

https://dspace.org/

Licensing Open Data in Developing Countries: The Case of the Kenyan and City of Cape Town Open Data Initiatives

Willmers, Michelle; Van Schalkwyk, François; Schonwetter, Tobias

2015-12-15

http://hdl.handle.net/10539/19312

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 (AJIC), 16, 26-37. https://doi.org/10.23962/10539/19312

Downloaded from a DSpace repository. https://dspace.org/

LICENSING OPEN DATA IN DEVELOPING COUNTRIES: THE CASE OF THE KENYAN AND CITY OF CAPE TOWN OPEN DATA INITIATIVES

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

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

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

ABSTRACT
Open data practice is gaining momentum in the public sector and civil society as an important mechanism for sharing information, aiding transparency, and promoting socio-economic development. Within this context, licensing is a key legal mechanism that enables re-use without sanction. However, there is evidence of a "licensing deficit" and this raises questions regarding best practice and sustainability in emerging African open data initiatives, particularly in the context of intermediaries being encouraged to exploit shared data for economic and social benefit. This article asks two main questions: (1) What is the current state of open licensing in two African open data initiatives; and (2) to what degree is it appropriate to focus on licensing as a key indicator of openness? Utilising a case study approach, the research explored licensing dynamics in the Kenya Open Data and the City of Cape Town Open Data initiatives, examining the contexts in which these initiatives were established and their resulting licensing frameworks. The cases reveal evidence of strategic engagement with content licensing, driven largely by the need for legal protection, adherence to international best practice and attraction of the user base required in order to ensure sustainability. The application of licensing systems in both contexts does, however, suggest an emerging system in which data providers are "learning by doing" and evolving their licensing practice as portals and their associated policy frameworks mature. The paper discusses the value of open data licensing as an indicator of organisational change and concomitant importance of taking into consideration the institutional dynamics when evaluating the organisational licensing frameworks of city, national and other governments.

KEYWORDS
open data, Kenya, Cape Town, licensing, open licences, Creative Commons

INTRODUCTION: THE ROLE OF LICENSING IN OPEN DATA PROVISION

This article is concerned with the current state of open licensing in two African open data initiatives and the extent to which licensing approach should be considered as an indicator of openness. Open data practice is gaining international momentum in the public sector and civil society as an important mechanism for sharing information, aiding transparency, and promoting socio-economic development. Vast amounts of data are being released for public consumption under the expectation that this activity can contribute to a better-informed citizenry, provide economic opportunities for intermediary partners, and improve administration. Within this context, open data licensing is a key legal mechanism for facilitating the lawful re-use of data (Davies et al., 2013; Dulong de Rosnay & Janssen, 2014). This article attempts to understand some of the contextual factors influencing the application of data licensing systems in an African open data context.

In the context of data sharing it is useful to distinguish between the notions of "gratis" and "libre". Gratis content is typically free for the user to download from the Internet without any cost, but with full copyright retained. Libre, on the other hand, refers to content which is, amongst other things, openly licensed and thus available for re-use, with certain provisos stipulated. While gratis data can be valuable in terms of simple information sharing, data shared in the libre context hold greater affordance for adaptation and remixing by intermediaries. Previous research (Davies, 2014) has demonstrated that such intermediaries are a crucial part of the open data ecosystem (Chattapadhyay, 2014; Roberts, 2014; Sein & Furoholt, 2012; Van Schalkwyk et al., 2013) and their engagement plays an important role in the sustainability of open data initiatives. Williams, Marcello and Klopp (2014) argue that how open access is provided to data is just as important as making it freely available (i.e., without cost). They point out that access is defined by context, connectivity and capabilities; and that these factors all come to bear on the effective utilisation of open data.

The World Bank (n.d.) uses the term "open data" with very specific meaning; data or content is open if anyone is free to use, re-use or redistribute it, subject at most to measures that preserve provenance and openness. It identifies two dimensions of openness: (1) The data must be legally open, i.e., placed in the public domain or licensed under liberal terms of use with minimal restrictions; and (2) the data must be technically open, i.e., published in electronic formats that are machine readable and preferably non-proprietary.

Open Definition (n.d.) lists "Licensing" as one of the three pillars that define an authentically open or libre resource – along with access (available via the Internet without charge) and open format (provided in a convenient and modifiable form such that there are no unnecessary technological obstacles to the performance of the licensed rights). This definition goes on to state that a licence is only authentically open if its terms satisfy the following conditions: use, redistribution, modification, separation, compilation, non-discrimination, propagation, application to any purpose, and free from charge.

In terms of international copyright law, copyright exists automatically in original works – meaning that no additional measures (other than physical creation) need to be taken for copyright protection to apply. Open data licences are used within this framework for copyright protected data to denote selected elements of traditional "all rights reserved" copyright protection which the author or copyright holder wishes to waive (such as the sole right to adapt or reproduce elements of that content). Open licences are applied to protect the creator and, at the same time, facilitate ease of re-use, thereby eliminating the need for permissions and additional contracting around application of content. It is a useful means of signposting whether and how data may be used by others, and whether there are any particular provisos or conditions associated with that use.

The absence of an open licence implies that all rights are reserved to the author or copyright holder, and serves as a potential barrier for re-use. It is therefore not only important that data are made open, but also that the potential users of such data are clear about being able to re-use data without fear of legal sanction (Janssen et al., 2012). In order for users to operate autonomously in this manner, licensing provisions should be expressed clearly and in alignment with other organisational terms of use or policies governing content distribution.

Open data licensing is expressed either through standard or bespoke licences. Standard licences are in some ways preferable in that they are re-usable, immediately recognisable amongst user communities without the need for complex legal interpretation, and (ideally) machine interoperable. The most common examples of standard licences include Creative Commons (CC),[1] Open Data Commons[2] and the Open Database License.[3] Bespoke licences are typically developed by governments and international organisations that engage in widespread open data sharing, and have a need for more detailed, specific terms of use. The benefits of these licences are that they enable an organisation to make explicit its particular concerns or provisos around the access and re-use of its data. The licence can also provide more detailed guidelines on how the data are to be applied and attributed. One example of a bespoke licence is the UK Open Government Licence.[4]

Terms of use can also contain elements of bespoke licensing. Typically associated with an entire website, platform or repository (as opposed to standard licensing, which is typically associated with an individual resource or digital object), terms of use can also complement standard licences. Some data providers take the approach of not licensing individual data sets, but instead articulating terms of use[5] for all content associated with a platform, portal or website. This is acceptable legal practice, but best practice in terms of open content sharing recommends that the licensing terms are embedded both in the actual data set (where possible) and in the descriptive metadata accompanying it.

There are indications that an increasing number of individuals and organisations worldwide are utilising open licensing when sharing content on the Internet. But while this practice is growing, there are still significant barriers to mainstream implementation. Creative Commons reported[6] in 2014 that there were 882 million CC-licensed works published on the Internet, with the most popular licensing provision (33%) being CC BY-SA (the CC licence requiring attribution and sharing of adaptations under similar licence conditions). In terms of geographical spread in application of CC licensing, the same Creative Commons report pointed out that only 1% of the CC-licensed content was associated with the African continent. Most of the content had come from North America (37%) and Europe (34%); with growing representation from the Asia-Pacific sector (16%), Latin America (10%), and the Arab World (2%). These figures are indicative of a deeper challenge that exists in Africa and other developing-country regions arising from low familiarity with the digital commons and a deficit in the skills, capacity and confidence required to engage strategically with open content licensing at various organisational levels (see, for example, Rizk, 2014).

A survey of development-related data sets in South Africa, conducted by Powell et al. (2012), found that while many NGOs, universities, research projects and government departments published data sets on their websites, few had explicit licensing statements. This aligns with the findings of the Open Data in Developing Countries (ODDC) initiative (Davies, 2014) with regard to the application of open licensing in developing-country contexts. ODDC found that in the spread of the open data initiatives surveyed:

> Very few datasets are clearly openly licensed, and there is low understanding of what open licenses entail. There are mixed opinions on the importance of a focus on licensing in different contexts. (Davies, 2014, p. 17)

The licensing deficit raises questions regarding best practice and sustainability in emerging African open data initiatives, particularly in the context of intermediaries being encouraged to exploit the shared data for economic benefit. In addition, the licensing deficit may be exposing the difficulties organisations bound by institutional dynamics

1 http://creativecommons.org/
2 http://www.opendatacommons.org/
3 http://opendatacommons.org/licenses/odbl/
4 http://www.nationalarchives.gov.uk/doc/open-government-licence/version/3/
5 http://web1.capetown.gov.za/web1/opendataportal/Images/OpenDataLicence2.pdf
6 https://stateof.creativecommons.org/report/

may face when attempting to come to terms with content licensing, and this may point to an often superficial and unsympathetic reading of open data licensing in governments across the globe.

This article therefore attempts to answer two main questions:

(i) What is the current state of open licensing in two African open data initiatives?
(ii) To what degree is it appropriate to focus on licensing as a key indicator of openness in the African open data context?

The next section of this article outlines our research methodology, followed by sections reporting and analysing findings from the two case studies of open data initiatives. We then provide discussion and conclusions.

METHODOLOGY

The study identified the Kenya Open Data Initiative (KODI) and the City of Cape Town Open Data (CCTOD) initiatives as its two principle sites of investigation. The sites' selection was based on the fact that they comprise two of the most prolific government open data initiatives in Africa. In addition, both aggregate data from a wide range of departments and agencies, and have demonstrated strategic engagement with open content licensing. At a national level, both South Africa and Kenya score consistently well in assessments by the Open Data Barometer[1] (a critical index that focuses on the context, availability and emerging impacts of open government data on the web), and by the Global Open Data Index[2] (which assesses the state of government data around the world). The prolific nature of these two countries in the open data landscape provided further incentive for site selection, as it provided confidence that we would find a level of activity sufficient to make the study viable.

Despite the fact that KODI is national-level and CCTOD is city-level, the two initiatives can be considered similar sites in technical and organisational respects and thus can serve as valuable lenses for understanding the dynamics at play in African public agencies concerned with open data provision.

Data on these two open data initiatives, and on their associated licensing dynamics, were collected via a two-phase data collection process, comprised of a general desk review followed by interviews. The desk review focused principally on obtaining a sense of the licensing practices and policy frameworks in the two cases. Where there were gaps in information, or where greater clarity was required to substantiate claims, interviews were undertaken with key informants working within each of the initiatives. Interviews were conducted either in person, via email, or telephonically, depending on geographic location and the preference of each interviewee.

FINDINGS

KENYA OPEN DATA INITIATIVE (KODI)

The KODI platform was launched in July 2011 with the intention of making Kenyan government data openly available through a single online portal (Kwamboka, 2013). The platform's launch came in the wake of a new national Constitution, adopted in 2010, which mandated a new era of public participation in government and altered the way in which Kenya's counties communicated with central government (Rose & Amolo, 2013). Under the new Kenyan Constitution, the right to information is enshrined in the Constitution's Bill of Rights. Article 35 of the Bill of Rights states that an individual has a right of access to information held by the State; and to information held by another person required for the exercise or protection of any right or fundamental freedom. It further imposes a duty on the State to publish and publicise any important information affecting the nation (Republic of Kenya, 2010).

Kenya, with East Africa's largest economy, is recognised as having a thriving information and communication technology (ICT) sector (Williams et al., 2014). The development of the KODI platform in 2010–11 took place in a context in which the ICT sector was expanding rapidly, and a number of factors collided to create an enabling environment. Jay Bhalla, Executive Director of the Open Institute and member of the government-appointed Task Force that led to the development of the KODI platform, attributes the birth of the KODI platform to a number of drivers (J. Bhalla, pers. comm., 2015), namely: (1) the new, more open policy environment enabled by the launch of the new Kenyan Constitution in 2010; (2) a strategic relationship between Kenya and the World Bank, which had embarked upon aggressive promotion of open data activity around this time; (3) a booming ICT sector enabled by the arrival of international undersea fibre optic cables that boosted Kenya's available bandwidth; (4) an explosion in the mobile telephony sub-sector; (5) significant relaxation of controls within legislation on investment in the ICT sector; (6) the emergence of ICT hubs and networks in need of data; and, most significantly, (7) the championing of the movement by the then-Permanent Secretary of the Kenya ICT Authority, Dr Bitange Ndemo.

Ndemo is largely acknowledged as being the father of the open data movement in Kenya, and he played a primary role in realisation of the KODI platform. In the absence of legal and policy frameworks for open data, the push to establish KODI was largely driven by Ndemo, who played the role of open data champion in government and lobbied intensely for support from the executive (Kenei, 2014). Bhalla of the Open Institute says that the principle arguments employed in the lobbying for the KODI platform were largely focused on job creation and a need to service the burgeoning ICT environment (with its associated intermediaries who wanted access to information in order to build applications for public consumption). The development and launch of the KODI platform can also be understood against a backdrop of significant activity in the area of e-government and in emerging innovation sectors that were driving economic activity. This momentum was supported and driven by Ndemo, who consistently employed arguments around economic benefit and job creation when confronted by government critics who were

1 http://barometer.opendataresearch.org/
2 http://index.okfn.org/

nervous about the risks that open data activity might bring (J. Bhalla, pers. comm., 2015).

The launch of the KODI platform made Kenya the first country in sub-Saharan Africa to have an open data portal, and the second on the continent after Morocco.[3] The ambition of the portal was to make core government developmental, demographic, statistical and expenditure data available in a useful digital format for researchers, policymakers, ICT developers and the general public (Mutuku & Mahihu, 2014), thus creating an "enabling infrastructure that could accelerate human and economic development throughout communities in Kenya" (Hopkins, 2012).

The KODI platform aggregates and shares data sets from Kenyan government ministries and agencies. The 2009 census data, as well as national and regional expenditure data and information on key public services such as education, health and agriculture, were some of the first data sets released (Mutuku & Mahihu, 2014). Originally much of this data, such as the census data, did not exist digitally (J. Bhalla, pers. comm., 2015), and so needed to be digitised in order to be curated and shared. The move towards more sophisticated data practice on the part of government ministries and agencies therefore coincided with the activity around the KODI platform.

KODI data are published online through the Socrata[4] platform, and users can view data sets at national, county and constituency levels in the categories of Education, Energy, Health, Population, Poverty, and Water and Sanitation. Data can be visualised using the online platform and also downloaded in a wide range of formats, including CSV, CSV for Excel, JSON, PDF, RDF, RSS, XLS and XML. At the time of writing, in June 2015, there were over 500 data sets available on the KODI portal. Data sets had been downloaded over 541,000 times and embedded into third-party websites and portals and viewed over 33,875 times. There had been 650 requests from the public and intermediaries for new data sets (S. Mawiyoo, pers. comm., 2015).

An online survey of reaction to the KODI platform and its open data sharing practice conducted as part of this study in early 2015 did, however, demonstrate significant negative response. Mutuku and Mahihu (2014) highlight the fact that the low quality of the available data hinders usage and limits value, but at the same time point out that intermediary technology has the potential to enhance access and usability of data. They also point out that supply of open data is still a challenge in Kenya because most government agencies are yet to fully embrace the constitutional directive around making data available, resulting in a large proportion of the available data sets being out of date.

The supply challenge is exacerbated by (1) unclear distinction between what can be safely shared and what is classified; (2) lack of technical capacity to produce/curate open data; and (3) lack of understanding of the exact mandate, as enshrined in law, for institutions to release data. For these reasons, "legal frameworks, including policy documentation with implementation plans and an access to information law, are necessary" (Mutuku & Mahihu, 2014, p. 31). This perspective is echoed by Brown (2013), who states that "Kenya's open data portal is floundering" because government agencies have been reluctant to release data (see also Mutuku & Mahihu, 2014; Wokabi, 2012), and because implementation of the multiple requirements of the new Constitution (which created a new devolved system of government) has hamstrung government officials who are trying to adjust to new roles and responsibilities, undermining officials' ability to incorporate open data into their workflows.

While the new Constitution recognises an individual's right to information, it has been argued that there is still a great need for a freedom of information (FOI) law to codify and implement this constitutional right. The absence of an FOI law has been cited as a significant inhibiting factor in the advancement of Kenya's e-government strategy as well as the sustained growth of the KODI platform (Brown, 2013; Kenei, 2014; Mutuku & Mahihu, 2014). At the time of writing, an Access to Information Bill along with a Data Protection Bill had been drafted and was awaiting debate in Parliament (S. Mawiyoo, pers. comm., 2015). Recognising the policy gap around the legal challenges of data publication, KODI also drafted an Open Data Policy in 2014, which, at the time of writing, was being evaluated by the ICT Authority. It is envisioned that this emergent policy and legislative framework will guide and set the standards for future release of public data.

In the current absence of a KODI-specific policy governing the licensing of the data shared via the portal, the legal framework is essentially being dictated by the ICT Authority's Terms of Use and Ndemo's directive, which was formulated in line with World Bank consultation and international best practice (both of which advocate for open licensing and third-party appropriation without restriction).

The KODI platform Terms of Use state that the Kenya ICT Board and the government agencies whose information is provided on the portal "impose no restrictions to the commercial and non-commercial reproduction, re-publication and re-distribution of any information published on the portal" (KODI, n.d.). Data sets have to date been published on the platform under either:

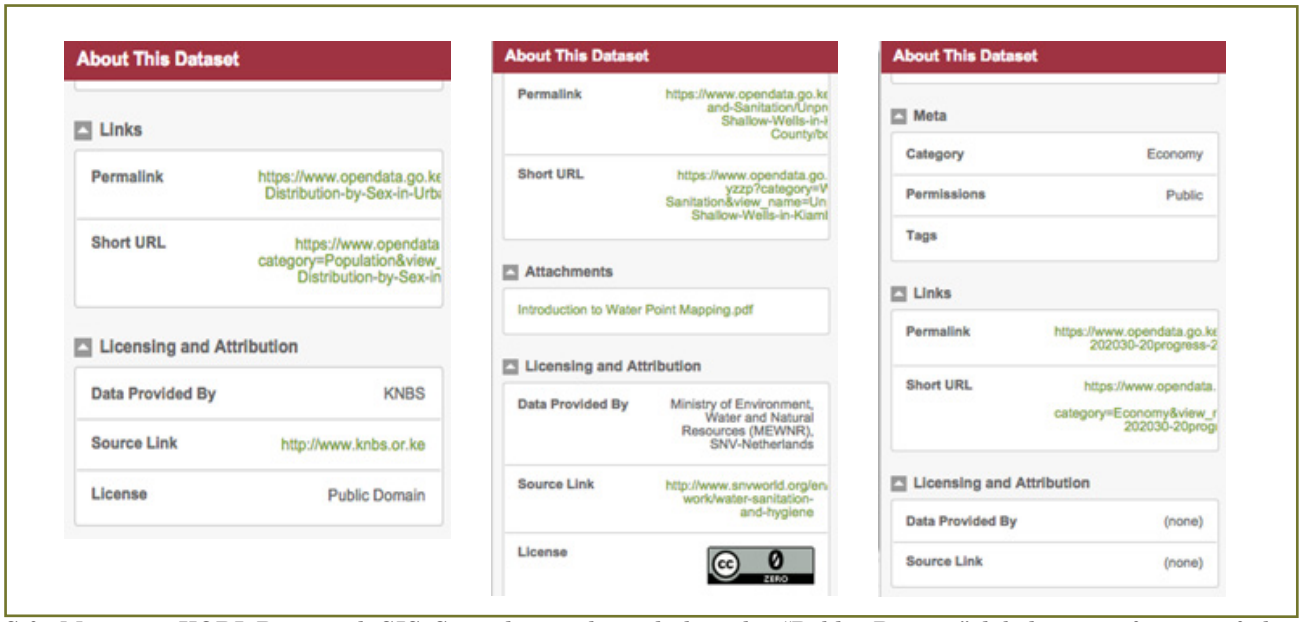
(1) the CC0 1.0 Universal Public Domain Dedication (CC0);
(2) a "Public Domain" statement; or
(3) in some cases no licensing provisions are indicated in the "Licensing and Attribution" metadata field (this typically takes place in instances where data suppliers have merely supplied pdfs of data sets).

Figure 1 shows metadata extracts from the "About" links on three KODI data sets, demonstrating the variations

3 https://opendata.go.ke/page/about
4 http://www.socrata.com/

outlined above. The first screenshot is from a data set[5] provided by the Kenya National Bureau of Statistics (KNBS) and is indicated as being in the "Public Domain". The second screenshot is from a data set[6] provided by the Ministry of Environment, Water and Natural Resources (MEWNR), SNV-Netherlands, and is identified as having a CC0 licence. The third is from a data set[7] provided by an unidentified source and does not carry any licensing statement.

FIGURE I:METADATA EXTRACTS FROM DATA SETS SHARED ON THE KODI PLATFORM



Sifa Mawiyoo, KODI Data and GIS Specialist, indicated that the "Public Domain" label was a feature of the administrative interface and had arisen due to the fact that the term was easier to grasp for metadata inputters who did not have knowledge of the CC licensing regime as well as being one of the first options in the Socrata platform's drop-down list of licensing options when ascribing metadata (pers. comm., 2015). Cases where there was no licensing statement had resulted from error or oversight in metadata input rather than a divergent strategic choice. Mawiyoo indicated that the KODI team was undertaking work to standardise metadata provisions in order to eliminate input errors of this nature, and that the approach to licensing management was evolving as the team became more familiar with licensing regimes.

In terms of KODI general practice, metadata were sometimes supplied by the relevant government ministry along with the data sets. If no metadata were provided, the KODI team attempted to source this information from a local contact at the source ministry, or otherwise undertook background research to populate metadata if the source ministry could not provide the relevant detail. In terms of overall process, the KODI team was acquiring data sets from the ministries and agencies, cleaning and formatting the data, ascribing metadata, and then publishing the data sets on the portal (S. Mawiyoo, pers. comm., 2015).

It is important to note that the "Public Domain" label was being employed as an administrative solution in the metadata context (largely based on its prominence in the standard Socrata platform interface) rather than being used to indicate the choice of the CC Public Domain Mark 1.0 – another tool provided by Creative Commons. CC0 and the CC Public Domain Mark are both CC public domain tools. CC0 licensing is applied in cases where the licensor is the copyright holder and wishes to waive all rights associated with a particular work worldwide (to the extent possible under law); while the Public Domain Mark is utilised in cases where a user has identified a work that is free of known copyright restrictions, and is merely "tagging" or identifying the copyright status of the work for other users.

ANALYSIS OF THE KODI LICENSING FRAMEWORK

In the absence of a national legal framework for open data provision, the KODI platform plays an important role in providing a platform on which user rights are entrenched in an open legal framework. The same data sets are, however, sometimes disseminated under different legal conditions via different delivery channels. An agency such as the Kenya National Bureau of Statistics (KNBS), for instance, shares data via its own website on a full copyright basis with all rights reserved. According to the KNBS website Terms of Use:

> You may download, print and store selected portions of the content of the site provided that you (l) only use these copies for your own personal, non-commercial use, (ll) do not copy or post the content on any network computer or broadcast the content in the media, and (lll) do not modify or alter the content in any way. (KNBS, n.d.)

KNBS does, however, simultaneously share selected data sets via the KODI platform, and KNBS data on that platform

5 https://www.opendata.go.ke/Population/Population-Distribution-by-Sex-in-Urban-Centres-an/yc6j-ekrh (retrieved 11 June 2015)

6 https://www.opendata.go.ke/Water-and-Sanitation/Unprotected-Shallow-Wells-in-Kiambu-County/bqhi-yzzp (retrieved 11 June 2015)

7 https://www.opendata.go.ke/Economy/Vision-202030-20progress-20report/p3t9-shd7 (retrieved 11 June 2015)

were at the time of writing indicated as being in the public domain. This disjuncture among licensing practices has come about as a result of a gap between policy and approach at the Kenya ICT Authority and at individual ministry or agency level, as well as a gap in understanding of open data licensing regimes. The ICT Authority and KODI position themselves in line with the international open data movement, meaning that they are forced to engage with open licensing systems, interpret a complicated licensing framework and ascribe appropriate metadata on behalf of data providers in order to facilitate third-party engagement. The state entities that provide the data have to date not faced the same strategic imperative.

While application of the "Public Domain" label in the KODI context arose from an inputting error when ascribing metadata, it points to a deeper challenge in terms of the complexities of not only getting to grips with licensing systems, but also developing an understanding of the technical principles of what it means for content to be in the public domain. The commonplace assumption might be that this refers to the fact that content is freely available for download on the Internet. There is, however, considerable complexity associated with the legal definition of this concept, particularly in the context of data publication.

In terms of international convention, a work is in the public domain either: (a) when the copyright holder has waived all rights associated with the work; or (b) the work, under various conditions, does not enjoy protection under the provisions of the Copyright Act in a certain jurisdiction. As such, content that is in the public domain in one country may not be in the public domain in another. Examples of instances when work does not enjoy protection under a Copyright Act include cases when the defined copyright period has expired, or when work is not eligible for copyright protection in the first place (such as instances where the work is a mere expression of fact). Copyright laws vary by jurisdiction, in terms of duration of protection and what constitutes copyrightable subject matter (Creative Commons, n.d.). As such, one of the only ways to determine with certainty whether something is in the public domain is when the copyright holder dedicates the work under a CC0 licence. CC licences do not, however, affect the status of a work that is in the public domain under applicable law (such as when duration has expired or the work is not eligible), because these licenses only apply to works that are protected by copyright.

In order for KODI to ensure cohesion and correct legal application of licensing terms it is crucial that government data-provider agencies develop a deeper understanding of licensing systems in order that they can provide a clearer directive on the usage rights and provisions ascribed to their content. According to Bhalla (pers. comm., 2015), government ministries and agencies essentially allowed their data to be released openly (under CC0 and Public Domain statements) via the KODI platform because of the directive from Ndemo, while there was little understanding of content licensing amongst these ministries and agencies. This sentiment was echoed by KNBS Senior Manager of Data Processing Mutua Kakinyi (pers. comm., 2015), who indicated that there was little to no working knowledge of CC and the public domain legal framework within the national statistics agency.

Significant challenges therefore exist in bringing about cohesion in licensing frameworks for release of government data, both within individual platforms and initiatives as well as across agencies. This will require investment in the up-skilling of agencies to deal with copyright issues and a coordinated approach around data release and ownership principles. Bhalla (pers. comm., 2015) indicated that the initiative around the FOI Act was, among other things, aimed at addressing this deficit and creating technical capacity in government ministries to release data in an openly licensed format. In line with this new initiative, all government services down to county level are to implement systems to facilitate digital data gathering and curation. This activity will need to be matched by capacity development in licensing systems across multiple levels of the agencies involved if data sets are to be shared in an optimally open and legally appropriate manner.

CITY OF CAPE TOWN OPEN DATA (CCTOD) INITIATIVE

The CCTOD portal[8] was launched in January 2015, in line with the City's Open Data Draft Policy[9] of February 2014 (City of Cape Town, 2014). The launch of the portal established Cape Town as the first city in Africa to establish an open data presence, and positioned it amongst an international group of cities that have in recent years launched similar initiatives in line with e-government strategies.

The City's Open Data Draft Policy recognises four issues necessitating the establishment of an open data portal: (1) the role played by data in the economy and society is changing; (2) innovators and entrepreneurs are using data sets to design new kinds of products, to enhance competitiveness, to build social capital, and to engage in civic life; (3) the City's useful information is often hidden and data access policies and procedures within the City impede public access; and (4) the City's various websites are often not user-friendly, with the information they contain sometimes out of date and/or not in machine-readable format (City of Cape Town, 2014). The CCTOD initiative aims to address these issues and to create "an enabling environment to attract investment that generates economic growth and job creation" (City of Cape Town, 2014, p. 3).

Recent research by Bagui and Bytheway (2013) indicates that:

> the use of mobile, web and social media technologies is widely expected to be an important feature of improving public participation in government in the City of Cape Town, but … the necessary transformation that would enable it is far from complete. (Bagui & Bytheway, 2013)

The CCTOD initiative is a step towards addressing the public participation goal as highlighted by Bagui and

8 https://web1.capetown.gov.za/web1/opendataportal/

9 "Draft" is contained in the title of the final published policy.

Bytheway (2013), and has significant synergies with a number of other strategic activities, such as the City's Smart City Strategy,[10] which aims for the City to advance its digital agenda and become more open in offering services to the public. In this sense, the CCTOD initiative forms part of the broader ambition to ramp up information management and e-government strategy at City level and to increase engagement with stakeholders from the broader community.

At the time of writing in June 2015, the CCTOD portal hosted 33 data sets, covering community services, natural resources and the environment, basic services and infrastructure, transportation, spatial planning, finance, health, safety and security, land administration, and political and administrative boundaries. The platform was custom-developed by the City's Information Services and Technology (IS & T) Department for the City's Development Information and Geographic Information Systems (DI & GIS) Department (the custodian of the portal), and typically makes data sets available as a combination of CSV, KMZ, XLSX and SHP files. All data sets are accompanied by metadata, which include data set name, document name, description, date added, time coverage, spatial coverage, subject, file size, format, usage considerations and update frequency. (Because the portal had only been in existence for five months at the time of writing, it was not possible to comment on usage and uptake factors, e.g., the degree to which departments are providing data, or the extent to which data are being downloaded and used.).

The City's open data team, principally located within the DI & GIS Department, has solicited data for the portal from various City departments through internal processes. In addition, users are able to request data sets via the portal through its "Suggest a data set" link. Data requests are reviewed by an Open Data Steering Committee comprised of City Council representatives and two external stakeholders from the local open data community. The Committee convenes on a quarterly basis to review requests for additional data sets. This regulatory mechanism has provoked some criticism from local open data advocates, as it is seen by some as an inhibiting, gatekeeping mechanism with the potential to slow down the release of data (Eyal, 2014). It does, however, play an important role in providing a sense of regulation and security for City departments sharing their data on the portal (X. Limberg, pers. comm., 2015).

The impetus for launching the CCTOD initiative came from parallel Western Cape Province and City of Cape Town activities initiated in 2013. It arose out of a process initiated by the City Mayor's Office after the City was awarded World Design Capital status in 2014, following the initiation of an Open Data Forum by the Western Cape MEC for Finance, Economic Development and Tourism (K. Smith, pers. comm., 2015). The provincial forum drew together a wide range of relevant stakeholders and interest groups and resulted in a proposal to develop an open data policy. This activity was then advanced further with the support of the Mayor's Office.

Following the recommendations of the provincial forum, the City's DI & GIS Department put together a discussion document around open data policies and practices and facilitated an extensive internal consultation in order to obtain organisational buy-in on the policy. DI & GIS was tasked with this function because of its central role in the City's information management, with the open data process seen as a means to bolster this area of activity within the City. The discussion document generated by DI & GIS served as the basis for the Open Data Draft Policy published in early 2014 (City of Cape Town, 2014).

According to Keith Smith, Head of DI & GIS, when the City's internal consultation process was launched, the idea of open data was fairly new to the City and was met with mixed responses (K. Smith, pers. comm., 2015). At senior management level in the City, there were individuals who were supportive, but there were concerns around potential impact on existing resources, possible risks (such as incorrect representation of data), and the priority level of the initiative (given the fact that there were a number of other competing imperatives). There was also some concern that some of the City departments that were selling data would lose out on revenue. Obtaining the necessary internal buy-in during the consultation process was aided by referencing international examples and by support from the Mayor's Office.

The legal conversation around copyright management and licensing formed part of the exploratory internal consultation process. Smith (pers. comm., 2015) stated that the overall approach to the open data project was largely one of learning by doing, and that copyright considerations were not included in the City's published Open Data Draft Policy because the idea was to have a policy that was as short and simple as possible. A decision was made to articulate legal provisions through a customised licence in a separate Terms of Use document (City of Cape Town, n.d.).

Neil Hoorn, the developer within DI & GIS tasked with the project management of the CCTOD portal, said that the licensing consultation process was informed by dual imperatives: (1) providing free access to the data and (2) protecting the City against any liability that might arise from inappropriate use of the data. In this sense, the licensing conversation aimed to address both internal (organisational) and external (user) considerations (N. Hoorn, pers. comm., 2015). The consultation process was also seen as an important mechanism for providing legal assurance to internal departments in terms of the anxieties associated with open data sharing (K. Smith, pers. comm., 2015).

The main actors in the articulation of the CCTOD licensing approach were the City's Legal Services department and DI & GIS. Consultation was undertaken with Creative Commons South Africa[11] in order to learn more about CC licensing, but CC licences were considered too generic as they were seen as not dealing with some of the specific issues raised by stakeholders during the discussion and review stage of the City's Open Data Draft Policy. Specifically, the open data team felt that the licensing needed to be more explicit with respect to illegal or inappropriate use of its data (N. Hoorn, pers. comm., 2015).

10 http://web.capetown.gov.za/eDocuments/Smart_City_Public_Private_Partnership_Conference_228200310231_389.pdf

11 One of the authors of this article, Dr. Tobias Schonwetter, participated in these discussions in his capacity as Legal Lead for Creative Commons South Africa.

This was in spite of the fact that CC licences do contain disclaimers and provisions limiting the liability of the licensor. It was eventually decided that the City should create a bespoke terms of use licensing statement that would make usage conditions as explicit as possible. Draft terms of use were considered by the Steering Committee and adopted with minor revision.

The CCTOD portal's adopted Terms of Use is a four-page document downloadable as a pdf via the "Terms of use" link on the portal home page (City of Cape Town, n.d.). It opens with a "Disclaimer" stating that the City "makes data available without any remuneration"; "makes no representations and warranties […] about the completeness, accuracy, reliability, suitability or availability of data on the website"; and will "not be liable for any errors, omissions, or inaccuracies in the data provided" (City of Cape Town, n.d., p. 1). The next section, on "Use of Data", outlines the usage provisions for the user, but does so from a regulatory perspective in that it emphasises what the user is not allowed to do. According to the Terms of Use, the user specifically undertakes (1) only to use the data for a lawful purpose; (2) not to use the data to commit a criminal offence; (3) not to use the data to infringe any lawful entitlement; (4) not to use the data to impersonate another misrepresent identity; and (5) not to alter, damage or delete any content or load any harmful programmes, computer code or files that may alter, damage, interrupt or limit access to data. The Terms of Use further state that "the User is required to explicitly state that the City does not warrant or guarantee the quality or accuracy of the data" (City of Cape Town, n.d., p. 2).

The Terms of Use document makes no explicit mention of commercial application of the data, but states that the "User may use the data contained on this site free of charge" (City of Cape Town, n.d., p. 2). From a legal perspective, this implies that commercial for-profit application is permissible. However, this might not be immediately apparent to a user who is not well versed in legal matters.

In addition, while the CCTOD portal's Terms of Use make it explicit that the conditions apply to visitors and users of the Open Data Portal section of the City's website, uncertainty about commercial application could be compounded by the copyright notice of the broader City website (of which the Open Data Portal is part). That City website copyright notice states that:

> Any redistribution or reproduction of part or all of the contents in any form is prohibited, other than the following:
> - you may print or download to a local hard disk extracts for your personal and non-commercial use only;
> - you may copy the content to individual third parties for their personal use, but only if you acknowledge our website as the source of the material […]. (City of Cape Town, 2008)

The appearance, in the licensing approach, of foregrounding organisational concerns over affordances for the user again arises in the "Use of Data'" provision in the Terms of Use, which states that the user undertakes not to "alter, damage or delete any content" (City of Cape Town, n.d., p. 2). This prioritisation of organisational concerns is also present in the "Intellectual Rights" sub-section of the Terms of Use, which explicitly prohibits copying and reproduction of data. The sub-section states that: (1) all intellectual property rights "shall remain vested at all times in the City"; and (2) "[t]he user shall not, under any circumstances whatsoever reproduce, copy or use the City's IP or permit the use of the City's IP by any third party without the City's prior written consent" (City of Cape Town, n.d., pp. 3-4).

ANALYSIS OF THE CCTOD LICENSING FRAMEWORK

The City's Open Data Draft Policy stresses the need to use open data to attract investment that generates economic growth and job creation (City of Cape Town, 2014), and the City informants interviewed for this research all stressed the importance of open data's economic innovation component, maintaining that the portal's data was available for any kind of use. Against this backdrop, the focus on a legally oriented statement that serves to address organisational anxieties around use and quality – as opposed to a user-oriented statement that is designed to facilitate re-use and economic exploitation – may be viewed as one of the principal shortcomings of the City's licensing approach. While the intention behind this provision was to counter any malicious hacking or illegal use of the site and its contents, it is possible that third-party intermediaries who wish to remix and adapt the content will view this provision as a major stumbling block for any such activity, or be required to address a query to the City in order to clarify their rights in this regard.

In this sense, the Terms of Use may serve as an effective means for protecting the City's interests, but cannot be said to be an optimal open licensing solution in that the fundamental provision of commercial exploitation is not made explicit and adaptation appears to be disallowed, despite the insistence of City representatives to the contrary. Uncertainty of this kind stands to be a potential barrier in terms of applying City data in an open data ecosystem where freedom and flexibility around appropriation of data are required. If intermediaries were to utilise the data they would most likely be required to engage the City in conversation around the scope of usage rights and permissions in order to obtain clarity, rather than just relying on a licence statement. While this may be acceptable and even desirable for the organisation during the initiative's fledgling stage, it raises questions around scalability and long-term viability. It is also counter-productive in terms of the fundamental purpose of open licensing systems, which is to circumvent permission seeking.

While the City's open data provision falls within the realm of gratis (i.e., available on the Internet for access without

charge), it does not appear to be authentically libre content in line with international protocols around re-use. There also appears to be a tension or contradiction between the ambitions of the City's Open Data Draft Policy (which speaks to economic exploitation and innovation) and the Terms of Use (which regulate activity in terms of commercial application and adaptation). At present, the large number of anxieties around organisational risk associated with sharing open data has manifested in a relatively conservative licensing approach, when evaluated against international open data protocols. This is, however, not uncommon in the licensing strategies of city and other large-organisation data initiatives. In the case of the City of Cape Town, organisational realities and the groundbreaking nature of the initiative make the licensing of content a particularly challenging component in terms of balancing risk management and user rights.

DISCUSSION

The cases examined reveal evidence of strategic engagement with content licensing. In both the KODI and CCTOD initiatives, there are indications that licensing has been addressed as a key component of the strategies for establishment of the initiatives. These strategic engagements have been driven largely by a mixture of the need for legal protection, the desire to adhere to international best practice, and the need to attract the interactive user base that is required to build sustainability.

While there is strategic engagement with the licensing question in both initiatives, the actual implementation of licensing systems in the KODI and CCTOD contexts also suggests an element of "learning by doing" in both cases, through which licensing practices are evolving as the portals and associated policy frameworks mature. While this may not be ideal for the user community or in line with strict open data principles, it is understandable in the context of large intra-institutional initiatives that are innovating in an emerging terrain. The case studies have identified a patchwork of licensing systems being applied, with a mix of both standard and bespoke licences and some vacillation in how licensing is expressed. One of the main problems in this situation is licensing compatibility: differently licensed materials (particularly the combination of CC and bespoke licences) pose a challenge for users needing to integrate content with divergent licensing provisions. In addition, the lack of clarity is a significant challenge for optimal third-party engagement as well as internal buy-in and long-term sustainability.

The data generated by our interviews suggest that this patchwork scenario 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 outside of the institution. In the case of the KODI initiative, this situation is compounded by the absence of a legislation and policy framework to govern open data provision and the terms under which data is released, resulting in a divergent licensing approach. In the case of the CCTOD initiative, the Terms of Use display a tension between legal protection of the data-providing organisation on the one hand and user application on the other. Given the foregrounding by the international open data movement of licensing as a key indicator of openness, this situation raises questions about how one balances ease of understanding for the user, adherence to formal protocols on best practice, and cohesion amongst intra-organisation entities.

In response to the first of our two research questions - as to what is the current state of open licensing in two African open data initiatives - the findings from this study suggest that the current state of licensing is nascent and practice is manifesting in a non-uniform fashion. There are also challenges in promoting the understanding and application of open licensing systems amongst members of data-provider organisations – particularly as it relates to the ambition of publishing authentically libre content that can be legally exploited for commercial application. Current engagement with licensing appears to be limited to specific senior divisions of the data-providing organisations, and an imperative exists for this knowledge to filter through to other levels of the organisations. This need to conscientise and develop capacity in various sectors of the data-provider environment around open licensing coincides with the need to gain the trust and buy-in of various sectors of the data-providing organisation in order to ensure a sustainable stream of data provision.

In response to our second research question – as to whether is it appropriate to focus on licensing as a key indicator of openness in the African context – the challenges that exist around organisational coherence, and the imperative to transition from gratis to libre practice so as to be in line with international open data protocols, suggest that licensing stands to be a crucial means of consolidating activity and aligning practice across government as an organisation. Licensing has the more obvious potential to be a valuable regulating, standardisation and security mechanism, but it can also serve as an indicator of the extent to which open data practice is being embedded across the organisation and the extent to which effective policy development has occurred.

The act of assigning an open licence to a data set indicates an understanding of what constitutes open data, in particular the fundamental principle of re-use; and it indicates that organisational actors have cognitively come to terms with and accepted the consequences of data being re-used without restriction.

From our case studies, there is evidence of two types of pressures that organisational actors are vulnerable to with regard to open data practice. First, there are exogenous pressures, i.e., pressures to adopt and internalise similar structures and procedures as those of other organisations within the institutional field. In the case of Kenya, a supranational agency (the World Bank) is seen to be exerting this pressure to conform, and in the case of the City of Cape Town there is an aspiration to emulate the open data initiatives of international, high-status cities.
Second, are endogenous pressures, i.e., pressures from within the organisation, applied to internal operational

units. In Kenya, the then-Permanent Secretary of the ICT Authority applied pressure for government to conform to international "best practice", and in Cape Town, the Mayor's Office directed the City to match its northern counterparts. The pressure exerts directly and indirectly on those in the operational units of governments (the departments and agencies). But at the same time, actors in these units are institutionally bound and tend to favour compliance with institutional norms and values. Should a conflict exist between institutional norms and values and the pressures being exerted on actors from outside of the institution, then it is likely that the institutional actors will "decouple" in order to create a buffer to protect themselves from these conflicting pressures for change. In the case of Kenya, there is already evidence of departments not releasing data to KODI.

From a licensing perspective, Dulong de Rosnay and Janssen (2014) point out that the obstacles to unrestricted open data use caused by lack of harmonisation in data licensing are legal and technical, but they are also institutional and cognitive. Based on their findings, they recommend awareness and education at all levels (policy-, operational- and user-level).

Of interest in Kenya is the expectation that the new FOI Act, coupled with a boost in technical capacity, will be sufficient to result in a more streamlined and coherent data licensing landscape. Our findings from the two cases we studied seem to suggest that in addition to such measures, attention also needs to be directed towards how to align the norms and values of all actors across government as institution, in order to harmonise open data licensing as an important step to embedding open data practice as a taken-for-granted activity.

CONCLUSIONS

The case studies presented in this paper suggest that the African open data licensing landscape will continue to be expressed through a mix of standardised and bespoke licences, as well as other customised statements such as terms of use – in some cases applied at object level to the individual data set, and in other cases on a site-wide or platform basis. This situation reflects a significant set of under-appreciated institutional dynamics of the data publisher as well as a significant need for capacity development in understanding of open licensing systems amongst data-provider organisations.

It will be valuable to undertake further research on whether more cohesion manifests itself in the licensing environment as the African open data community grows and providers overcome the anxieties associated with the unknown.

ACKNOWLEDGEMENTS

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.

The authors wish to thank KODI Data and GIS Specialist Sifa Mawiyoo for his willingness to engage, time spent in document review, and contribution to sourcing data; as well as Jay Bhalla, Executive Director of the Open Institute, for providing valuable understanding of the KODI context. We also wish to acknowledge the significant contribution of senior City of Cape Town DI & GIS team members, Keith Smith and Neil Hoorn, for their crucial insight into the CCTOD initiative.

REFERENCES

Bagui, L., & Bytheway, A. (2014). Exploring e-participation in the City of Cape Town. *The Journal of Community Informatics*, *9*(4). Retrieved from http://ci-journal.net/index.php/ciej/article/view/982/1052

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

Cabinet Office. (2012). *Open data white paper: Unleashing the potential*. Presented to UK Parliament by the Minister of State for the Cabinet Office and Paymaster General. Retrieved from https://www.gov.uk/government/publications/open-data-white-paper-unleashing-the-potential

Chattapadhyay, S. (2014). Access and use of government data by research and advocacy organisations in India: A survey of (potential) open data ecosystem. Presentation at the 8th International Conference on Theory and Practice of Electronic Governance, 27-30 October.

City of Cape Town. (2008). City of Cape Town website copyright notice. Retrieved from http://www.capetown.gov.za/en/Pages/copyright.aspx.

City of Cape Town. (2014). *Open data draft policy*. Retrieved from http://www.capetown.gov.za/en/PublicParticipation/Documents/HYS_Open_Data_Draft_Policy_version_21_February_%202014.pdf

City of Cape Town. (n.d.). Open data terms of use. Retrieved from https://web1.capetown.gov.za/web1/opendataportal/Images/OpenDataLicence2.pdf.

Creative Commons. (n.d.). Public domain. Retrieved from https://wiki.creativecommons.org/wiki/Public_domain

Davies, T. (2014). *Open data in developing countries: Emerging insights from phase I: ODDC report*. Retrieved from: http://www.opendataresearch.org/content/2014/704/open-data-developing-countries-emerging-insights-phase-i

Davies, T., Perini, F., & Alonso, J. (2013). *Researching the emerging impacts of open data: ODDC conceptual framework*. ODDC Working Papers #1. Retrieved from http://www.opendataresearch.org/content/2014/667/researching-emerging-impacts-open-data-oddc-conceptual-framework

Dulong de Rosnay, M., & Janssen, K. (2014). Legal and institutional challenges for open data across public sectors: Towards common policy solutions. *Journal of Theoretical and Applied Electronic Commerce Researc*h, *9*(3), 1-14.

Eyal, A. (2014). Cape Town's open data policy. Time to celebrate? Retrieved from ttp://code4sa.org/2014/09/27/capetown-opendata-policy.html

Hopkins, C. (2012). Kenya launches sub-Saharan Africa's first national open data initiative. Retrieved from http://readwrite.com/2011/07/08/kenya_launches_africas_first_national_open_data_in

International Council for Science (ICSU). (2014). *Open access to scientific data and literature and the assessment of research by metrics*. Retrieved from http://www.icsu.org/general-assembly/news/ICSU%20Report%20on%20Open%20Access.pdf

Janssen, M., Charalabidis, Y., & Zuiderwijk, A. (2012). Benefits, adoption barriers and myths of open data and open government. *Information Systems Management*, *29*, 258-268.

Kenei, S. (2014). Tracing the evolution of open data in Kenya. Retrieved from http://devinit.org/old_site/tracing-evolution-open-data-kenya/

Kenya National Bureau of Statistics. (KNBS). (n.d.). Terms of use. Retrieved from http://www.knbs.or.ke/index.php?option=com_content&view=article&id=193&Itemid=1070

Kenya Open Data Initiative (KODI). (n.d.). Terms of use. Retrieved from https://opendata.go.ke/terms-of-service

Korn, N., & Oppenheim, C. (2011). *Licensing open data: A practical guide* (version 2.0). JISC, Discovery, and Research Libraries UK (RLUK). Retrieved from http://discovery.ac.uk/files/pdf/Licensing_Open_Data_A_Practical_Guide.pdf

Kwamboka, L. (2013). *Open data: How Kenya did it*. Kenya Open Data Initiative. Retrieved from http://datascience.co.ke/wp-content/uploads/2015/03/Open-Data-%E2%80%93How-Kenya-Did-It.pdf

Mutuku, L., & Mahihu, C.M. (2014). *Open data in developing countries: Understanding the impacts of Kenya open data applications and services*. Nairobi: iHub Research. Retrieved from http://www.opendataresearch.org/sites/default/files/publications/ODDC%20Report%20iHub.pdf

Open Definition. (n.d.). Open definition (version 2.0). Retrieved from http://opendefinition.org/od/

Powell, M., Davies, T., & Taylor, K.C. (2012). *ICT for or against development? An introduction to the ongoing case of Web3*. IKM Working Paper No. 16, March. Retrieved from http://wiki.ikmemergent.net/files/1204-IKM-Working_Paper_16-WEB3-Mar_2012-2

Republic of Kenya. (2001). Copyright Act No. 12 of 2001. Retrieved from http://portal.unesco.org/culture/en/files/30229/11416612103ke_copyright_2001_en.pdf/ke_copyright_2001_en.pdf

Republic of Kenya. (2010). The Constitution of Kenya. Retrieved from https://www.kenyaembassy.com/pdfs/The%20Constitution%20of%20Kenya.pdf

Rizk, N. (2014). From de facto commons to digital commons? The case of Egypt's independent music industry. In J. De Beer, C. Armstrong, C. Oguamanam, & T. Schonwetter (Eds.), *Innovation and intellectual property: Collaborative dynamics in Africa*. Cape Town: UCT Press. Retrieved from http://www.openair.org.za/images/9781775821427.pdf

Roberts, A. (2014). Making transparency policies work. Freedominfo.org. Retrieved from http://www.freedominfo.org/2014/10/making-transparency-policies-work/

Rose, J., & Omolo, A. (2013). *Six case studies of local participation in Kenya*. World Bank. Retrieved from https://openknowledge.worldbank.org/handle/10986/17556

Sein, M., & Furuholt, B. (2012). Intermediaries: Bridges across the digital divide. *Information Technology for Development*, *18*(4), 332-344. doi: 10.1080/02681102.2012.667754.

Van Schalkwyk, F., Willmers, M., Czerniewicz, L., & Musundwa, S. (2013). Viscous open data: The flow of data in a public university

governance ecosystem. Paper presented at ICT for Development Conference, December, Cape Town.

White African. (2011). Africa's first national open data initiative: Kenya. Retrieved from http://whiteafrican.com/2011/07/07/africas-first-national-open-data-initiative-kenya/

Williams, S., Marcello, E., & Klopp, J.M. (2014). Toward open source Kenya: Creating and sharing a GIS database of Nairobi. *Annals of the Association of American Geographers 104*(1), 114–130. Retrieved from http://www.tandfonline.com/doi/abs/10.1080/00045608.2013.846157

Wokabi, C. (2012, 10 November). Kenya open data Initiative has hit a dead end, says PS. *Daily Nation*. Retrieved from http://www.nation.co.ke/business/news/Open+data+initiative+has+hit+a+dead+end/-/1006/1617026/-/11tfwouz/-/index.html

World Bank. (n.d.). Open data defined. Retrieved from http://toolkit.dev.zognet.net/en/essentials.html