

Case Study

Case Study: Tobacco Economics Control Project

Cameron Neylon ‡

‡ Curtin University, Perth, Australia

Corresponding author: Cameron Neylon (cn@cameronneylon.net)

Reviewable

v1

Received: 17 Oct 2017 | Published: 19 Oct 2017

Citation: Neylon C (2017) Case Study: Tobacco Economics Control Project. Research Ideas and Outcomes 3: e21703. <https://doi.org/10.3897/rio.3.e21703>

Abstract

The Tobacco Control Economics Project is a project that seeks to gather evidence on tobacco use and economics in southern Africa. It is a project of the University of Cape Town with support from the DataFirst repository based at the University of Cape Town. Its aim is to gather data that already exists, sometimes in digital form, frequently in offline records or in some cases paper records, and bring them together as an open resource.

The project faces challenges of data gathering as well as permissions. Frequently data is or should be “available” in some form but control over the data is relinquished only unreluctantly. In many cases the legal standing of data is unclear. Many of the challenges relating to the bringing together of the data involve ascertaining what the legal standing of a dataset is or gaining permissions for its re-use.

DataFirst is a longstanding data sharing infrastructure with professional and experienced data management staff. Challenges of ensuring continued funding and maintenance are similar to those of data infrastructures globally. The infrastructure meets international standards and provides leadership to other services and platforms in this space.

Keywords

research data, data management, open data, tobacco, health data, economic data, South Africa, Southern Africa

Main Findings

DataFirst is the site with the most previous experience and best existing infrastructure amongst the case studies. The participants work within a strong culture of data sharing and best practice, including training.

- Sites of strong infrastructure are usually found to be sites of strong practice.
- Strong long term infrastructures can be built in the context of developing and transitional countries
- Provision of support, training and capacity is being funded on a project basis and is therefore fragile.
- Even in the context of strong practice a formal requirement for a data planning process can lead to new issues surfacing.

Awareness and pre-existing capacity for managing and examining data

The project contact has extensive experience of data management, data management planning, and best practice in data handling. DataFirst is a world-class facility for data management. Therefore this contributing project represented this most experienced and expert part of the pilot project. A series of versions are available in the project data package (Neylon 2017) and a final published version is available (Woolfrey 2017).

A logistical oddity is that the main contact had not previously carried produced a Data Management Plan for a project. This was because they were generally responsible for executing an existing plan or advising on their development. Similarly to the Brazilian Virtual Herbarium the infrastructure nature of DataFirst meant that standardised approaches are not always appropriate.

The development of data management plans

There were some technical issues involved in the use of the DMPAssistant tool and the project contact elected to use the UK Digital Curation Centre DMP Online Tool instead. The technical issues appeared to be to do with authentication rather than network access or bandwidth so are probably not serious. Network access did not seem to be a major issue for South African projects in contrast with other African projects.

The development of the DMP was a useful exercise in surfacing issues to do with rights in data. Several project partners who held relevant data sets had previously stated they would contribute them to the project. However the DMP exercise led to the issue of rights being raised and clarification being sought. It is worth noting that this is not an explicit part of the DMP rubric but nonetheless the planning process was fruitful in providing a structure within which issues were surfaced.

Tools and systems: Experience of use in developing world context

DataFirst is a long standing infrastructure built in the context of UCT and South Africa. It is therefore well situated to operate effectively in this context. The challenges of gathering data from other sources within Southern Africa are varied, however the project has been structured specifically to tackle this.

In general, the experience of the main contact with systems such as those being deployed here meant that they were familiar and comfortable with the systems and questions. Nonetheless as noted an authentication problem lead them to use the DCC tool rather than that provided by Portage. Small technical challenges can be problematic and the availability, and awareness, of an alternative tool potentially saved a significant amount of time. Online/offline capability is therefore useful, though not critical in this case.

Challenges of implementation and data sharing

The main challenge that emerged for data sharing for the TCEP was a lack of clarity around permissions for data use. The DMP process was helpful in driving an explicit discussion of permissions status amongst the project leads. Datasets that were assumed to be usable were discovered to have either significant restrictions or to have no explicit permissions at all.

In terms of issues specific to developing and transitional nations there was less practical awareness at government level of Open Government Data. Control over access and limiting permissions persisted even in some cases where policy implied a requirement for openness. This is by no means restricted to developing and transitional nations, however it may be more prevalent and therefore more important as a limiting factor in these contexts.

As a data infrastructure DataFirst provides an excellent platform for data sharing and faces many of the same sustainability challenges as other infrastructures. Long term sustainability is not guaranteed and much funding is on a project basis, although it notes the ongoing support of its parent institution (Macdonald et al. 2016). Many of the characteristics of DataFirst that deliver best practice to an international standard depend on the continuity of a very small staff. This creates a risk for future capacity for data management and sharing.

Changing culture and the role of policy

DataFirst has as an implicit goal the delivering of training and capacity that supports a culture of greater data sharing. In this sense policy implementation can be a support for existing activity as it strengthens the motivation of external grantees to engage with the

infrastructure. Where data management planning, data deposition, or sharing are mandated there will be a motivation for local researchers to engage with local provision.

Policy in the context of DataFirst would support local action and be compatible with other policies that user groups might be subject to. Provided that policy imposition is associated with both information on support that DataFirst or similar groups can provide, and that DataFirst has the capacity to deliver that support, there is an opportunity to support culture change. The key issue is capacity and compatibility of policy frameworks.

Grant title

Exploring the opportunities and challenges of implementing open research strategies within development institutions (Neylon and Chan 2016).

References

- Macdonald S, Dillo I, Lafferty-Hess S, Woolfrey L, Vardigan M (2016) Demonstrating Repository Trustworthiness through the Data Seal of Approval. IASSIST Quarterly 40 (3): 6-13. URL: http://www.iassistdata.org/sites/default/files/vol_40-3_6_13.pdf
- Neylon C (2017) Dataset For Idrc Project: Exploring The Opportunities And Challenges Of Implementing Open Research Strategies Within Development Institutions. International Development Research Center. Zenodo <https://doi.org/10.5281/ZENODO.844393>
- Neylon C, Chan L (2016) Exploring the opportunities and challenges of implementing open research strategies within development institutions. Research Ideas and Outcomes 2: e8880. <https://doi.org/10.3897/rio.2.e8880>
- Woolfrey L (2017) Data Management Plan: Opening access to economic data to prevent tobacco related diseases in Africa. Research Ideas and Outcomes 3: e14837. <https://doi.org/10.3897/rio.3.e14837>