 by the UCT IP Unit

DOI http://dx.doi.org/10.6084/m9.figshare.1528088

The funding for this work has been provided through the World Wide Web Foundation 'Exploring the Emerging Impacts of Open Data in Developing Countries' research project, supported by grant 107075 from Canada's International Development Research Centre (web.idrc.ca). Find out more at www.opendataresearch.org/emergingimpacts

CONTENTS

ACRONYMS AND ABBREVIATIONS

HSRC	Human Sciences Research Council (South Africa)
IPR	Intellectual property rights
KIPPRA	Kenya Institute for Public Policy Research and Analysis
KNBS	Kenya National Bureau of Statistics
ODB	Open Data Barometer (of the World Wide Web Foundation)
ODDC	Emerging Impacts of Open Data in Developing Countries (a project of the World Wide Web Foundation)
OKF	Open Knowledge Foundation
Stats SA	Statistics South Africa
Web Foundation	World Wide Web Foundation

1 INTRODUCTION

As governments open up vast and complex datasets, the expectation is that our lives as citizens will improve as a consequence of the data being made publicly available. However, there are several stumbling blocks in the path of extracting benefits from open data. On the side of the provider these barriers may include the effort and cost required to convert closed to open data; the cost of providing a user-focused context to ensure the uptake of complex datasets; poor data quality; absence of legal and policy frameworks; a lack of capacity to implement and sustain open data practices; and resistance by data custodians to opening data.[1] On the side of the data user, barriers include lack of access, low levels of data literacy, lack of human, social and financial capital to effectively use open data, and also to open up and combine several datasets that together can create value for citizens.[2]

One barrier that may be impeding the provision of open data, and one that we believe has received insufficient attention in the research on change process surrounding open data at the organisational level, is the constellation of institutional domains in which government as a complex organisation functions. In other words, from an open data perspective, we believe that too little attention has been paid by the open data movement to the institutional dynamics of governments and other public agencies; nor has the research community drawn sufficient attention to the institutional dynamics at play in the *implementation* of open data initiatives in public agencies.[3]

Given the promise of open data – that is, the potential of open data to increase the credibility of institutions in the eyes of citizens through greater the transparency and greater accountability – we would suggest that the open data movement needs to develop a more nuanced understanding of how to institutionalise open data practices, particularly if open data practice is to become an enduring and taken-for-granted course of action by government agencies. While useful as mapping exercises, current enabler and barrier analyses on open data have not, we believe, drilled deep enough to provide reliable insights sensitive to the implementation context. There have undoubtedly been early successes on the open data supply side, but these successes are not yet unqualified successes; questions remain regarding the sustainability of open data supply and about how ubiquitous supply is across government organisations. In the African context, there are fewer large-scale success stories, although there is evidence of commitments to opening up government data and of open data portals at state and/or regional government levels.

From an institutional theory perspective, the argument could even be made that introducing new technologies (including opening government-held data) may well serve to reinforce existing institutional norms and structures – and in the case of government, command and control

1 Janssen M, Charalabidis Y & Zuiderwijk A (2012) Benefits, Adoption Barriers and Myths of Open Data and Open Government. *Information Systems Management* 29: 258-268; Magalhaes G, Roseira C & Strover S (2013) Open Government Data Intermediaries: A Terminology Framework. Paper presented at ICEGOV2013, October 22-25, 2013, Seoul, Korea.

2 Gurstein M (2011) Open data: Empowering the empowered or effective data use for everyone? *First Monday* 16:2; Magalhaes G, Roseira C & Strover S (2013) Open Government Data Intermediaries: A Terminology Framework. Paper presented at ICEGOV2013, October 22-25, 2013, Seoul, Korea; Canares, M (2014) Opening the Local: Full Disclosure Policy and its Impact on Local Governments in the Philippines. Accepted for presentation at the Eighth International Conference on Theory and Practice of Electronic Governance, October 27-30, 2014;

3 See Luna-Reyes and Gil-Garcia who make a similar point. LF & Gil-Garcia JR (2014). Understanding the Co-Evolution of Institutions, Technology, and Organizations: The Enactment of the State Government Portal of Puebla. *The Proceedings of the 14th Annual International Conference on Digital Government Research.*

rather than openness and accountability, may be reinforced as open data is perceived, used and implemented within existing institutional arrangements.[4] In addition, the rate of change demanded by movements such as the open data movement, could be described as being at loggerheads with the rate of change in institutional domains that are typically highly resilient and slow to change. This introduces the possibility of implementation, driven by political will, being forced through at a superficial or symbolic level, without penetrating and becoming a steady state of practice across an organisation.

It is the measurement of the institutional conditions surrounding emerging open data practice in governments that this research project undertakes in order to develop a deeper understanding of the embedded barriers and enablers of change.

With its focus on developing country contexts, the research of the Emerging Impacts of Open Data in Developing Countries, combined with the research proposed by this study, offers the opportunity to contribute to our theoretical understanding of change processes in institutions. In particular, insight can be gained as to the socio-technical conditions under which open data initiatives in public agencies are more or less likely to succeed in the institutional domains under examination.

4 Janssen, Charalabidis & Zuiderwijk (2012). See also the Mysciencework article *L'open data est-il un leurre politique* ? http://rue89.nouvelobs.com/2014/08/04/lopen-data-est-leurre-politique-254056

2 CONCEPTUAL FRAMEWORK

Governments – be they federal, national, regional or metropolitan – can be defined as a collection of purposeful organisations. While governments may be distinct from other organisations because of the 'legitimate coercion'[5] at their disposal as law-makers and enforcers, governments are not necessarily distinct in their structural arrangements as organisations, albeit that their arrangements may tend toward complexity in relation to firms.

As organisations, governments do not function in isolation and are subject to external environmental pressures to adapt to new environmental conditions. These pressures may come from several sectors in society or may even be driven by changes in the physical environment. At the same time, organisational actors are subject to the pressures exerted on them by their particular institutional domain. In the case of government agencies, the state bureaucracy is the most obvious institutional arrangement in play. Institutions, in this sense of the word, play a critical role as stable, enduring social arrangements that provide the rules, practices and structures that shape the behaviours and beliefs of organisational actors. In the words of Olsen (2007:3) '[a]n institution is an enduring collection of rules and organized practices, embedded in structures of meaning and resources'.[6]

Structures of resources make actors more or less capable of acting according to prescriptive behavioural rules and laws. The degree and form of institutionalisation impact both on motivation and capacity to follow institutionalised rules and codes of behaviour.[7]

Institutions enjoy a relatively high degree of autonomy; they are relatively invariant to a changing environmental context, and are not typically disrupted by a turnover of individuals within the institution. Those with agency within organisations tend to conform to shared, common, taken-for-granted values and beliefs about what is acceptable or normal behaviour and practice for the relevant institutional domain. The institution's survival depends on conformity to these rules and values. However, this is not to suggest that institutions themselves are immune to external influences; they exist and function within a broader social context and are subject to pressures that emanate from changes in society. That institutions endure – they are both sustainable and stable – is attributed to their slow rate of change as they buffer themselves against external pressures to protect their values.

Organisational actors are therefore exposed to two countervailing pressures: environmental pressures driving change and institutional pressures valuing conformity.

The structural composition, values, norms and culture of organisations should be considered when examining the interaction between institution and organisation if one is to consider how change will manifest across the organisation. In the realm of technological change, and of open data in particular, Tim Berners-Lee asserts that change 'has to start at the top, it has to start in the middle and it has to start at

5 Streeck & Schmitter (1985: 20) in Scott WR (2014) *Institutions and Organizations* (4th edn). Thousand Oaks, CA: Sage.

6 Olsen JP (2007a). Understanding Institutions and Logics of Appropriateness: Introductory essay. Working Paper no. 13. ARENA, University of Oslo.

7 March JG & Olsen JP (1989) *Rediscovering Institutions: The organizational basis of politics*. New York: Free Press.

the bottom'.[8] And Hogge contends that a three-tiered drive for change was responsible for the early success in the in the implementation of the government-led open data initiatives in the US and the UK.[9] In her analysis, the top level consists of the administration espousing the political will for change at the highest levels of government; the middle level consists of government bureaucrats; and civil society makes up the third tier. She also speculates about donor agencies constituting a fourth tier.

Along with the church and the military, universities are one of the oldest institutions. In his chapter on change in universities in *The Higher Education System: Academic Organisation in Cross-National Perspective* (1983), the social scientist Burton Clark identifies three structural levels of authority in academic systems, each with different predispositions to change: the under structure (or the academic disciplines); the middle structure (university governance structures); and the super structure (the state). Clark's three levels of authority are the structural sites of embedded meaning and resources. Two of these levels – the middle and under structures – combine to make up a university as organisation.

Clark's approach is therefore not dissimilar from the requirements for change in governments suggested by Berners-Lee and observed by Hogge. Clark's framework enables the operationalisation of the research questions in a manner which is sensitive to the unique arrangements and complexities of public organisations.

For the purposes of this research project, it is suggested that public agencies as organisations be understood as being structured in the same way as universities.

In Clark's higher education framework the super structure assumes its structural properties in the form of the state and its apparatuses. For this study with its focus on government, it does not make sense to conceptualise of the super structure as being composed of the state. However, public agencies are nevertheless subject to constant environmental change pressures from donors, global financial institutions (such as banks and rating agencies), supranational agencies (such as the United Nations or the European Union) and from attentive citizens. In this study, this typically global constellation of organisations will be taken to represent the super structure.

The under structure consists of those tasked with the day-to-day running of the various government departments and agencies, and all government and agency employees who execute tasks in fulfilling the government's mandate to the public. The middle structure is composed of senior government leadership, generally referred to as the executive (the president, the prime minister, ministers, members of parliament or the equivalents of these high-level positions) as well as administrative government entities tasked with the responsibility of running government as an organisation.

It is the middle structure that is confronted more directly by the pressures emanating from the environment and that must either buffer the under structure from these pressures or translate the pressure into action in the under structure. The middle structure assumes a mix of beliefs from both the super and under structures. This mix will vary in its leaning either upwards or downwards, and the middle structure's natural direction of change – that is, whether it will buffer or translate pressures for organisational change – will depend on the extent and direction of the middle structure's beliefs towards the super and the under structures.

The response of the under structure to pressures from the middle structure will also vary depending on the alignment of its beliefs with those of the middle structure, as well as its taken-for-granted institutional beliefs. When the beliefs of the middle and under structures are in alignment, resistance is likely to be lower and change is more likely to occur. If, however, the beliefs of the under structure are out of kilter with those of the middle structure – possibly because of the institutional norms and values held by the under structure, or a disconnect in the middle structure's understanding of the daily realities in the under structure – then resistance in the under structure is more likely to be high and change is less likely to occur.

This is a process described as "decoupling" in institutional theory. It explains how organisations are able to conform to institutional pressures at a structural level in order to achieve legitimacy but may nevertheless at an internal level operate independently and autonomously.[10]

8 Hogge B (2010) Open Data Study: New technologies. Transparency and Accountability Initiative. p. 10.

9 Ibid.

10 Scott W R (2014) *Institutions and Organizations* (4th edn). Thousand Oaks, CA: Sage. 185-188.

Different departments and agencies internal to government will respond differently to change pressures depending on (i) the influence exerted by their discreet institutional domain, and (ii) *the degree of influence* of the institutional domain over a specific government department. For example, the judiciary holds a different set of institutional values from the military; and a department of defence, rooted firmly within the military as institution, may be bound by the stronger institutional forces than may be the case in another government department such as housing.

However, it would be limiting to assume that those actors in the under structures of governments fall under the conforming influence of only a single institution. Government departments or agencies in the under structure are subject to what appear to be at least two institutional pressures.

The first is to conform to those institutional norms and values intrinsic to the government as an institution – 'the bureaucratic state' is, according to Friedland and Alfred,[11] one of the core institutions of society. The second is to conform to pressures exerted by a second additional, discreet institutional domain specific to the department or agency.

In this sense, the behaviour of actors in the under structure is not simply determined almost mechanically by the pressures of the bureaucratic state, but is more open-ended and predicated on the effects of at least two institutions on actors in the under structure.

This is in line with the perspective of 'institutional logics'[12] which suggests that organisations and their individual actors are exposed to multiple institutionalised norms, and conceives of society as an inter-institutional system.

From this perspective, change occurs at the crossroads of multiple external as well as internal established beliefs, structures and practices. As Thornton and Ocasio state: "Viewing society as an inter-institutional system allows sources of heterogeneity and agency to be theorised and to be observed from contradictions between the logics of different institutional orders."[13]

In addition to institutional pressures, is a third pressure which may have a bearing on the actions of actors in the under structure. It would be an oversimplification to presume that the under structure is immune to pressures from outside of the organisation, and only has to contend with institutional pressures, even if they are numerous. While the middle structure fulfils an intermediation or translation function between the super and under structures, it cannot completely isolate the under structure from the effects of external, environmental pressures for change.

These processes of conforming and resisting by actors located in the two organisational levels to both endogenous and exogenous pressures are presented in Figure 1.

The figure reflects how the middle structure (or government executive) can be positioned either closer to the super structure, indicating a certain degree of compliance with the super structure, or further away from super structure indicating resistance to the demands for change emanating from the super structure.

Each of the government agencies in the under structure are represented as being either closer to the middle structure or further away from it. A position closer to the middle structure reflects a degree of compliance and cohesion with the institutional effects of the bureaucratic state on the organisation as a whole. Should an agency be positioned further away from the middle structure on this vertical axis, its position could be interpreted as being indicative of a strategy of decoupling.

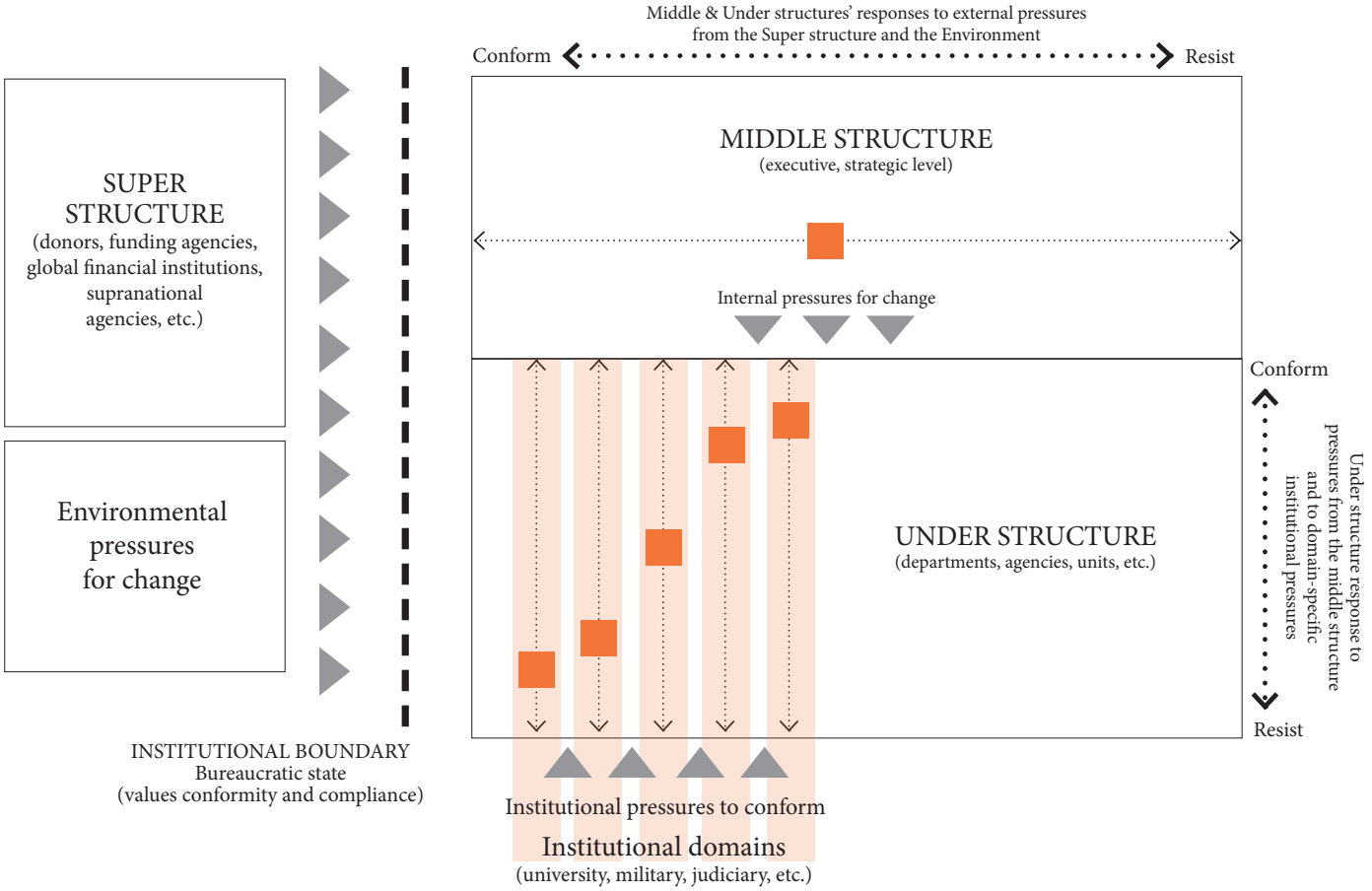
Such a position could also be indicative of the strong effects of that agency's specific institution in diluting the institutional effects of the bureaucratic state. That agencies are clustered or unevenly dispersed along the vertical axis is an acknowledgement of the varied and different effects on government departments and agencies that multiple institutional domains may trigger.

11 Friedland & Alford in Thornton PH & Ocasio W (2013) Institutional Logics. In: R Greenwood, C Oliver, K Sahlin & R Suddaby (eds), *The Sage Handbook of Organizational Institutionalism*. Thousand Oaks, CA: Sage. 99-129.

12 Thornton PH & Ocasio W (2013) Institutional Logics. In: R Greenwood, C Oliver, K Sahlin & R Suddaby (eds), *The Sage Handbook of Organizational Institutionalism*. Thousand Oaks, CA: Sage. 99-129.

13 Ibid: 104.

Figure 1: Conceptual framework



In addition to their vertical position, the departments or agencies that make up the under structure of government may vary in their horizontal positioning. They may be closer to or further away from the super structure and the environment.

That agencies or departments are closer or further away from the super structure acknowledges the bearing that external, non-institutional pressures may have on actors in the under structure; and that the departments and agencies in the under structure do not necessarily respond in the same way as the middle structure to the constellation of endogenous and exogenous pressures.

This horizontal position reveals an additional condition: it shows whether departments or agencies are to a greater or lesser degree in alignment with the middle structure. In other words, whether they are responding in the same way to external pressures for change.

3 RESEARCH QUESTIONS

The main interest of this study is to provide empirical evidence of whether open data practice is being embedded at all levels of government with reference to the institutional theory. The main question posed is therefore:

1. Is open data practice being embedded in African governments?

In order to answer the primary research question, three sub-questions are posed:

2. What are the possible indicators of open data practice being embedded?
3. What do the indicators reveal about resistance to or compliance with pressures to adopt open data practice?
4. What are different effects of multiple institutional domains that may be at play in government as an organisation?

4 RESEARCH DESIGN

In order to answer the question of whether open data practice is being embedded in African governments, the research conducted was primarily quantitative. The project developed indicators to assess the success of open data practice being embedded in government. However, the project also relied on qualitative data in order to complement the quantitative data, as it was felt that the quantitative alone would not allow reveal some of the complexities that were important to surface in relation to contextual differences between and within organisations.

INDICATORS

This research project makes use of indicators to assess whether open data practice is being embedded in governments. Indicators are a means of quantifying the complex properties or states of social arrangements such as organisations (including governments). Indicators may reflect a property or particular state either at a specific point in time or as these properties and states change over time. These properties or states are subject to the influence of extraneous conditions.

The indicators are grounded in institutional theory which maintains that institutionalisation processes can be regulative, normative, or cognitive.[14] Scott describes these processes as the three pillars of institutional theory.

Regulative processes involve formal rule-setting, monitoring, and sanctioning activities. Individuals may acknowledge the existence and even the validity of institutionalised rule systems without necessarily believing the rules are fair, right, or appropriate. Institutions function effectively as individuals determine the cost of violating formal or informal rules is too high. Thus, institutionalisation occurs as individuals find it expedient to comply with the rules.

Normative processes are grounded in a collective sense of what is appropriate. Similar to regulative processes, normative processes involve a sense of following rules. Individuals follow normative rules, however, because they perceive that following the rules is morally appropriate. Thus, institutionalisation occurs as individuals deem it socially responsible to honour informal obligations.

Cognitive processes involve widespread acceptance of the value of an activity. Institutionalisation occurs as individuals take it for granted that a certain way of doing an activity is the best way. Further evidence is provided when individuals carry aspects of the activity into other endeavours, or when other individuals or organizations adopt similar activities.

For this study we regard the institutionalisation of open data practice as requiring the widespread acceptance of practice as the best and most appropriate form of practice. The end-point of such practice can be interpreted as the often referred to "open by default". The study therefore situated itself within the cognitive pillar of institutional theory.

A set of three indicator categories were developed, each of which relates to the uptake of and possible tension inherent in open data practice at both the middle and under structure levels of government as organisation. The indicators categories are:

14 Scott (2014); Colbeck C (1999) Assessing Institutionalization: Indicators of lasting reform. 29th ASEE/IEEE Frontiers in Education Conference, 10–13 November 1999, San Juan, Puerto Rico. http://fie-conference.org/fie99/papers/1241.pdf

1. *Open data policy and regulations:* existence of approved public policy and relevant legislation [middle structure] versus approved internal policies [under structure]

2. *Open licences*: existence of policy position specifically on open data licencing [middle structure] versus the application of open data licences [under structure]

3. *Openness:* Open data readiness (commitments and resourcing) [middle structure] versus open data implementation (number of datasets being published) [under structure].

Table 1 presents the indicators by organisational level and category.

Policy and regulations: Existence of approved public policy and relevant legislation versus approved internal policies

A formal commitment to open data at organisational level is expressed in the form of policy. The process of drafting and approving a policy document to deal specifically with the release of open data is indicative of a commitment at the level of the middle structure to embedding open data practice within and throughout the organisation. Policies may vary in the degree of their commitment to openness – some policies may take a default open stance while others may be more cautious and risk averse, and may contain clauses to guide a more restrictive type of data release.

Policies alone do not determine how government and its departments and agencies are able to express commitments to open data practice. The laws of a particular country dictate the limits of lawful action, and even if such laws do not pertain directly to open data, they exert influence on how agencies and department enact open data practice. It is therefore necessary to consider both the country-specific policy and its regulative environment.

While policy and regulations may be indicative of a commitment to open data practice at a strategic level, it is at the level of the under structure that such practice is given agency. The existence of department- or agency-level policies on data sharing (or on ICTs more broadly) and on open data specifically, will be taken as indicative as evidence of open data practice being embedded in the under structure of government as an organisation.

Licensing: Formulation versus application

In all cases, the premise of data re-use without restriction or encumbrance is critical – data is only open if it is re-usable. Placing restrictions on the re-use of data by the public, civil society organisations, researchers, entrepreneurs or by other organisations in an institutional field, undermines the potential benefits of transparency, economic growth and development. It is therefore not only important that data is made open, but that the potential users of such data are clear about being able to re-use data without fear of sanction or prosecution. The application of any form of licence to a dataset may in itself seem restrictive rather than open as it appears to impose limitations, but licences provide users with an unambiguous signal as to how they may use the data. Moreover, unlicensed data is, by default, protected by the provisions of copyright law. In the case of open licences such as a Public Licence and, to a lesser extent, the Open Commons Attribution Licence, the licence provides unequivocal assurance that users are free to re-use the data made available.

A further benefit of assigning licences to datasets is that they enable use across borders and territories. According to

Table 1
Indicators of open data practice being embedded in government

	Policy and regulations	Licensing	Openness
Middle structure	M1 Existence of open data policies and legislation (public)	M2 Extent to which public policy makes provision for open licences	M3 Readiness: Degree in terms of commitments and resourcing as enabling open data practice
Under structure	U1 Existence of open data policies (internal)	U2 Extent to which open licences are applied to datasets	U3 Implementation: Number of open datasets published

the ODDC, "The creation of a unified legal framework around open datasets is seen as a particularly important issue as data travels across borders where different intellectual property rights apply to datasets, and as datasets are combined with each other. Incompatible licences, it is argued, can create significant challenges in determining the legal status of derivative datasets, yet much of the value of open data comes in combining different datasets."[15]

The act of assigning an open licence to a dataset indicates (i) an understanding of what constitutes open data, and of the fundamental principle of re-use; and (ii) that the department or agency applying the licence has come to terms with and accepted the consequences of data being re-used *with certain limited restrictions*. From an institutional theory perspective, the presence and application of an open licence could be interpreted as an indicator of a cognitive shift of re-use without restriction as an accepted component of open data. Applying open licensing could, therefore, be interpreted as signalling the critical shift from open data provision to government being open.[16] Moreover, as Hogge writes, the application of open licences could be indicative of a common set of beliefs about open data between the middle and under structures:

> This leads the researcher to speculate whether [. . .] it was in fact the contentious issue of data licensing that gradually brought together allies from the grassroots and public administrator communities, building a stronger base of expertise and shared goals [. . .].[17]

The policy presence and application of open licences will serve as the indicator of the acceptance of open data processes, and will form a core component of this study.

Openness: Readiness and implementation

There are only two indexes that seek to measure and rank the openness of governments vis-a-vis open data specifically: (1) The Open Data Barometer (ODB) of the World Wide Web Foundation; and (2) The Global Open Data Index of the Open Knowledge Foundation (OKF). These indexes provide data on the readiness of governments to support open data practice – an indicator of their openness at the level of the middle structure.

Without adequate support and resources, civil servants are not able to translate policy into practice. The allocation of resources at the level of the middle structure in order to capacitate government departments and agencies both in terms of finances and human capital, is taken into consideration in the Open Data Barometer as an important component of open data practice being institutionalised.

The indexes also provide data on the number of datasets published as well as the extent to which these datasets can be described as open – an indication of implementation and the extent to openness is manifest in the under structure. The Indexes therefore provide data that may reveal potential difference in readiness and practice between the middle and under structures in government.

OTHER EVALUATIONS OF OPEN DATA PRACTICE

The above indicators were referenced against indicators developed by three types of evaluation instruments: open data readiness assessments, open data indexes, and diagnostic tools. Specifically, the indicators of the Benchmark on Readiness for Open Agency Data (BROAD)[18] instrument developed by Step Up Consulting to determine the capacity and performance of national government agencies in the provision of open data to the public, were referenced. BROAD itself was developed based on a review of five evaluation instruments focused on open data, namely: the Web Foundation's Open Data Barometer[19], the United Nations Open Government Readiness Assessment[20], the Open Data Institute's Open Data Maturity Model[21], the Center for Technology in Government's Information Sharing Dissemination Worksheets[22], and the World Bank's Open Data Readiness Assessment[23]. In addition to the BROAD tool, the Common Assessment Framework developed through a workshop hosted by Web Foundation and GovLab NYU in May 2014[24], was also referenced, as was Ubaldi's Open Government Data:

15 Davies T, Perini F & Alonso J (2013) Researching the Emerging Impacts of Open Data: ODDC Conceptual Framework. ODDC Working Papers #1.

16 Janssen, Charalabidis & Zuiderwijk (2012).

17 Hogge (2010: 12).

18 Step Up Consulting (forthcoming). Benchmark on Readiness for Open Agency Data.

19 http://www.opendatabarometer.org

20 http://www.unpan.org/DPADM/Themes/OpenGovernmentDataandServices/OpenGovernmentData(OGD)ReadinessAssessment/tabid/1748/language/en-US/Default.aspx

21 http://theodi.org/guides/maturity-model

22 http://www.ctg.albany.edu/publications/guides/infosharing_toolkit/infosharing_toolkit.pdf

23 http://opendatatoolkit.worldbank.org/en/odra.html

24 http://thegovlab.org/towards-common-methods-for-assessing-open-data/

Table 2
Selected open data assessments: Focus, theoretical framework and unit of analysis

Indicator source	Focus	Theoretical framework	Unit of analysis
BROAD	Capacity	No	Government: under [Agency]
Common Assessment Framework	Harmonising existing assessments in terms of readiness, use and impact	No	Government: no distinction
UN Open Government Readiness Assessment	Assessment: Readiness	No	Government: Middle [Top-level] and Society [Values]
World Bank Open Data Readiness Assessment	Assessment: Readiness	No	Government: middle and under
Open Data Barometer	Ranking: readiness, use and impact	No	Government: no distinction
Ubaldi Open Government Data	Value	No	Government: no distinction

Towards Empirical Analysis of Open Government Data Initiatives.[25]

Generally speaking, the indicators proposed by this study correlated with those used in the assessments, diagnostic tools and indexes examined. This is perhaps not surprising given that the indicators proposed for this study were fewer than those used in the other evaluation instruments because of the focus on very specific aspects of open data practice in this study. However, the indicators proposed in this study are also distinct from those used in other instruments in two important aspects. First, the indicators in this study were theoretically informed. Second, the indicators in this study took into account the possible effects of two discreet organisational levels on the embedding of open data practice in government.

The open data evaluations that were open-data specific, focused on government open data, and provided sufficient documentation are summarised in the table below according to their focus, whether there is evidence of a theoretical framework underpinning the design of the assessment methods[26], and the assessment's unit of analysis (that is, whether it treats government as a single organisational entity or whether it differentiates between the under and middle structures).

Table 2 shows that the evaluations listed have been developed without reference to any theoretical framework that might provide a foundation for how the measurement or diagnostics can be used to explain how open data practice may be resisted or become embedded in governments. The table also reveals that very few assessments acknowledge the organisational dynamics of governments by including in the assessments the contribution of actors responsible for the day-to-day functioning of government departments and agencies. While the BROAD tool does acknowledge the distinction between the middle and under structures, its approach does not accommodate a single assessment of both levels, and therefore cannot capture any of the potential interplay between these organisational levels. The World Bank assessment is the only one that acknowledges both organisational levels, with a dedicated category of indicators for "Institutional structures, responsibilities and capabilities within government". However, this makes up only one of the eight indicator categories, and the category itself is weighted "Importance: High"; whereas the "Leadership" and "Demand" categories, which capture dynamics exogenous to the under structure, are weighted as "Very important". Furthermore, many of the indicators in the "Institutional" category are given a relatively low weighting; only one indicator is regarded as "Very important" and it has a technical capacity bias that is likely to be insensitive to either the normative or cognitive factors that affect behavioural change in the under structure.

25 Ubaldi B (2013) Open Government Data: Towards Empirical Analysis of Open Government Data Initiatives, OECD Working Papers on Public Governance, No. 22, OECD Publishing. http://dx.doi.org/10.1787/5k46bj4f03s7-en

26 It is acknowledged that notions of what constitutes a theoretical framework differ – between academia and other research organisations, and even between academic disciplines. It is also acknowledge that the theoretical underpinnings of a particular assessment method may not be made explicit in documents intended for consumption by practitioners. When conducting this brief analysis of existing open data practice evaluation methods, the focus was on locating evidence with the existing documentation available online, and of any attempt by the authors or organisation to link the proposed indicators to a particular explanation or theory of how the unit of analysis functions as a social system.

SAMPLE

The sample consisted of two national governments and selected organisational units from each.

South Africa and Kenya constituted the two countries in the national government sample. Their selection was based on the fact that South Africa and Kenya (along with Ghana) are often regarded as being the sub-Saharan countries that are furthest down the road in terms of open data initiatives.[27] Their selection is supported by the availability of existing research literature as well as documentation produced by the ODDC Phase 1 (no case studies on Ghana were included in the ODDC Phase 1 research).

Data for the under structures of each national government was collected from two government departments/ agencies: the national statistical agency and a public research agency. The selection of these two organisational sub-units is premised on two criteria:

1. that each sub-unit is associated with a different institutional domain, increasing the probability of different levels of alignment with the middle structure; and

2. that each sub-unit possesses a different level of autonomy in relation to the middle structure, thus potentially impacting on the extent to which it will conform with the pressures for change being exerted by the middle structure.

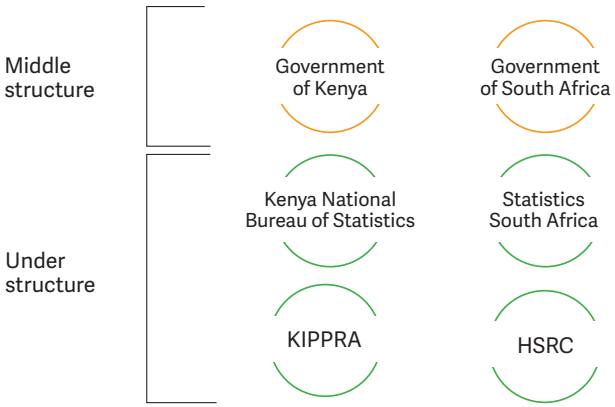
The composition of departments/agencies was as follows: Human Sciences Research Council (institution: science; country: South Africa); Statistics South Africa (institution: statistics; country: South Africa); Kenya Institute for Public Policy Research and Analysis (institution: science; country: Kenya); and Kenya National Bureau of Statistics (institution: statistics; country: Kenya).

The Human Sciences Research Council (HSRC) was established in 1968 as South Africa's statutory research council and is currently the largest social sciences and humanities research institute on the African continent. Its mandate is to inform effective formulation and monitoring of government policy, evaluate policy implementation, stimulate public debate through effective dissemination, foster research collaboration, and to help build research capacity and infrastructure for the human sciences.[28]

27 http://www.opendatabarometer.org/report/analysis/rankings.html

28 http://www.hsrc.ac.za/en/about/what-we-do

Figure 2
Sample by level of analysis



Statistics South Africa (StatsSA) is the national statistical service of South Africa, producing official demographic, economic, and social censuses and surveys. It is a national government department accountable to the Minister in the Presidency responsible for the National Planning Commission and the activities of the department are regulated by the Statistics Act (Act No. 6 of 1999), which ensures independence from political interference in production and dissemination methods and practices (Stats SA, 2013).

The Kenya Institute for Public Policy Research and Analysis (KIPPRA) was established as an autonomous public institute in May 1997. Its aim is to be a leading institute in public policy research and analysis by providing public policy advice to government and other stakeholders by conducting independent research and undertaking capacity building in order to contribute to achievement of national development goals. It also sees itself as a point of contact encouraging the exchange of views between the government, private sector and other civil society.[29]

The history of organised statistical activities in Kenya dates back to the 1920s, but was administered by various government departments until the Kenya National Bureau of Statistics (KNBS) was established under the Statistics Act of 2006. The Act established KNBS as a semi-autonomous government agency incorporated under the Ministry for Planning, National Development and Vision 2030, with the core mandate of collecting, compiling, analysing, publishing and disseminating statistical information for public use.[30]

29 http://www.kippra.org/About-KIPPRA/about-kippra.html

30 http://www.knbs.or.ke/index.php?option=com_content&view=article&id=79&Itemid=504

DATA COLLECTION

In the case of indicators M3 and U3, secondary data extracted from two open data indexes – the Open Data Barometer (ODB)[31] and the Global Open Data Index[32] – were used. In the case of the HSRC and KIPPRA, no relevant secondary data could be found in either index. In these cases, the research team identified a dataset produced by the department/agency that is produced regularly and/or is specific to the mandate of the department/agency, and the ODB and ODI methodologies were then replicated for each of these specific datasets by the research team.

In the case of Indicators M1, M2, U1 and U2, indicator data was collected by means of desk research focusing on the existing ODDC Phase 1 case study reports. Semi-structured interviews and extended email correspondence with key personnel in each of the governments also generated data. Email addresses of relevant, senior officials were sourced from the websites of the units included in the study. Formal requests to participate were sent to each official. In some cases these officials participated in the interview process; in other cases the research team was referred to other personnel within the unit. A set of questions informed by the indicators were developed to guide the interview and email correspondence processes between the research team and the respondents (see Appendix 1). From an analysis of the interview data collected for each unit, the research team scored each unit on each of the six indicators.

The section below provides detail on how each of the six indicators was quantified. In all cases, the method of assessment was based on the principle of a continuum of compliance rather than on an absolute dichotomous comply/resist basis. Scales were developed as a means of quantifying the compliance continua for the policy and licensing indicators (M1, M2, U1 and U2). In creating the scales, the number of categories was determined by examining the data collected from the desk research as well as from the interviews conducted. This was an iterative process of using the interview instrument to collect a first round of data, followed by email correspondence and additional interviews to fill in gaps in the data. This process was followed to ensure that the scales were informed by the realities on the ground. It also meant that the scales were not uniform in the sense that they were made up of an equal number of values; the number of values was determined by the extent to which each individual value made sense in relation to the secondary data rather than because they had to conform to an equal number across all four scales.

M1: Policy and regulation

In order to collect data on policy and regulations pertaining to open data at the middle structure level, a 7-point scale was developed to answer the question: "Does the government have an open data policy in place?" Terms of Use are not typically associated with either policy or legislation. However, following our grounded approach to developing the scales, it was felt that Terms of Use as they appear on a government open data platform or portal are both relevant and useful to include because these platforms and portals constitute spaces for the middle structure to make public statements on how open government data can be used, particularly if there are no specific policies or provisions in legislation on the matter.

0	1	2	3	4	5	6	7
No	No, but Terms of Use in place with indirect references applicable to open data	No, but other policies or legislation in place that make reference to openness more broadly	No, but Terms of Use in place with explicit references to open data, or principles of open data	No, but other policy or legislation that makes explicit reference to open data, or principles of open data, is under development	No, but other official and approved policies in place that make explicit reference to open data, or principles of open data	Yes, specific open data policy exists but policy in still in draft form	Yes, an approved open data policy exists

M2: Licensing

In order to collect data on open data licensing in the middle structure, a 10-point scale was developed to answer the question: "In a policy or similar government document, is it made explicit how the open data it publishes should be licensed?" In this case, the scale used the Creative Commons licensing framework as a reference and added additional categories to capture likely instances that are not captured by any single iteration of a Creative Commons licence.

0	1	2	3	4	5	6	7	8	9	10
No, there is not any statement in policy or elsewhere that directs how open data should be licensed	Yes, all data is under full copyright	Yes, but data may not be used for commercial gain, nor may it be adapted shared alike (CC-BY-NC-ND)	Yes, but data may not be used for commercial gain (CC-BY-NC)	Yes, but data may not be used for commercial gain (CC-BY-NC)	Yes, but data may not be adapted (CC-BY-ND)	Yes, but data must be shared using the same provisions (CC-BY-SA)	Yes, but data must be attributed (CC-BY)	Yes, but data must be attributed and a data-specific licence is applied (ODC-BY)	Yes, CC-0 'licence' or similar public domain custom licence	Yes, Public Domain for data/ databases (PDDL)

31 World Wide Web Foundation, Open Data Barometer Global Report (2nd Edition) 2015 http://www.opendatabarometer.org

32 http://index.okfn.org/

M3: Readiness

The Open Data Barometer's "Readiness" sub-index for "Government" was used as a proxy for the extent to which the middle structure of governments can be said to be implementing open data initiatives.

In carrying out a qualitative Open Data Readiness assessment across a number of countries from 2010 to 2013, the Web Foundation developed a six-dimensional framework for looking at the Political, Organisational, Legal, Social, Economic, and Technical context within a country in order to understand factors that may facilitate or inhibit the development of an open government data initiative, and the successful use of open data. These six dimensions informed the selection of indicators in the readiness section of the Open Data Barometer.

The variables used in the "Readiness" sub-index for "Government", along with their variable names, are as follows:

- ODB.2013.C.INIT (Expert survey question): To what extent is there a well-resourced open government data initiative in this country?
- ODB.2013.C.CITY (Expert survey question): To what extent are city or regional governments running their own open data initiatives?
- WEF.GITR.8.01 (Secondary data): Importance of ICT to government vision (World Economic Forum Global Information Technology Report 2014; Variable 8.01; Taken from WEF expert survey)
- UN.OSI (Secondary data): UN E-Government Survey, Government online services index (2014 edition)

Variable ODB.2013.C.INIT seeks to measure an explicit government commitment to open data:

> An open data initiative is a programme by the government to release government data online to the public. It has four main features:
> The government discloses data or information without request from citizens. This may be according to a release schedule or ad hoc.
>
> 1. The Internet is the primary means of disclosure. Mobile phone applications may also be used for disclosure.
> 2. Data is free to access and reuse, e.g. open licences;
> 3. Data is in a machine readable format to enable computer-based reuse, e.g. spreadsheet formats, Application Programming Interface (API).
>
> Look for all these features in the policy you are assessing for it to receive the maximum scores.[33]

The variable also seeks to reflect the extent to which significant resources are allocated by national government for an open government data initiative. These include a sufficient budget as well as personnel and facilities to carry out the mandate of the open data initiative, including technical personnel with appropriate qualifications for dealing with open data issues.

We exclude the ODB.2013.C.CITY variable. The ODB is interested in the national landscape as an enabling environment for open data practice and it therefore makes sense to include enabling conditions at the regional and metropolitan levels. However, we are interested in national government as a discreet organisation. Autonomous organisational arrangements outside of national government are therefore deemed to have no bearing on the readiness of the national government as a bounded environment for enabling or inhibiting open data practice.

The other two secondary datasets are also excluded because while they measure important enabling conditions such as ICT and e-government services, they do not speak directly to governments' *open data* commitments.

U1: Policy

In order to collect data on policies pertaining to open data at the under structure level, a 9-point scale was developed to answer the question: "Does the government department or agency have an open data policy in place?"

0	1	2	3	4	5	6	7	8	9
No	No, but data sharing policy addressing dissemination exists	No, but other policies in place (e.g. ICT policy) that make reference to openness more broadly (such as open source, public benefits of ICT, etc.)	No, but Terms of Use in place with references applicable to open data	No, but Terms of Use in place with explicit references to open data, or principles of open data	No, but other policies in place that make explicit reference to open data, or principles of open data	Yes, specific open data policy exists, but pending approval	Yes, specific open data policy exists but lacks clarity around quality of openness	Yes, specific open data policy exists but there are questions around nature/ authenticity of openness	Yes, an approved open data policy exists

33 Web Index Survey Handbook (Open Data Barometer extract) www.opendatabarometer.org/report/about/

U2: Licensing

In order to collect data on open data licensing in the under structure, a 10-point scale was developed to answer the question: "Does the government department or agency apply open data licences to the datasets it publishes?" As for indicator M2, the scale used the Creative Commons licensing framework as a reference and added additional categories to capture likely instances not captured by any single iteration of a Creative Commons licence.

0	1	2	3	4	5	6	7	8	9	10
No, there is no evidence of open licences being assigned to datasets	No, all data is subject to full copyright as specified in the Terms of Use or in a site-wide licence	No, all data is subject to copyright exceptions as specified in the Terms of Use or in a site-wide licence	Yes, CC or custom open licence but applied site-wide rather than attributed to specific data objects	Yes, but data may not be used for commercial gain or adapted	Yes, but data may not be used for commercial gain [attributed to specific data objects]	Yes, but data may not be adapted [attributed to specific data objects]	Yes, but data must be shared using the same provisions [attributed to specific data objects]	Yes, but data must only be attributed [attributed to specific data objects]	Yes, CC-0 'licence' or similar custom licence [attributed to specific data objects]	Public Domain licence [attributed to specific data objects]

U3: Implementation

In order to measure the level of open data implementation in the under structure, the openness of specific datasets that fall under the ambit of the relevant department or agency were assessed. KNBS and StatsSA are the government agencies responsible for national statistics.

We relied on secondary data from two sources to calculate the score for StatsSA and KNBS:

1. Open Data Barometer (ODB). In the 2014 Open Data Barometer, expert survey researchers were asked to complete a detailed checklist. The 10 checklist questions are shown in the tables in the Findings section of this report. Following validation, the ODB weights the checklist responses. The weighting is designed to emphasise the four questions that pick out key aspects of the Open Definition. Census Data falls within the mandates of both statistical agencies. We therefore use only the Census Data scores from the Open Data Barometer as an indicator of the extent to which these two government agencies are publishing open data. However, to avoid duplication of the same indicator, we removed the question related to licensing as this is already covered in indicator U2.

2. Global Open Data Index (ODI). The Global Open Data Index tracks whether data is released in a way that is accessible to citizens, media and civil society. It is a crowd-sourced survey of open-data releases around the world that is peer reviewed by a network of local open data experts. The Global Open Data Index measures the existence of 10 government datasets (national statistics, government budget, election results, legislation, national map, pollutant emissions, company register, transport timetables, post codes, and government spending) according to nine questions, the answers to which are weighted. One of these datasets is described as "National Statistics". KNBS and StatsSA are the government agencies responsible for national statistics in our study. Key national statistics include demographic and economic data (e.g. GDP, unemployment, population, etc.). Aggregate data (e.g. GDP for whole country at a quarterly level, or population at an annual level) is also considered acceptable in this data category. We therefore use only the National Statistic scores from the Global Open Data Index. As for the Open Data Barometer, questions related to licensing are excluded in our calculations.

In the case of KIPPRA and the HSRC, specific datasets taken to be the responsibility of the agency and published on a regular basis were identified and assessed using the same ODB and ODI methodologies described above.

LIMITATIONS

Developing indicators and representing these in diagrams that purport to represent reality is always fraught with danger. It is therefore important keep in mind that what we have attempted to represent or indicate are, and can only be, approximations and simplifications of the complexity of society and the behaviour of individuals. It is for this reason that, in addition to the quantitative measures used, the research team elected to complement the quantitative data with qualitative information gained from documents, reports and interviews.

Thornton and Ocasio[34] point to the limits placed on sufficiently in-depth published studies because of article word-count limits imposed by journal publishers. We would argue that the same applies to contract research, although in this case the constraint is not number of words but number of days. Commissioned research projects are time-bound and this places limits on the number of cases and on the number of interviews that can be conducted.

34 Thornton & Ocasio (2013).

Ideally, we would have included more government agencies in our sample and we would have been able to interview more of the leaders and employees of those agencies, spending more time embedding ourselves in the organisations being studied.

We acknowledge that when devising the indicators presented in this study, we did not differentiate between organisations (governments) of the developing and the developed world. We saw organisations as universal in their structures and functions. Without the inclusion of governments from developed countries in our sample, we cannot make any claims as to whether the indicators are sensitive to national contexts or whether they are only sensitive inter-organisational differences. To this extent, the inclusion of "Africa" in the title of this report is a red herring.

5 FINDINGS

In this section we present the findings from the quantitative analysis in the form of the indicators developed, as well as the findings from the interviews. We feel that there is value in sharing findings from the interviews as they complement and add a degree of richness that is absent in the quantitative findings.

The quantitative findings are presented separately for each of the country cases before findings from the interviews are shared.

SOUTH AFRICA

Figure 3 shows the findings for South Africa. Table 3 shows the indicator scores for each of the three countries, and for both the middle and under structure of the South Africa government as organisation. The figure and table are followed by detailed descriptions of how the scores were calculated for each of the indicators.

M1: Policy and regulations

The Constitution of the Republic of South Africa[35] (1996) provides the overarching foundation for access to information in South Africa. Clause 32 (1) of the Constitution states that: "Everyone has the right of access to (a) any information held by the state; and (b) any information that is held by another person and that is required for the exercise or protection of any rights." In line with the spirit of the Constitution, South Africa is a signatory to the Open Government Partnership[36], an international platform for governments committed to being more open, accountable and responsive to citizens.

In terms of national legislation, a number of Acts regulate citizens' rights around access to state information and protection of personal data, including:

- Copyright Act No. 98 of 1978[37]. This Act makes explicit which works are eligible for copyright in South Africa and regulates the copyright provisions pertaining to those works.

Table 3
South Africa: Score by indicator

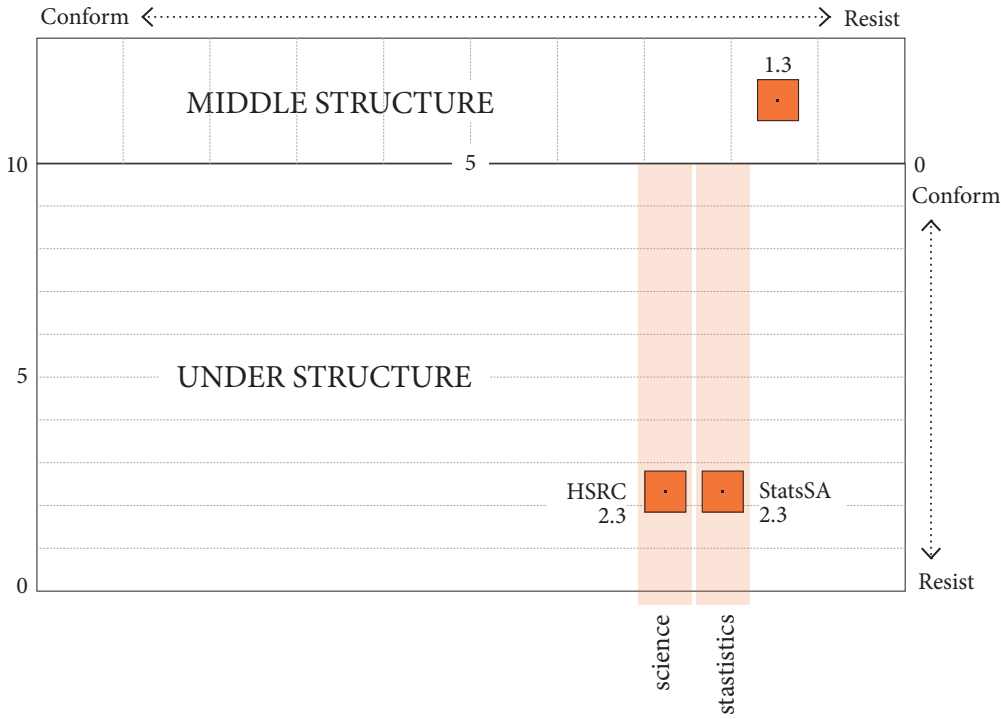
LEVEL	POLICY & REGULATIONS		LICENSING		OPENNESS		TOTAL	
Middle structure	M1		M2		M3		3.857/30	1.3/10
	2/7	2.857/10	0	0/10	1/10	1.000/10		
Under structure: StatsSA	U1		U2		U3		6.950/30	2.3/10
	1/9	1.111/10	1/10	1/10	75/155	4.839/10		
Under structure: HSRC	U1		U2		U3		6.950/30	2.3/10
	1/9	1.111/10	1/10	1/10	75/155	4.839/10		

35 http://www.thehda.co.za/uploads/images/unpan005172.pdf

36 http://www.opengovpartnership.org/

37 http://www.nlsa.ac.za/downloads/Copyright%20Act.pdf

Figure 3
South Africa: Extent to which open data practice is being embedded in the middle and under structures



- Promotion of Access to Information Act No. 2 of 2000 (PAIA)[38], under which organisations are required to grant access to records if those records are required by the requestor to exercise or protect their legal rights.
- Intellectual Property Rights (IPR) from Publicly Financed Research and Development Act No. 51 of 2008[39], intended to provide for more effective utilisation of intellectual property emanating from publicly financed research and development.
- Protection of Personal Information Act of 2013 (POPI)[40], under which "organisations in South Africa are required to be open and transparent about how they handle personal information, and allow individuals to access and correct their personal information which the organisations hold"[41].
- There is also a controversial piece of proposed legislation, the Protection of Information Bill[42] (commonly referred to as the Secrecy Bill), which sets out to replace the Protection of State Information Act of 1982. The Bill articulates a range of measures to protect classified information and while it is not yet enacted, it is relevant in the access to information legislative framework in that it demonstrates a counter-approach to the PAIA and POPI by foregrounding state secrecy over freedom of expression and the Constitutional right around access to information.

South Africa does not have a specific open data policy, with access of state data being determined by the legislative framework outlined above depending on contextual circumstances such as nature of data, intended application, and privacy factors.

In light of this overall context, South Africa receives a score of 2 on the "M1: Open data policy" scale, placing it in a category where no explicit open data policy exists, but other policies or legislation are in place that make reference to openness more broadly.

38 http://www.dfa.gov.za/department/accessinfo_act.pdf

39 http://research.ukzn.ac.za/Files/IPR_from_PFR_D_Act_No_51_of_2008_1.sflb.pdf

40 http://www.kpmg.com/ZA/en/IssuesAndInsights/ArticlesPublications/Protection-of-Personal-Information-Bill/Documents/POPI%20ACT.pdf

41 http://www.willis.co.za/privacy/

42 http://www.ssa.gov.za/Portals/0/SSA%20docs/Media%20Releases/2010/Summary%20of%20the%20Protection%20of%20Information%20Bill.pdf

M2: Licensing

The South African policy and legislative framework does not explicitly address licensing of open data released by state departments and government agencies.

Various pieces of legislation are, however, invoked by national research councils and agencies in their articulation of copyright frameworks. Relevant to the cases in this study, the HSRC references the IPR from Publicly Financed Research and Development Act of 2008 in their adoption of an "All rights reserved" copyright regime, based on the Act's provisions around data gathered using Parliamentary grant financing[43]. The Statistics South Africa Access to Information Manual[44], on the other hand, has been drafted in line with the framework dictated by the Promotion of Access to Information Act of 2000 (PAIA), and gives effect to the Constitutional right of access to information.

South Africa receives a score of zero on the "M2: Open data licences" scale, a category indicating a situation where there is no statement in policy or elsewhere that directs how open data should be licensed.

M3: Openness: Readiness

The Open Data Barometer 2014 score for South Africa's readiness to support and implement open data is 1.00 out of a possible 10.

U1: Policy – StatsSA

StatsSA has no formal open data policy, but has published an Access to Information Manual[45] in line with the requirements stipulated in section 14 of the Promotion of Access to Information Act 2 of 2000.[46] The purpose of this Manual is to give effect to the constitutional right of access to information, as outlined in section 32 of the South African Constitution, and to assist persons requesting access to Stats SA records.

The Stat SA Access to Information Manual makes no explicit mention of open data, nor does it make reference to IP rights and copyright management of data. Section 12 of the Manual states that "[a]s required by section 15 of the Information Act the following documents held by Stats SA are automatically available to the public and may be downloaded from the website or can be obtained from Stats SA's library",[47] but there is no list of documents provided. It is therefore unclear what information the public has automatic legal access to.

While the Stats SA Manual addresses data sharing, it does so in a context where Stats SA retains full copyright over data and request forms must be submitted in order to access data (in some cases accompanied by payment). In light of this situation, Stats SA receives a score of 1 on the "U1: Open data policy scale", placing it in a category where there is no explicit open data policy, but data-sharing guidelines addressing dissemination exists.

U1: Policy – HSRC

The HSRC does not have an explicit open data policy in place, but the overarching Human Sciences Research Council Act of 2008 recognises "the effective dissemination of fact-based results of research" as one of Council's core objectives.[48] Within this context, it specifically lists the objective to "develop and make publicly available new data sets to underpin research, policy development and public discussion".[49]

The HSRC has also articulated a Data Sharing Policy, which states that the HSRC "supports data sharing and therefore ensures that data is managed, organized and preserved for further reuse by the HSRC, as well as by the wider scientific and stakeholder community".[50] The policy further states that all activity in this regard is "in line with the requirements of the Intellectual Property Rights from Publicly Financed Research and Development Act (Act No. 51 of 2008) which requires that 'intellectual property emanating from publicly financed research and development is identified, protected, utilised and commercialized for the benefit of the people of the Republic'."[51]

While HSRC policy addresses data sharing, it does so in the context of an "All rights reserved" copyright framework, meaning that data is not shared in line with internationally-

43 HSRC Data Sharing Policy 2014. Unpublished document.

44 http://www.statssa.gov.za/wp-content/uploads/2013/07/Access_to_information.pdf

45 StatsSA (2003) Access to Information Manual. http://www.statssa.gov.za/wp-content/uploads/2013/07/Access_to_information.pdf

46 http://www.dfa.gov.za/department/accessinfo_act.pdf

47 StatsSA (2003: 22).

48 Government of South Africa (2008) Human Sciences Research Council Act No. 17 of 2008. Available at: http://www.hsrc.ac.za/uploads/pageContent/179/Download%20the%20HSRC%20Act,%20No.%2017%20of%202008.pdf

49 Ibid: 6.

50 HSRC Data Sharing Policy 2014.

51 Ibid: 1

accepted open data principles. In light of this situation, the HSRC receives a score of 1 on the "U1: Open data policy scale", placing it in a category where there is no explicit open data policy, but a data-sharing policy addressing dissemination exists.

U2: Licensing – StatsSA

Stats SA does not apply open data licences to the datasets it publishes. The Stats SA website[52] through which data delivery is administered carries no terms of use, but does have a "Copyright Statistics South Africa" statement as a footnote on the home page.

The Access to Information Manual states:

> 8.1 A requester must use Form C of Annexure B to the Regulations regarding the Promotion of Access to Information. [. . .] The request must be submitted to the information officer of STATS SA, together with any applicable request fee. (p.9)
>
> [. . .]
>
> 9. A requester is required to pay the applicable request and access fees. [. . .] STATS SA may withhold access to any record requested until the requester has paid the applicable fee. (p.11)

StatsSA datasets are available via third-party platforms and services such as DataFirst[53] and Google Public Data[54], but in instances such as these datasets are also shared under full copyright in line with the terms of use articulated by those platforms.[55] Stats SA is therefore not seen to be an open data practitioner.

In line with this situation, StatsSA receives a score of 1 on the "U2: Open data licences" scale, meaning it is in a category where the government agency in question does not apply open data licences to the datasets it publishes, and all data is subject to full copyright, as specified in the terms of use or in a site-wide licence.

52 http://www.statssa.gov.za/

53 https://www.datafirst.uct.ac.za/dataportal/index.php/catalog/central#_r=1435153324651&collection=&country=&dtype=&from=1960&page=1&ps=&sk=Stats+SA&sort_by=nation&sort_order=&to=2014&topic=&view=s&vk=

54 http://www.google.com/publicdata/directory?q=South+Africa#!dp=Statistics+South+Africa

55 See http://www.google.com/intl/en/policies/terms/ and http://www.datafirst.uct.ac.za/

U2: Licensing – HSRC

The HSRC does not apply open data licences to the datasets it publishes. Instead, all users are required to agree to the terms and conditions of use described in an End User License[56], through which the organisation retains full copyright.

The Terms and Conditions of the HSRC End User License for accessing data requires that the end users agree to 19 conditions. These conditions are summarised in the Access Conditions of each metadata record as follows:

- Data and documentation will not be duplicated, redistributed or sold without prior approval from the HSRC.
- Data will be used for statistical and scientific research purposes only and confidentiality of individuals/organisations in the data will be preserved.
- The HSRC will be informed of any publications resulting from work based on the data and documentation.
- The HSRC will be acknowledged as the data source, in line with HSRC citation specifications.
- Electronic copies of any publications resulting from HSRC data will be sent to the HSRC for archiving and bibliographic purposes.
- The data collector, HSRC and relevant funders bear no responsibility for use of data or for interpretations.
- Retrieval of the data signifies agreement to comply with the above-stated terms and conditions, and the user provides assurance that the user of data will conform with widely-accepted standards of practice and legal restrictions intended to protect confidentiality of respondents.

The Access Conditions make no explicit mention of commercial activity, but the statement that data may only be used for academic purposes implies that commercial activity is prohibited. The End User License is more explicit in this regard, as it requires that the end users agree to:

> 2.3. [. . .] use and/or make personal copies of any part of the Data Collections only for the purposes of not-for-profit research, teaching

56 http://curation.hsrc.ac.za/modules/DCaccess/pnincludes/EndUserLicense2011-12-22.pdf

and/or educational development. To obtain permission prior to utilising part and/or all of the Data Collections for commercial purposes by contacting the Data Service Provider in order to obtain an appropriate license from the rights holder in question and/or their permitted licensee if one is available.

With respect to intellectual property rights, the Human Sciences Research Council Act of 2008 states that "[t]he rights in respect of any invention or design or any works eligible for copyright protection by an employee of the Council in the course and scope of the employee's employment vests in the Council," but goes on to state that "[t]he Council may in its discretion, but subject to such conditions as the Minister may determine, assign or dispose of any of its rights" (p.14).

The HSRC receives a score of 1 on the "U2: Open data licences" scale, placing it in a category where the government agency in question does not apply open data licences to the datasets it publishes, and all data is subject to full copyright, as specified in the terms of use or in a site-wide licence.

U3: Openness: Implementation – StatsSA

The scores for the extent to which Census Data and National Statistics can be described as open according to the Open Data Barometer and the Global Open Data Index are provided in the tables below. The combined score for StatsSA is 75.00 out of a possible 155.00. Of note is the big discrepancy between the ODB and ODI scores.

StatsSA Open Data Barometer score

Q	Question	Weight	Score
a	Does the data exist?	5	0
b	Is it available online from government in any form?	10	5
c	Is the dataset provided in machine-readable formats?	15	10
d	Is the machine-readable data available in bulk?	15	0
e	Is the dataset available free of charge?	15	0
g	Is the dataset up to date?	10	0
h	Is the publication of the dataset sustainable?	5	0
i	Was it easy to find information about this dataset?	5	0
j	Are (linked) data URIs provided for key elements of the data?	5	0
	TOTAL	**85**	**15**

StatsSA Open Data Index score

Question	Answer	Weight	Score
Does the data exist?	Yes	5	5
Is data in digital form?	Yes	5	5
Publicly available?	Yes	5	5
Is the data available for free?	Yes	15	15
Is the data available online?	Yes	5	5
Is the data machine readable?	Yes	15	15
Available in bulk?	No	10	0
Is the data provided on a timely and up to date basis?	Yes	10	10
SCORE for 2014		**70**	**60**

U3: Openness: Implementation – HSRC

The dataset South African Social Attitudes Survey (SASAS)[57] was identified as a dataset likely to be published regularly and one that falls within the mandate of the agency. The scores for this dataset according to the Open Data Barometer and the Global Open Data Index are provided in the tables below. The combined score for the HSRC is 75.00 out of a possible 155.00. The ODB and ODI scores are consistent across both methods.

HSRC Open Data Barometer score based on own calculations

Q	Question	Weight	Score
a	Does the data exist?	5	5
b	Is it available online from government in any form?	10	10
c	Is the dataset provided in machine-readable formats?	15	15
d	Is the machine-readable data available in bulk?	15	0
e	Is the dataset available free of charge?	15	0
g	Is the dataset up to date?	10	0
h	Is the publication of the dataset sustainable?	5	5
i	Was it easy to find information about this dataset?	5	5
j	Are (linked) data URIs provided for key elements of the data?	5	0
	TOTAL	**85**	**40**

HSRC Open Data Index score based on own calculations

Question	Answer	Weight	Score
Does the data exist?	Yes	5	5
Is data in digital form?	Yes	5	5
Publicly available?	Yes	5	5
Is the data available for free?	No	15	0
Is the data available online?	Yes	5	5
Is the data machine readable?	Yes	15	15
Available in bulk?	No	10	0
Is the data provided on a timely and up to date basis?	No	10	0
SCORE		**70**	**35**

57 http://curation.hsrc.ac.za/Dataset-299-metadata.phtml

KENYA

Table 4 shows the indicator scores for each of the three countries, and for both the middle and under structure of the Kenyan government as organisation. Figure 4 shows the findings for Kenya. The figure and table are followed by detailed descriptions of how the scores were calculated for each of the indicators.

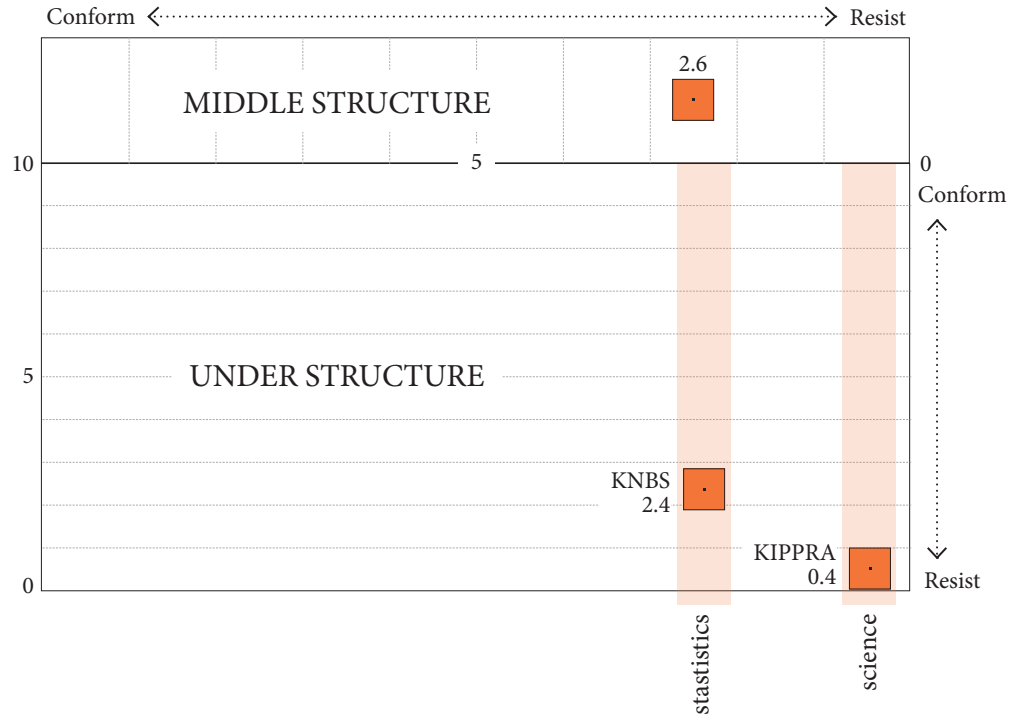
Constitution states: "Every citizen has the right to access information held by the state … The state shall publish and publicize any important information affecting the nation."

The Constitution of Kenya contradicts the Official Secrets Act (1970, revised 2009), which previously gave the

Table 4
Kenya: Score by indicator

LEVEL	POLICY & REGULATIONS		LICENSING		OPENNESS		TOTAL	
Middle structure	M1		M2		M3		7.714/30	2.6/10
	4/7	5.714/10	0	0/10	2/10	2.000/10		
Under structure: KNBS	U1		U2		U3		7.272/30	2.4/10
	1/9	1.111/10	1/10	1/10	80/155	5.161/10		
Under structure: KIPPRA	U1		U2		U3		1.290/30	0.4/10
	0/9	010	0/10	0/10	20/155	1.290/10		

Figure 4
Kenya: Extent to which open data practice is being embedded in the middle and under structures



M1: Policy and regulations

Access to information in Kenya is regulated by the Constitution of Kenya (2010)[58], which requires that public institutions allow access to information. Article 35 of the

government the right to withhold data from the public. Recent changes in government approach towards information-sharing have signalled a new era of transparency, as is reflected in the Freedom of Information

58 http://www.wipo.int/wipolex/en/details.jsp?id=8559

Bill (2012)[59] and the Access to Information Bill (2012)[60], both of which aim to give effect to Article 35 of the Constitution. Majeed points out that under the regime of previous president Daniel Arap Moi (1978–2002), "the government was clearly on one side and the citizen on the other … citizens had no business accessing [government] information."[61]

While the Kenya Constitution recognizes an individual's right to information, the Kenyan legislature still needs to develop a framework to implement and codify this constitutional right.[62] It is expected that the proposed Freedom of Information Act will address this situation. There appears to be a good deal of support for the Act, which will hopefully "define what information citizens can access, how public institutions can avail data to the public, or the consequences to be faced for their violation or failure to comply".[63]

In addition to the Constitution as well as the Freedom of Information and Access to Information Bills, dissemination of government information is also regulated by the Kenya Copyright Act (2001) in that it governs the use and distribution of government content by third parties.

In light of this overall context, Kenya receives a score of 4 on the "M1: Open data policy" scale, placing it in a category where no explicit open data policy exists, but other policy which makes explicit reference to open data, or principles of open data, is under development.

M2: Licensing

There is no national policy or legislation that addresses licensing of government open data in Kenya, although it is envisioned that the proposed Freedom of Information Act will address IP frameworks and legal conditions around accessing state information and government data, which may inform licensing approach.[64]

Kenya receives a score of zero on the "M2: Open data licences" scale, a category indicating a situation where there is no statement in policy or elsewhere that directs how open data should be licensed.

M3: Openness: Readiness

The Open Data Barometer 2014 score for Kenya's readiness to support and implement open data is 2 out of a possible 10.

U1: Policy – KNBS

KNBS does not have an explicit open data policy, but has published a Data Dissemination and Access Policy (2012).[65] This policy was formulated as a direct response to the national Kenya Open Data initiative, and section 3 (p.2) of the policy, pertaining to "Objectives", states:

> The broad objective of this dissemination policy is to ensure timely and quality data provision to data users to be achieved through:
>
> **a.** Provision of a framework for availing data to the public in conformity with the government's open data initiative

The KNBS Data Dissemination and Access Policy outlines the dissemination principles for official statistics, distribution mechanisms and features pertaining to data access. While the policy demonstrates an overt in-principle commitment to "wide and easy access to official statistical data" as a standing priority (p.5), this is done in the context of an "All rights reserved" framework which prohibits both adaptation and commercial application of datasets. Data cannot therefore be considered authentically open in line with international protocols.

In light of this situation, KNBS receives a score of 1 on the "U2: Open data licences" scale, placing it in a category where the government agency in question does not apply open data licences to the datasets it publishes, and all data is subject to full copyright, as specified in the terms of use or in a site-wide licence.

U1: Policy – KIPPRA

The foundational piece of legislation governing KIPPRA

59 http://www.law-democracy.org/wp-content/uploads/2010/07/Kenya.FOILawJan12.pdf

60 http://www.cickenya.org/index.php/legislation/item/333-the-access-to-information-bill-2013#.VXfzimSqqko

61 Majeed R (2012) Disseminating the Power of Information: Kenya Open Data Initiative, 2011–2012. http://www.princeton.edu/successfulsocieties/content/superfocusareas/traps/ME/policynotes/view.xml?id=206

62 Brown G (2013) Why Kenya's open data portal is failing – and why it can still succeed. http://sunlightfoundation.com/blog/2013/09/23/why-kenyas-open-data-portal-is-failing-and-why-it-can-still-succeed/

63 Mutuku L & Mahihu CM (2014) Open Data in Developing Countries: Understanding the Impacts of Kenya Open Data Applications and Services. p. 26. Available at: http://www.opendataresearch.org/content/2014/731/understanding-impacts-kenya-open-data-applications-and-services

64 J Bhalla, pers. comm., 2015.

65 http://knbs.or.ke/index.php?option=com_phocadownload&view=category&download=372:knbs-data-access-and-data-dissemination-policy&id=23:policies-manuals&Itemid=599

activities is the Kenya Institute for Public Policy Research and Analysis Act (2006, revised 2012)[66], which provides for the establishment of the think tank and outlines provisions on incorporation, powers and functions. This Act does not, however, make mention of dissemination, copyright management or terms of use around accessing data. KIPPRA itself has no open data policy and, while some datasets are provided via the "Resource Centre"[67] link on its website, it cannot be said to be an open data practitioner.

KNBS receives a score of 1 on the "U2: Open data licences" scale, placing it in a category where the government agency in question does not apply open data licences to the datasets it publishes, and all data is subject to full copyright, as specified in the terms of use or in a site-wide licence.

U2: Licensing – KNBS

KNBS does not apply open data licences to the datasets it publishes. Section 9.4 of the Data Dissemination and Access Policy, pertaining to "Terms and Conditions Governing Use of Public Data" (p.6) states that:

> a. Data and other materials provided by KNBS shall not be sold to other individuals, or organizations without written authority from the Director General.
>
> b. Data shall be used for statistical purposes only and not for investigation of specific individuals or organizations or any other purpose.
>
> [. . .]
>
> e. Requests for micro-data shall be serviced upon completion of and submission of the Application Form for Microdata.

Section 9.5, pertaining to "Copyright and Citation Requirements" states that "data users shall acknowledge that any available intellectual property rights, including copyright in the data are owned by the KNBS" (p.7). Added to this, the KNBS website Terms of Use[68], state that:

> 1. Everything you see, read or hear on this site is protected by copyright, design right and/or other intellectual property rights unless otherwise stated and may only be used in accordance with these Terms.
>
> [. . .]
>
> 3. You may download, print and store selected portions of the content of the site provided that you (I) only use these copies for your own personal, non-commercial use, (II) do not copy or post the content on any network computer or broadcast the content in the media, and (III) do not modify or alter the content in any way.

KNBS receives a score of 1 on the "U2: Open data licences" scale, meaning it is in a category where the government agency in question does not apply open data licences to the datasets it publishes, and all data is subject to full copyright, as specified in the terms of use or in a site-wide licence.

U2: Licensing – KIPPRA

There are no signs of open data licensing being applied to KIPPRA datasets, nor does there appear to be any form of metadata making access conditions explicit. The KIPPRA website, through which a very small selection of datasets are shared, carries a footnote stating that all content is copyright KIPPRA 2015 and that all rights are reserved. There is otherwise no documentation or statement around copyright.

KIPPRA receives a score of zero on the "U2: Open data licences" scale, meaning it is in a category where the government agency in question shows no evidence of open licences being assigned to datasets.

U3: Openness: Implementation – KNBS

The scores for the extent to which Census Data and National Statistics can be described as open according to the Open Data Barometer and the Global Open Data Index are provided in the tables below. The combined score for KNBS is 80.00 out of a possible 155.00. As in the case of StatsSA, of note is the big discrepancy between the ODB and ODI scores, although in the case of KNBS, ODB returns the higher score while in the case of StatsSA it was ODI that returned the higher score.

66 https://www.google.co.za/url?sa=t&rct=j&q=&esrc=s&source=web&cd=4&cad=rja&uact=8&ved=0CC4QFjAD&url=http%3A%2F%2Facbf-pact.org%2Fprojects%2Feastern-southern-africa%2Fkenya%2Fkenya-institute-public-policy-research-and-analysis&ei=3MqKVcnlHcTV7gaSw4_oDA&usg=AFQjCNGc8aoYTaCwz9D77Oeo-7JTE10WfA&sig2=v97M0yEdVsP-YkUZWrKZQQ&bvm=bv.96440147,d.ZGU

67 http://www.kippra.org/Resources/datasets.html

68 http://www.knbs.or.ke/index.php?option=com_content&view=article&id=193&Itemid=1070

KNBS Open Data Barometer score

Q	Question	Weight	Score
a	Does the data exist?	5	5
b	Is it available online from government in any form?	10	10
c	Is the dataset provided in machine-readable formats?	15	15
d	Is the machine-readable data available in bulk?	15	15
e	Is the dataset available free of charge?	15	15
g	Is the dataset up to date?	10	10
h	Is the publication of the dataset sustainable?	5	0
i	Was it easy to find information about this dataset?	5	0
j	Are (linked) data URIs provided for key elements of the data?	5	0
	SCORE for 2014	**85**	**70**

KNBS Open Data Index score

Question	Answer	Weight	Score
Does the data exist?	Yes	5	5
Is data in digital form?	Yes	5	5
Publicly available?	No	5	0
Is the data available for free?	No	15	0
Is the data available online?	Unsure	5	0
Is the data machine readable?	Unsure	15	0
Available in bulk?	Unsure	10	0
Is the data provided on a timely and up to date basis?	Unsure	10	0
SCORE for 2014		**70**	**10**

U3: Openness: Implementation – KIPPRA

The dataset Policy Timeline and Time Series Data for Kenya[69] was identified as a dataset likely to be published regularly and one that falls within the mandate of the agency. The scores for this dataset according to the Open Data Barometer and the Global Open Data Index are provided in the tables below. The combined score for KIPPRA is 20.00 out of a possible 155.00. The ODB and ODI scores are consistent across both methods.

KIPPRA Open Data Barometer score based on own calculations

Q	Question	Weight	Score
a	Does the data exist?	5	5
b	Is it available online from government in any form?	10	0
c	Is the dataset provided in machine-readable formats?	15	0
d	Is the machine-readable data available in bulk?	15	0
e	Is the dataset available free of charge?	15	0
g	Is the dataset up to date?	10	0
h	Is the publication of the dataset sustainable?	5	0
i	Was it easy to find information about this dataset?	5	5
j	Are (linked) data URIs provided for key elements of the data?	5	0
	TOTAL	**85**	**10**

69 http://www.kippra.org/index.php?option=com_docman&task=cat_view&gid=13&Itemid=

KIPPRA Open Data Index score based on own calculations

Question	Answer	Weight	Score
Does the data exist?	Yes	5	5
Is data in digital form?	No	5	0
Publicly available?	Yes	5	5
Is the data available for free?	Unsure	15	0
Is the data available online?	Unsure	5	0
Is the data machine readable?	No	15	0
Available in bulk?	Unsure	10	0
Is the data provided on a timely and up to date basis?	No	10	0
SCORE		**70**	**10**

GENERAL FINDINGS

One of the research questions this study posed was whether it could develop indicators on open data practice in government as an organisation. The findings presented show that it was possible to develop such indicators and that the indicators have revealed differences at various points of comparison between the two cases of the governments of South Africa and Kenya.

However, operationalising the indicators was not without its challenges. In particular, the development of fair and representative categories for licensing and policy/regulation scales were challenging. The differentiation between policy, legislation and terms of use was itself challenging at times. As was the relative position of no licence versus full copyright as an indicator of government and its agencies embedding open data practice. For example, should a clear and unambiguous statement about how data may not be shared (i.e. full copyright) score more highly than the absence of any statement or guidelines?

Our resolution of such concerns was to seek recourse to the conceptual framework: which scenario indicated a greater likelihood of the middle or under structure dealing with the complexities of data sharing. Following this approach, the presence of an explicit copyright statement shows evidence of some engagement, even if it is not towards openness.

The discussions on absence of clear open licensing terms versus full copyright surfaced an additional insight that resonated with the data that emerged from interviews. Government employees, often at the coal-face in terms of actioning data sharing policies, expressed confusion due to the absence of any guidelines on data licensing. It was clear from the interviews that this absences creates confusion

and/or moments for (mis)interpretation by actors in the under structure.

While our investigation into the open data initiatives, particularly in the case of Kenya, revealed strategic engagement with licensing issues, it also revealed that a patchwork of licensing systems was being applied, with a mix of both standard and bespoke licensing, and some vacillation in how licensing is being expressed.[70] This has arisen partly because of complex organisational realities, but also (in the case of the Kenya Open Data Initiative [KODI]) because of the absence of any clear policy or guidelines to regulate licensing activity. In the Kenya context this resulted in a disjunctive licensing approach between the data provider (KNBS) and the data publisher (KOD), with KOD administrators and technical team members being placed in the position of having to ascribe licensing provisions for third-party users when the provisions of data-provider agencies were not clear or conflicted with open data licensing practice.

70 See Willmers M, Van Schalkwyk F & Schonwetter T (2015) Licensing Open Data in Developing Countries: The Case of the Kenyan and City of Cape Town Open Data Initiatives. *The African Journal of Information and Communication* Thematic Issue: 'African Intersections between IP Rights and Knowledge Access'.

6 DISCUSSION

Kenya and South Africa reveal very different pictures. In the case of South Africa, the middle structure appears to be buffering itself from external pressures to embed open data practice in government. Relatively speaking, the two agencies studied, however, appear to be more responsive to such pressures and there are indications of open data practice taking root. However, in the case of both agencies, there are clearly still some barriers to full acceptance of open data practice (particularly in terms of the licensing component), even if relative to the government executive, acceptance is greater.

Nevertheless, despite the middle structure's buffering position, the units in the South African government under structure are conforming to a greater degree to external pressures. This confirms the fact that the under structure not only responds to institutional pressures, but also to pressures from the environment.

Based on the interviews with the HSRC, there is evidence that government agencies could be further down the track in terms of open data practice, were it not for the fact that the middle structure is holding back such efforts. It seems evident that the both the legislative and policy frameworks in the country are rooted in the commercial value of data as a government resource. In other words, the focus is on innovation and not on the transparency and engagement of open data as an important dimension of open government. This could be interpreted as being attributable to the marketisation of government, in so doing introducing commercialisation as the ideal type in terms of the institutional logic of the state. However, the institutional logics of the agencies in question – both with a strong disposition towards an ideal type of sharing as a driver of science advancement (for example, through the ability to replicate scientific discoveries and to interrogate more closely scientific claims) – is at loggerheads with the middle structure.

In the case of Kenya, the picture is diametrically opposite to that of South Africa in the sense that it is at the middle structure level that open data practice is more advanced. This indicates that the government executive is responding to external pressures by conforming rather than buffering itself from pressures to embed open practice.

At the level of the under structure, the units in the Kenyan sample are not responding uniformly, as was the case in South Africa. In the case of KIPPRA, there is little evidence of it conforming to pressures – either from the middle structure or directly from external stakeholders – to embed open data practices. Whether this is because of multiple and possibly conflicting institutional pressures is difficult to say based on the data available.

Findings from a larger sample of agencies in each government is needed in order to make more conclusive statements about the possible effects of multiple institutional pressures.

What is clear, however, is the extent to which KNBS is in greater alignment with the Kenyan government executive, than is KIPPRA. Neither of the South African agencies are in alignment with its government's executive. A possible explanation for the alignment between KNBS and the Kenyan government middle structure is that KNBS and the government's open data initiative (KODI) share data. To some extent there appears to be a pact or some form of mandate between the executive and KNBS, and therefore between the middle and under structures, in Kenya. In

South Africa, there is no evidence of the government executive engaging with agencies in the under structure around open data practice.

An analysis of the breakdown of the indicators shows that KNBS scores relatively well on the Implementation indicators, but poorly on the policy and licensing indicators. It is also this aspect that differentiates KNBS from KIPPRA. This could be interpreted as evidence of open data practice taking root despite clear policy and licensing guidelines being in place. The Kenyan executive in contrast scores relatively well in terms of the policy and legislative dimension, but poorly in terms of the other indicators. In other words, while there is alignment at the aggregated level, there is poor alignment in terms of strategies and priorities when it comes to open data practice between the middle and under structures in the Kenyan government.

It appears that the patchwork licensing scenario witnessed is largely due to the fact that data providers are still making sense of (1) the kinds of protection they require, (2) how licensing systems actually work and (3) the complexities of weaving together different stakeholder demands, from both within and from outside of the institution.

At the super structure level this seems to point to the fact that national governments of developing countries are the recipients of multiple 'organizing visions' on good government, open data and development from international donors, NGOs and other countries considered pioneering and exemplary. The licensing patchwork can be read as variations in interpreting the alleged strategic benefits seen as desirable development and implementation tactics. At this level, it is therefore not simply a matter of a national government conforming to or resisting a universally accepted dogma on open data, good government and development. In fact the issue of open data and its utilisation for socio-economic development by NGOs and private sector actors is full of controversies. Rather than complying or resisting 'best practice', one could conceptualise the licensing patchwork specifically or open data policy more generally as context-specific adaptations.

These context-specific adaptations can fruitfully be understood in the conceptual terms of mindful innovation. Westney's[71] study of the building of the modern Japanese state, which consciously drew and chose appropriate for its circumstances modern state institutions from European countries, provides similar evidence. Although Westney's book predates the concept of mindful innovation and therefore does not use this term, it nevertheless provides an account of a country that was influenced by the dominant norms on infrastructures for a modern state but proactively sought to not only imitate but also choose and adapt by introducing design elements suitable to its context.

The conceptual framework presented in this report is analogous with the multi-level approach that is characteristic of the institutional logics school of thought within institutional theory. However, even though the conceptual framework presented here acknowledges the existence and importance of organisational relevance of the under structure, it does not sufficiently surface the potential agency of the actors who occupy the under structure. And it would be short-sighted to ignore the agency of individual actors in institutional settings.

It is easy to assume that change stems from the environments of organisations and that public agencies' practices may change because of pressures from their environment. From this perspective the adoption of open data practice is a matter of conforming to external pressures, and is therefore precarious because, according to the institutional theory predicated on isomorphism as the binding logic of institutional actors, such pressures are expected to be resisted. The under structure of public agencies is, in this sense, understood to be stable and contented in the taken-for-granted beliefs that underpin the behaviour of its actors, and change may result from the isomorphic institutional pressures of mimicry and coercion (regulation). These are foundational ideas in institutional theory. And yet, the dynamic and proactive role of the public agencies undergoing change should not be neglected. Several theoretical developments draw attention to the agency of and within public sector institutions:

a) The perspective of 'institutional logics'[72] that proposes multiple institutionalised norms. Change occurs at the intersection of external as well as internal established beliefs, structures and practices.

b) The notion of 'mindful innovation' in conjunction to that of 'organising vision'.[73] Taken together,

71 Westney DE (1987) *Imitation and Innovation: The Transfer of Western Organizational Patterns to Meiji Japan*. Harvard University Press.

72 Thornton & Ocasio 2013

73 Swanson EB & Ramiller NC (1997) The Organizing Vision in Information Systems Innovation. *Organization Science 8*(5): 458-474; Swanson EB & Ramiller NC (2004) Innovating Mindfully with Information Technology. *MIS Quarterly 26*(4): 553-583.

these concepts suggest that isomorphic forces, while very important, do not have an unquestioned effect on organisations. Organisational actors do not necessarily adopt or resist in a 'mindless' way organising visions regarding specific new technologies and their alleged effects which are formed in their environment, but they consider, explicitly or tacitly, their suitability in their particular circumstances.

c) The notion of institutional 'entrepreneur' which proposes a view of institutional change by 'embedded agency'.[74]

Overall, these relatively recent contributions to the institutional theory of organisations suggest that those located in the under structures are not passive receivers of pressures for change from their environments. Their members are capable of agency, which is manifested in their capacity to take initiatives vis-à-vis developments in their environment. In other words, organisational actors in the under structure may be similar but they are not the same; similar to the extent that they exhibit conformity and compliance to institutional rules, norms and values, but different to the extent that each actor has the capacity to interpret differently those moments where intra-institutional rules, norms or values conflict, or where new impulses from the environment conflict with institutional rules, norms or values. For example, Avgerou's[75] research on the Mexican state corporation PEMEX, was able to identify internal dynamics of change in the interaction of organisational actors conveying alternative rationalised ways of organising with ICT.

In our findings we see evidence of actors in the under structure grappling with the complexities of laws and licensing as they pertain to open data practice. Instead of providing clear paths for compliance with exogenous prescriptions, they provide moments for sense-making and interpretation, and, as such, rely more on endogenous individual or collective normative and cognitive coercive elements for their effects. The result is a non-cohesive and unaligned organisation, and this could be interpreted as being indicative of de-institutionalisation of the organisation as conformity and adherence to systematised behaviour are undermined.

Government organisations are not as uniformly adhering to some historically established public administration norms or regulations. The public sector has been subject to sustained efforts of 'new public management', 'digital era government', etc., and has been exposed to, often by working side-by-side with the private sector, various new ideas and practices which are seen as legitimate.

Public agencies may therefore take a critical 'mindful' stance towards imposed policy and work out locally appropriate practices. It is therefore equally important to develop an understanding of government organisations as involving interactions of multiple institutionalised logics, stemming from their own members. The desirable effects expected from open data is not merely a matter of non-resisting, or on a top-down imposed policy of data openness. Whether open data policy leads to desirable outcomes regarding accountability, better government and data-driven innovation by non-government actors depends on the capacity of public agencies to transform themselves to a techno-organisational and knowledge-economy setting that enables both their employees and the public at large to utilise data for various objectives. This goes far beyond non-resisting openness. It requires enacting change through their own agency.

74 Greenwood R & Suddaby R (2006) Institutional Entrepreneurship in Mature Fields. *Academy of Management Journal* 49(1): 27-48; Hardy C & Maguire S (2008) Institutional Entrepreneurship. In R Greenwood, C Oliver, R Suddaby & A Sahlin-Andersson (Eds), *The Sage Handbook of Organizational Institutionalism*: Sage.

75 Avgerou C (2002) *Information Systems and Global Diversity*. Oxford: Oxford University Press.

7 IMPLICATIONS FOR POLICY AND RECOMMENDATIONS

POLICY IMPLICATIONS

For international government open data initiatives

The governments of Kenya and South Africa are both members of the Open Government Partnership (OGP), indicative on the surface of an executive-level commitment to openness (including open data) by both governments. However, in the case of the South Africa, the findings presented in this report indicate a low levels of open data practice. And in the case of Kenya, the findings point to higher levels of embeddedness but a poor degree of alignment between the executive and government departments. For OGP and other initiatives such as the International Open Data Charter, the research highlights three important considerations to take into account when formulating plans to improve the uptake and impact of open data in government:

1. It should not be presumed that executive-level commitments are tantamount to an acceptance of the tenets of openness in general and of open data in particular in government departments and agencies, and that open data practice will automatically take root throughout government as a result of such commitments.

2. It may be more effective to focus on departments and agencies when it comes to embedding open data practice in government. However, it should not be assumed that all departments or agencies will respond similarly to policies requiring them to publish open data. Multiple institutional logics are likely to be in force. Therefore, even if there is a shift in the institutional logic of the bureaucratic state, the institutional logic of the professions may still idealise a more closed stance.

3. While government departments and agencies are institutionally bound and, as such, more likely to resist change and to conform to existing practices underpinned by taken-for-granted norms and values, the possibility of agency makes possible the de-institutionalisation of existing practice and, therefore, by implication, the institutionalising of new practice. Advocates for change in the form of greater openness could therefore benefit from identifying and engaging with those actors in institutional settings most likely to champion openness. Future research may therefore consider the addition of indicators of the de-institutionalisation of traditional public administration regimes that considered data to be the 'closed' property of government. In other words, new research could build on and enhance the search for capacity for local action and locally meaningful innovation within the institutionalist theoretical framing.

For funders and donor agencies

Funders of open data initiatives and advocates of open data practice should consider the efficacy of their actions in instilling organisational change of the kind where open practice is regarded as the right or obvious course of action for government departments and agencies to follow. Our research has shown that it is important to be mindful of all levels of government when designing change initiatives, and that any initial success may result in unevenness between government agencies, and between government agencies and the executive.

For governments

Governments should be cautious when relying on open data assessments and rankings as indicators of success. In developing country contexts, assessments and rankings may capture and reflect high-level commitments to open data and fledging initiatives that result in a batch of open datasets being published, but they are unlikely to reflect the fundamental organisational dynamics and complexities that may hinder, or at least delay, long-term, sustainable – in other words, embedded – open data practice. In developing countries, some governments are frequent visitors to the top-end of global rankings and assessments. Again, these positions, while undoubtedly correlated to some positive activity in relation to open data practice, cannot be presumed to reflect the kind of institutional change that equates to open data practice being taken-for-granted universally.

RECOMMENDATIONS

1. International agencies should provide technical support to governments as they develop open data policies, including directives on the licensing of open government data. The working groups for the Open Government Partnership and the International Open Data Charter have an important role to play in this regard, particularly in developing countries.
2. While policy variation between governments may be inevitable given different local contexts, open data licensing within each government should be clear and unambiguous if use of open government data is to be facilitated.
3. Advocacy around the opening of data should focus on the dissemination of knowledge about the open of data, and its benefits, in an attempt to engage all institutional actors at a cognitive level. This is a different approach to advocacy efforts that target top-level government officials, seek their endorsement of open government data practice and then expect such practice to filter through government on the back of policy.
4. Target all levels of government, and consider bottom-up approaches or multiple-level approaches when designing advocacy campaigns or support interventions. Consideration should also be given to developing differentiated approaches for different government departments and agencies. For example, the strategy for advocating open data practice in the judiciary may be very different from a strategy to do the same in the national statistical agency.
5. Identify committed and interested actors within institutional contexts, with sufficient status, to promote the de-institutionalisation of long-held values which may be entrenching practices that hinder openness.
6. Employ the measurement instrument presented in this report for self-assessment of open data practice in government. Other assessments may benefit from incorporating elements from the measurement instrument presented here into their instruments in order to make their measurements more sensitive to the institutional dynamics at play in governments.

8 CONCLUSION

In the introduction to this report it was stated that insufficient attention has been paid to researching the change process surrounding open data, in particular the constellation of institutional domains in which government as a complex organisation functions. To address this insufficiency, we asked the following questions: Is open data practice being embedded in African governments? What are the possible indicators of open data practice being embedded? What do the indicators reveal about resistance to or compliance with pressures to adopt open data practice? And, what are different effects of multiple institutional domains that may be at play in government as an organisation?

This study has shown that it is possible to develop a set of indicators that assesses the extent to which open data practice is being embedded in government, and that these indicators, because they are grounded in a particular conception of how organisations function, provide evidence of the effects of multiple institutional domains in determining differing levels of open data practice in the departments and agencies of governments.

The findings show that while there are differences between the governments in the sample and between the agencies studied, open data practice is not being embedded to the extent that some may expect or wish to be the case. At both organisational levels examined in this study, where policies or supportive regulations are in place, the directives and application of open data licences is either nonexistent, erratic or open to interpretation. And where there is evidence of open data readiness and implementation, this occurs in a context where licensing is vague, conflicting or inconsistent. Both findings indicate that both organisational levels of government have yet to accept and come to terms with the consequences of data being re-used without restriction.

Some may question the importance of licensing as an indicator of open data practice being embedded in government; and even more so if there is evidence of open data policy and implementation. We maintain that retaining licensing as an indicator, in conjunction with the other two indicators, remains critical. This is not only because licensing provides clear directives and signposts for (re)users of open data; retaining licensing as an indicator is critical because licensing captures, at a cognitive level, compliance with a non-default and fundamental tenet of open data: unrestricted re-use. Moreover, it is at the cognitive level that multiple institutional logics are interpreted by organisational actors. Licensing captures the moments of sense-making that organisational actors face when pressures to publish data in open formats is brought to bear on them. And unless actors in both the government executive as well as in government departments and agencies are able to resolve mutually reinforcing, locally appropriate logics of action, expressed in unambiguous licensing that promotes the responsible re-use of data without constraint, then open data practice will continue to be piecemeal and symbolic.

APPENDIX 1
QUESTIONNAIRE

Guiding questions for semi-structured interviews:

1. How was your organisation's open data policy process initiated?
    a. Who introduced the idea of developing an open data policy?
    b. Why did they initiate this process?
    c. When was the idea of developing an open data policy first suggested?
    d. Was specific reference made at any point to other organisation's open data initiatives?
    e. Following the initiation of the policy process, has any one person in the organisation emerged as an "open data champion"?
2. Describe the process from initiation to final approval of the organisation's open data policy.
    a. What were the stages in the process?
    b. What was the level of resistance and what were the main concerns raised?
    c. What benefits were advanced to develop an open data policy?
3. Were there any specific discussions on the type of licensing to assign to open data published?
    a. What were the issues raised around licensing?
    b. Who were the main actors in discussions on licensing of the organisation's open data?
4. Was any consideration given to resources required in implementing the open data policy?
    a. If resources such as human resources and financial resources were discussed, how did the organisation deal with the allocation of resources?
    b. What resources were made available or what commitments were made in terms of future resources to support the open data policy?

ABOUT THE AUTHORS

François van Schalkwyk
PhD candidate, Stellenbosch University; Researcher, Centre for Higher Education Transformation, Cape Town; and Research Manager: Africa, World Wide Web Foundation

Michelle Willmers
Project Manager and Researcher, Open Data Africa Initiative, Intellectual Property Unit, Faculty of Law, University of Cape Town.

Tobias Schonwetter
Director, Intellectual Property Unit, Faculty of Law, University of Cape Town

ODDC Exploring the Emerging Impacts of Open Data in Developing Countries
WORLD WIDE WEB FOUNDATION
od OpenData for development
IDRC CRDI