

Case Study

Case Study: HarassMap

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: HarassMap. Research Ideas and Outcomes 3: e21702.
<https://doi.org/10.3897/rio.3.e21702>

Abstract

HarassMap is an NGO based in Cairo that collects and maps crowdsourced data on sexual harassment in Egypt. Alongside this crowd-sourced data gathering it also offers training, workshops and advocacy programs, working with relevant parties to reduce the acceptability of all forms of sexual harassment.

The project has been running since 2010 based on the Ushaidi platform. Over this time it has collected a very large number of mapped events reported largely by anonymous members of the public. The data has value both in terms of its richness; mapping data, category of harassment and descriptions are all recorded; and also as a longitudinal dataset that can inform on the success of interventions as well as the development of new forms of harassment. The project has been approached in the past by a number of researchers interested in using the data it has collected.

The interest from HarassMap in Pilot Project participation was originally to obtain technical support to address how best to share data. While some technical advice was offered the focus on practice and planning was still useful. Identifying what data resources the project had, and in what form, allowed them to develop an online portal through which data can be made available to researchers on request.

Keywords

harassment, data management, open data, data sharing, Egypt, harassmap

Main Findings

[HarassMap](#) had a well structured data set and an interest in making it available. The main blockage was a way of monitoring and managing use, which is of concern for an advocacy organisation, and being able to manage the overhead. Both elements tended towards the conclusion of a managed access local data store that could be made available for researchers. The different in culture between an advocacy NGO and a research performing organisation such as a university was a subtle but consistent issue throughout.

- Policy design and implementation need to consider the different types of organisations that a funder supports.
- Projects express a wish to retain control over access to and use of the data they produce. This is a stronger impulse for advocacy-focussed projects where the agenda of a downstream user may be opposed to that of the organisation.
- Limited technical capacity and experience creates challenges for selecting, designing and implementing data sharing platforms and tools.

Awareness and pre-existing capacity for managing and examining data

The project team at HarassMap utilises an existing open source platform, Ushaidi (<https://www.ushahidi.com/>) to collect and manage data. This means that the data was structured to begin with. However there was limited experience with data management planning or consideration of archiving and preservation. As a project not based in a university or research institution there were differences in experience and culture, particularly where requests for data access came from research institutions.

There was also limited technical experience in managing online systems for data sharing and management. This was the initial motivation for engagement with the Pilot Project. This made it challenging to make choices on platforms and systems for data sharing and availability. In common with Derechos Digitales the engagement with open culture and open source movements is also associated with the application of Creative Commons licenses, with the apparent intent to cover data usage as well. Given the limitations for CC in application to data (Creative Commons Wiki Contributors 2013) this points to potential benefits for guidance on choices for data licensing.

The development of data management plans

The process of developing the DMP took some time. This was in part due to a range of people from the NGO being involved in the process and also in part due to the realisation of the expanding set of digital objects that perhaps should be covered. The local inexperience was also a contributing factor.

The DMP itself focuses mainly on replication and availability via a local server (Wael 2017). This enables local control over the data, addressing concerns with potential misuse, but also raises issues of long term preservation and continuity.

Tools and systems: Experience of use in developing world context

The online tools were adequate for the task of generating a DMP. It was noted that the questions were confusing, and also not appropriate for a retrospective planning process (). The design of the tool for use during the grant submission process was a potential issue. There was also limited interest in updating or modifying the DMP as a result of further work or changes in plan. For instance the initial intent was to clean up and make available data from a limited time period, however after the initial cleanup it was decided to continue this for all available data.

Challenges of implementation and data sharing

There were a number of issues to navigate with respect to large scale data sharing and management. One was the need to go back and clean up existing data resources. While these had been collected there was limited thought given to re-use as opposed to updating the existing online map. The DMP process highlighted issues of ongoing management and it was stated that there was an intent to change practice to manage this cleanup process earlier in the data collection process. There was evidence of a difference in culture between the capture of data in an NGO context and the collection of data resources in a research context.

The choices over what platforms and approaches to use to share data was also a challenge. The NGO has concerns over misuse of the data and there was a strong desire to maintain control over access. This was seen as a benefit both from the aspect of control, but also as a way of being able to report on downstream users to funders. The choice was made to commission a server to make the data available. This means the NGO retains control but also places the burden of maintenance with the project. The question of central provision for data sharing and/or archival and whether it is efficient for individual projects to develop their own infrastructure is seen across the case studies. The benefits of gaining experience in technical provision in a local context need to be balanced against the risks of data loss, potential lack of local expertise for server management, and potential efficiency gains of centralised provision.

A final issue related to language. The text and audio reports of harassment are in a mixture of Arabic and English. The simple data platform that was commissioned did not include sufficient resources to enable search across both language sets. This is in some ways a minor issue as a user can use search terms for both languages. However it illustrates the challenges of discoverability for data resources in non-English languages.

Changing culture and the role of policy

HarassMap is an advocacy focused NGO with a specific agenda. This creates a context different to that which most data management and sharing policies have been developed. Differences in thinking about the data and its uses emerged throughout the pilot project. In this sense the conventional nature of a generic DMP had a role in transmitting the assumptions and culture of a research project. This may be of value for consistency from a funder perspective but may not be appropriate where there are a diverse set of grantees. Most tools and training resources for data management are built with the assumptions of conventional research funders providing grants to universities and other research performing organisations.

The project contact did note benefits from engaging with the process, including a new awareness of the scope and benefits of good data management practice. Benefits that were noted included thinking about the data management process as a whole, considering whether the crowdsourcing tools were collecting sufficient information, and considering the different data assets being produced, from raw through to cleaned and processed data.

Unusually compared to the other projects it was not felt that engaging with the DMP process at the beginning of the project would have been useful (see notes of Data Management Interview, in the project dataset, Neylon 2017). There was felt to be a need to explore the tools and data being collected before a detailed planning process would have been valuable. In part this appears to be related to the cultural differences to traditional research performing organisations noted above. An emphasis was placed on the iterative nature of the process of improvement and analysis.

The role of funder policy in this case provided a stimulus but the impulse to engage was driven mainly by an existing interest in managing sharing with other researchers. Both the HarassMap and Derechos Digitales case studies point to issues of how the diversity of IDRC grantees can be supported by policy and infrastructures.

Grant title

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

References

- Creative Commons Wiki Contributors (2013) Data. <https://wiki.creativecommons.org/wiki/Data>. Accessed on: 2017-10-10.
- 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>
- Wael R (2017) Data Management Plan: HarassMap. Research Ideas and Outcomes 3: e15133. <https://doi.org/10.3897/rio.3.e15133>