

Chapter Title: Open data and records management — activating public engagement to improve information: case studies from Sierra Leone and Cambodia

Chapter Author(s): Katherine Townsend, Tamba Lamin, Amadu Massally and Pyrou Chung

Book Title: A Matter of Trust

Book Subtitle: Building Integrity into Data, Statistics and Records to Support the Achievement of the Sustainable Development Goals

Book Editor(s): Anne Thurston

Published by: University of London Press, Institute of Commonwealth Studies. (2020)

Stable URL: <https://www.jstor.org/stable/j.ctv14t45v5.12>

---

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact [support@jstor.org](mailto:support@jstor.org).

Your use of the JSTOR archive indicates your acceptance of the Terms & Conditions of Use, available at <https://about.jstor.org/terms>



This book is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License (CC BY-NC-ND 4.0). To view a copy of this license, visit <https://creativecommons.org/licenses/by-nc-nd/4.0/>. Funding is provided by AHRC, Knowledge Unlatched Select 2020: HSS Frontlist.



*University of London Press, Institute of Commonwealth Studies* are collaborating with JSTOR to digitize, preserve and extend access to *A Matter of Trust*

## **7. Open data and records management – activating public engagement to improve information: case studies from Sierra Leone and Cambodia**

*Katherine Townsend, Tamba Lamin, Amadu Massally and Pyrou Chung*

Open data initiatives support transparency, innovation, the promotion of a knowledge-based society and the advancement of democratic principles. Data in the hands of citizens can facilitate empowerment and support improved government efficiency and accountability. Open data promotes transparency by enabling citizens to reanalyse data underpinning government decisions and to monitor the impact of government policies. Citizens with access to the same government data used by policy-makers are more informed and better able to participate in and contribute to policy-making. Through their access to the administrative data generated by government, they are also able to identify incidents of corruption. For instance, citizen advocacy groups could potentially draw upon and analyse data derived from government payroll records, budgetary records, teacher employment records and other sources to assess the level of corruption in hiring teachers.

Open data provides an excellent vehicle for civic engagement, for information sharing, for rapid response and for supporting citizens' rights. The quality and integrity of the data underpinning open data initiatives and the ability to trace decisions documenting how the data were collected, processed and manipulated is key to achieving these benefits through time. The data management community is committed to the goal of achieving high-quality data, especially in the context of the Sustainable Development Goals (SDGs), but the role and importance of records management in providing evidence of the quality and integrity of the data through time has not yet received adequate attention.

In an open data environment, citizens should be able to trust that their governments are providing data with sufficient quality and integrity through an open data initiative that they can use it confidently. If it cannot be demonstrated that the processes used to produce the data and the data themselves are trustworthy, citizens' trust of the government can be damaged. Records, if well-managed, can provide the evidence needed to have confidence in the data. The role of records, and the steps required to manage them effectively, are

only beginning to be appreciated by those implementing open data initiatives. This chapter draws on examples from Sierra Leone and Cambodia to illustrate the positive impact of open data initiatives for citizens and, at the same time, the role that records management can play in ensuring that data quality and integrity can be demonstrated through time.

The example from Sierra Leone focuses on the goal of achieving free and fair elections, while the Cambodian example concentrates on successful land investment mapping. Each of these examples begins with an overview of an open data initiative and its role in advancing democratic values, knowledge dissemination and accountability. Both then go on to explore the extent to which records management can help strengthen the processes and the data produced and how a high level of quality and integrity can be sustained through time. A concluding section uses the two examples to consider the nature of the potential relationship between the open data and records management communities and the benefits for maximising the value of open data initiatives for citizens and governments alike.

## Sierra Leone

### *Open data in support of free and fair elections*

Sierra Leone is a small country in West Africa with seven million people. The nation has held elections every five years since 1996. In addition to the National Electoral Commission (NEC), there is a civil society organisation, the National Elections Watch, that aims to represent the people and to watch over NEC actions. Sierra Leone also has a Right to Access Information Commission (RAIC) that is responsible for making data and government transparent, including elections data. RAIC hosts an Open Data Council, comprising representatives from various private sector organisations, government agencies, academia and NGOs to make public sector data available and useable. Support for developing the RAIC and establishing the Open Data Council has been driven in large part by the need of the people of Sierra Leone to access government data and records and hold the public sector accountable.

At the time of the 2017 national elections in Sierra Leone, the government's open data portal had been down for months. The data were not being updated, and concerned and frustrated citizens and private sector organisations had decided to work together to start a parallel open data portal for anyone to use. This duplication caused some contention, with the government perceiving that its role had been overtaken. Nevertheless, discussion and debate occurred on a diversely represented WhatsApp group of approximately 250 professionals, journalists, government representatives and international actors.

Initially, the government wanted the alternative website taken down on the grounds that it should be the sole arbiter of open data. However, ultimately, the need for a consistent portal prevailed. The government had to bow to strong public opinion; as it was not providing the service itself and could not identify the harm being done by someone else playing this role, the site should continue. One company, LAM-TECH Consulting, which has now rebranded itself as TpISENT (SL) Limited, took on the primary role of hosting data and updating it with new datasets. After running the portal for several years, the team decided to focus on election monitoring and established the Sierra Leone Open Elections Data Portal (SLOEDP).

The portal is a resource that makes it possible for anyone to collect, aggregate, share and socialise elections data in an open format. The data adhere to the National Democratic Institute's open elections data format, which follows nine principles, namely that the data are timely, granular (at the finest level of detail possible), available for free on the internet, complete and in bulk, analysable, non-proprietary, non-discriminatory, licence-free and permanently available.<sup>1</sup> Unusual among many open data initiatives in developing economies, this movement began without any global donor funding but rather through people coming together through volunteer effort, self-funding and small-scale crowdfunding. The result was a solution sponsored and provided by Sierra Leoneans, for Sierra Leoneans.

In Sierra Leone, traditionally predictions of election results have been made by individual parties and candidates, and election results also are announced by the candidate and party. Misinformation, mismatched results and confusion had led to charged debate and even violence and bloodshed. SLOEDP's primary goal was to reduce violence by making information more accessible and easier to understand and trust. Its team included individual volunteers and organisations, all of them invested in election monitoring in Sierra Leone with the aim of preventing or at least reducing violence during the election cycle. They were well aware that violence would reduce the prospects for long-term, durable peace and stability and would undermine economic growth by limiting the purchasing power of citizens.

A strong WhatsApp group developed, including members of the Federation of Civil Society and Media Organisations (NaFCSMO-SL), Democracy and Development Associates – Sierra Leone (DADA-SL), the Open Government Initiative (OGI), the Women's Situation Room – Sierra Leone (WSRSL) and Njala University students. For instance, WSRSL, a women-led approach to preventing and reducing violence during the electoral cycle, was committed to reducing cases of violence, particularly sexual violence, and increasing the

1 <http://www.openelectiondata.net/en/guide/principles/>.

number of women in electoral processes. Even after the election, discussions among the group continued to be open and transparent, with little or no indication of censorship. Tamba Lamin, an experienced Sierra Leonian business analyst, site builder, trainer, content manager and passionate supporter of open data, described the group as a place, 'where we talk publicly about what has happened and share what we feel'. All parties are fully aware of each other's activities and of SLOEDP's efforts to make the data public, including converting PDFs to machine-readable data.

A major example of open data driving public recordkeeping involved the availability of the list of candidates running for office. The list was not publicly available online. According to SLOEDP, the NEC had advised, via Twitter, that the majority of the population was illiterate and did not have access to the internet. The more effective approach, the NEC advised, would be to post the list on a wall at polling stations. However, SLOEDP, the civil society platform, then took the initiative to submit what is believed to be the first freedom of information (FOI) request ever made in Sierra Leone to discover NEC's election records.

When the response was slow, Tamba Lamin went to Twitter and asked a CNN reporter to raise the issue with the NEC directly. Two days later, following a series of publicly viewed tweets, the NEC did respond to SLOEDP's FOI request with the count and voter roll of each of the stations. SLOEDP posted the information on its own website and shared it on Twitter as well as through shared WhatsApp groups. Shortly afterward, the NEC posted the information on its own website.

To effectively monitor the election, the team at SLOEDP introduced the ingenious process of training and paying motorcyclists to take photographs of each polling station and post them via multiple WhatsApp groups. Managing 16 WhatsApp groups, one for each district, it aggregated more than 10,000 snapshots of the actual results of each station. SLOEDP's methodology and capacity meant that it could cover more than presidential elections and could, moreover, produce results within 24–48 hours, as opposed to the week that the NEC required. The team at SLOEDP suspected that the NEC favoured a single political party, and that NEW (National Election Watch), which should have been an independent representative of civil society needs, had aligned too closely with the NEC. On election day, the NEC launched a new website with most of its historical content gone.

The value of public spaces for discourse on contentious issues cannot be overstated. Sierra Leone's Open Data Collaboratives WhatsApp group is immensely popular. It has reached its capacity of 250 persons, with many more waiting to join. The forum is a true marketplace of ideas, which draws together open data players as one forum to explore civic issues and how to solve them.

This group has been the most engaged and active of Sierra Leone's WhatsApp groups, with the most robust discussions. The smaller WhatsApp groups, which are focused on different geographic divisions within Sierra Leone, have also been vital in providing support, answering questions and facilitating coordination among the group monitors. SLOEDP has made its platform and methodology available for anyone to use via an open source licence, which has meant that anyone can use the platform, provided that they make their findings and any improvements and modifications publicly available.

In 2018, 23 elections were held across Africa. Similar efforts and initiatives for civil society-driven monitoring and data collection to help in election monitoring have occurred across the continent, from Nigeria, to Tunisia, to Kenya and more.<sup>2</sup> With a greater commitment to open data, ideally driven by government but with leadership and initiative from civil society, historical records of elections can be produced to help support fair elections and better systems for running them. With election monitoring more transparent, more public involvement during the election cycle, greater knowledge about candidates and issues, and larger voting turnout, more peaceful, trustworthy outcomes should become the norm.

### *The potential records management contribution*

SLOEDP has had significant success as a grassroots initiative emerging from citizens' efforts, and WhatsApp groups have flourished because individuals and various interest groups have seen the value of collaboration and used available technology to support their common communication objectives. They have tended to assume that the elections data and the WhatsApp communications can be trusted because they were designed and managed by individuals and groups with a stake in their success; the high level of trust in the quality and integrity of elections data makes sense in relation to the existing data. What is open to question, however, is the extent to which such a high level of trust can be sustained through the long term. It will be important for those managing the applications to be able to demonstrate, not just now, but at any given point in the future, that the data generated and used in an open data application such as that developed by SLOEDP can be trusted.

Examples of questions that can be addressed from a records management perspective are:

- trust in the government's portal was eroded considerably and irrevocably when citizens discovered the NEC's portal was down and the data had not been updated. Have steps been taken to ensure this doesn't happen in the case of SLOEDP?

2 <http://www.eisa.org.za/calendar2018.php> and <http://www.ifes.org/news/elections-watch-2018>.

- is the new citizen-driven portal based on generally accepted standards that focus on ensuring the quality and integrity of the data and the processes supporting the portal? Can the quality and integrity of the data be sustained through time? Do the standards address the kinds of records that will be needed to document the data and supporting processes so that evidence of data quality and integrity can be assessed through time?
- what steps will those managing SLOEDP take to address the National Democratic Institute's principle on 'permanently available'? How will the elections data and the records documenting their characteristics be preserved in an accessible manner through time in spite of changes to the technology? What policies and standards will be needed to preserve the WhatsApp communications given that they will provide an important resource for future research?
- how do the SLOEDP data relate to the data generated by the NEC? Are records in place to document the relationship so that future users will be able to discern the difference?
- what training will be needed for SLOEDP volunteers and organisations to equip them to manage the quality and integrity of the data and records effectively through time?
- are records in place to document the methods used to capture, organise and maintain the photos that SLOEDP volunteers have taken and to demonstrate their quality and integrity?
- are governance structures in place to ensure that accountability is assigned for the quality and integrity of the data generated and collected by the WhatsApp groups and by SLOEDP?

The answers to these and related questions will help guide what needs to be in place to ensure that the data are sustainable through time. The availability of authentic, complete and accurate records that can serve as evidence of the quality and integrity of the data will be fundamental to the answers.

## **Lower Mekong, Cambodia: land investment mapping**

### *The open data initiative*

As in Sierra Leone, Cambodia's approach to improving the availability of its public information illustrates how progress can be driven by an ambitious civil society. Open Development Cambodia (ODC) established a website in 2011,<sup>3</sup> to compile as many public resources as possible about government activities and international organisations that contribute to Cambodia's development.

3 <http://www.opendevdevelopmentcambodia.net>.

The site pulls information from academia, newsrooms, the private sector, local and international non-government organisations, and government resources. As more and more academics and international institutions cite ODC for research and for work, pressure has increased on the Cambodia government to provide more consistent information on its own activities. When applied to a targeted issue where there is a recognised need for accountability, and when these documents are open and available to the public, the effect is transformational, as has been the case with land investment information across the Lower Mekong.

The Lower Mekong Region, comprising Cambodia, Myanmar, Lao PDR, Vietnam and Thailand, has been experiencing rapid and unfettered development that is transforming the region and these countries' economies, while fundamentally changing the region's environmental landscape. The majority of the population is composed of smallholder farmers, fishermen and Indigenous forest communities who depend for their livelihoods on the Mekong River, the adjacent land and the rich natural resources of the Mekong ecosystem. The governments in the area rely heavily on an economic development model that depends on exploiting the land and natural resources for economic gain, which places economic development at odds with the local communities as they lose access to their traditional resources. The situation has been exacerbated by land tenure systems that are in transition from customary and communal use based on possession rights to various titling schemes in the different countries. These fluctuations make the poor especially vulnerable. Civil conflict in some of the countries has added a layer of complexity to the already fragile institutional and social framework that supports land-focused development.

When the ODC platform was launched in 2011, there was poor access to information, and ineffective public participation processes, where they existed at all, had intensified the situation. Publicly available data relevant to development were limited and difficult to access or to track systematically, which created difficulties, both for citizens and for the decision-makers themselves. Even in countries where some data on economic development were available, they tended to be generated and controlled by the governments, donors or the private sector, so that decision-making was not transparent. The result was rapid environmental changes with significant implications for both local communities and biodiversity.

Today, the initiative has six sites: five national level sites, one each in Cambodia, Lao PDR, Myanmar, Thailand, Vietnam, and a regional level site for the Mekong. The Open Development platform aggregates, organises and presents a wide range of information, while conforming to open data principles.<sup>4</sup> The

4 <http://www.opendevdevelopmentcambodia.net>; <http://www.openelectiondata.net/en/guide/principles/>.



data's usefulness is enhanced through maps, infographics and other visualisations and by being juxtaposed with related data. The Open Development platform has fostered an increase in public demand for information and has influenced the governments to provide it. In part, this has been due to the platform's objectivity. Data presented by advocacy groups have tended to be perceived as biased and to be discredited. The Open Development platform, however, provides the necessary combination of content, training and infrastructure to engender credibility. Objective information is presented by recognised and impartial sources, including governments in the Mekong Region, despite the fact that historically they have impeded access to information.

The ODC platform has targeted a wide public audience with the aim of developing greater awareness about the work of the Cambodia government and the actions of the international community invested in Cambodia. There has been a real effort to present the information in a way that anyone can access and understand, so that people can become more engaged in decisions that affect their own lives and welfare. ODC, which is the most mature site of the platform, provides good examples of how the open data initiative has functioned in practice. It has worked with a variety of stakeholders, including the government and NGOs, to pioneer open access to data on economic land concessions, mining and hydropower in the country.

The usefulness of these data, which are all associated with natural resource development contracts, has been enhanced by the way they have been presented. For instance, census data have been displayed across a period of years against a map showing the locations of economic land concessions. This has allowed users to see where local communities have declined or disappeared in relation to economic development. ODC has also reached out to local communities, journalists, university students and human rights NGOs and has trained them in how to use the datasets on the site. Through these activities, it has received valuable feedback from users on how to improve the site's usability and on new datasets that could be relevant to natural resource development, for instance data on environmental protection.

Data sourced from the government are presented in the same way as all other data presented on the site – openly and with context. Some of ODC's followers include government technocrats, who, in the past, often had limited access to information. ODC allows them to see how their plans relate to one another, not infrequently across siloed ministries with related goals. The site encourages ministries to be more forthcoming in sharing their data. For example, Cambodia's Ministry of Agriculture, Forestry and Fisheries (MAFF) increased its online information on economic land concessions from a few dozen to almost 100 after ODC and others set an example by publishing wider datasets.

ODC seeks to present data with context but without editorial comment. The intention is not to support analysis and opinions, but rather to provide resources for the public, for data journalists or for experts on particular issues so that they can provide their own perspectives and draw their own conclusions. Data are shared, whether sourced from the government or elsewhere, and are perceived as credible and objective across sectors, without being provocative or biased. The website has remained available online, regardless of changes in government or policy.

ODC has developed resources to make the data easier to understand and more attractive to a wider audience. For example, it developed an animated map showing the rate of the decrease in forest cover in Cambodia through time. The data used were not new, but the method of presentation, in the Cambodian national language, Khmer, as well as in English, was. It shows clearly the discrepancy between policy and reality. The launch of the map was covered by two major Cambodian newspapers, and within weeks, it reached almost 2,000 users, with almost a third of them able to access it in Khmer. The release of this information triggered action by local, national and international organisations, and as a result, the government was required to account for its decision-making. Eventually it began working with ODC to create an updated forest cover map.

### *The potential for a records management contribution*

As was the case in Sierra Leone, the ODC platform was developed in response to shortcomings in the government's ability to provide easy and timely access to information that had real value to citizens – in this case, environmental data for the Mekong Region. The ODC's success is reflected in the steady growth of its holdings, the high-quality and highly relevant data that it collects, maintains and makes available, and the increasingly diverse audience of users and contributors.

Nevertheless, ensuring the quality and integrity of the data and being able to prove their trustworthiness through time will inevitably present challenges, especially given the growing diversity in the types of data being collected, the potential for mashing up data from related and diverse sources, and the increase in the number of organisations participating in the initiative. In the future, there are likely to be questions about the ODC platform's ability to demonstrate through time, the quality and integrity of the data as well as the processes for supporting data collection, use and maintenance. The kinds of records management questions that could helpfully be raised are:

- how can the Open Data platform demonstrate through time that it is able to present objective information by recognised and impartial sources? What policies, standards, practices are required to support the

quality and integrity of ODC data now and in the future, and what records need to be in place to provide evidence of the level of objectivity and impartiality through the long term?

- one of the open data principles is 'permanently available'.<sup>5</sup> How will the data managed on the Open Data platform and the records documenting their characteristics be preserved in an accessible manner through time in spite of changes to the technology?
- disseminating data beyond a core audience, as for instance to the MAFF, requires that the data can be understood and that their integrity can be demonstrated to the new audience. How can the level of data integrity be documented reliably?
- combining or 'mashing' data, as in the case of census data being mapped onto digital economic land concession maps, must be handled carefully since dissimilar data sets will be based on different standards. Are records in place to document how the data were assembled, manipulated, mashed up with other data and displayed?
- does the governance structure in place for the ODC platform assign accountability for the quality and integrity of the data and the records that document the data and the supporting processes?

As in the case of Sierra Leone, the answers to these and related questions will help guide what ought to be in place to ensure that the integrity of the data can be sustained through time. The availability of authentic, complete and accurate records can serve as evidence of the quality and integrity of the data as well as the processes that support its collection, processing and dissemination.

## Key issues from the two case studies

The initiatives in Sierra Leone and in Cambodia that are described in this chapter illustrate the power of open data to promote democratic principles, enhance a knowledge-based society, stimulate the economy and fight corruption. At the same time, open data initiatives are not stand-alone projects. The data used in these initiatives have often been derived from, or are based on, data generated to support the administrative and operational activities of a government agency or some other participating organisation.

In the case of elections data in Sierra Leone, the results of the election and the election rolls generally will have been produced through a defined process carried out by an organisation mandated to administer elections, such as the NEC. The process would typically involve a sequence of steps, beginning with collecting the data and proceeding through data manipulation steps to produce the election rolls and election results. Various versions of the data, such as

5 <http://www.opendevelopmentcambodia.net>; <http://www.openelectiondata.net/en/guide/principles/>.

input data, verified and cleaned data, master edited data and published data will usually have been produced as a result of the process. Some data might be retained for short periods of time, while other data, with greater significance, might be retained far longer because of their significance.

The entire process should be supported by policies, procedures and technologies set up to administer the elections and manage both the process and the various forms of data that it generates. Records should be generated throughout the process to document the data, the steps involved and any decisions about how the process and the data were managed. Records will be necessary if evidence (or proof) is required regarding how the steps were carried out and that the data are reliable.

In the case of Cambodia, and based on the way that mapping systems are typically designed and managed, the process guiding map production would involve a sequence of steps beginning with collecting, processing, verifying and manipulating mapping data collected through field observations, satellite measurements or other processes and used to produce the digital base maps and economic, demographic, environmental and other data. Subsequent steps would involve merging the data, verifying their integrity and producing a range of digital and analogue maps on a variety of media (such as web, paper, digital media) to support a range of government objectives and respond to queries made by a wide range of users.

The process would normally generate an array of diverse data files. These would generally include raw cartographic data, source economic, demographic and environmental data, merged cartographic data files, verification and edit files, and analogue and digital files or products designed for access by the public, government officials and other interested groups. Throughout the process, records should be created to provide evidence that both the data and the process were managed properly, for instance that they were properly described, classified, retained, protected and preserved. The data and the process should be supported by effective quality and integrity controls.

The examples from Sierra Leone and Cambodia illustrate the importance of records in documenting the data, the sequence of process steps generating the various forms of the data and the decisions about how the data and the process were managed through time. When well-managed, records document the entire process, provide evidence of how the various steps were carried out, make it possible to assess the quality and integrity of the data generated at each stage and identify accountability for the data generated. The evidence that they provide should make it possible to demonstrate and manage the quality and integrity of the data and processes that produce them. By serving as authoritative, trusted sources of information, records can augment an open data initiative. They are an information source in themselves and a complementary component of any open data initiative.

Establishing a comprehensive approach to managing the quality and integrity of data in open data initiatives is challenging, especially when there are multiple players and disciplines involved (for instance, open data, data management, records management). A useful starting point would be to recognise that just as human and financial assets are managed by rigorously designed management frameworks, data provided through open data initiatives are a corporate asset that needs to be managed. From an asset management perspective, this will require laws and policies, standards and practices, systems and technologies, and qualified people, all geared to ensuring the quality and integrity of the data, the records and the processes.

Is such a framework too cumbersome and bureaucratic given the relatively small scale of a given open data initiative? Would it be enough to simply address quality and integrity issues in the context of the open data initiative? The answer depends on the level of acceptable risk: if the data generated to support the online mapping or the management of elections are flawed, then the mapping application or the Open Elections Data Portal will be flawed. What consequences would this have for the initiative, for its users or for the trust between the data providers and the data users? Such flaws can go undetected and can undermine what might otherwise be a healthy trust relationship. Once trust is eroded it is difficult to bring it back. A comprehensive framework for managing the data and records generated in the context of an entire process rather than just the process supporting the open data initiative would, through time, greatly reduce the risk that consumers will not trust the data that they access or receive.

## Conclusion

There is no doubt that open data initiatives can empower communities, equipping them with knowledge of key issues that affect them. It is important that the data they provide should be accurate and trustworthy, not only in the present but that the information should remain reliable and accessible through time. Open data initiatives can serve as important catalysts for galvanising organisations to address long-standing data quality and integrity issues, presenting valuable opportunities for the data management and records management communities to work together to address not only the quality and integrity of open data but of the data generated by administrative and operational activities of the government itself.

Designing a comprehensive management framework for ensuring the quality and integrity of data and records and the processes that support an open data initiative needs considerable care. An interdisciplinary approach with common strategies can have substantial benefits for both communities and for the citizens that they serve. Ultimately, a coordinated approach can serve the dual purpose of activating public engagement to improve the use of information and protecting its quality, integrity and accessibility through time.