

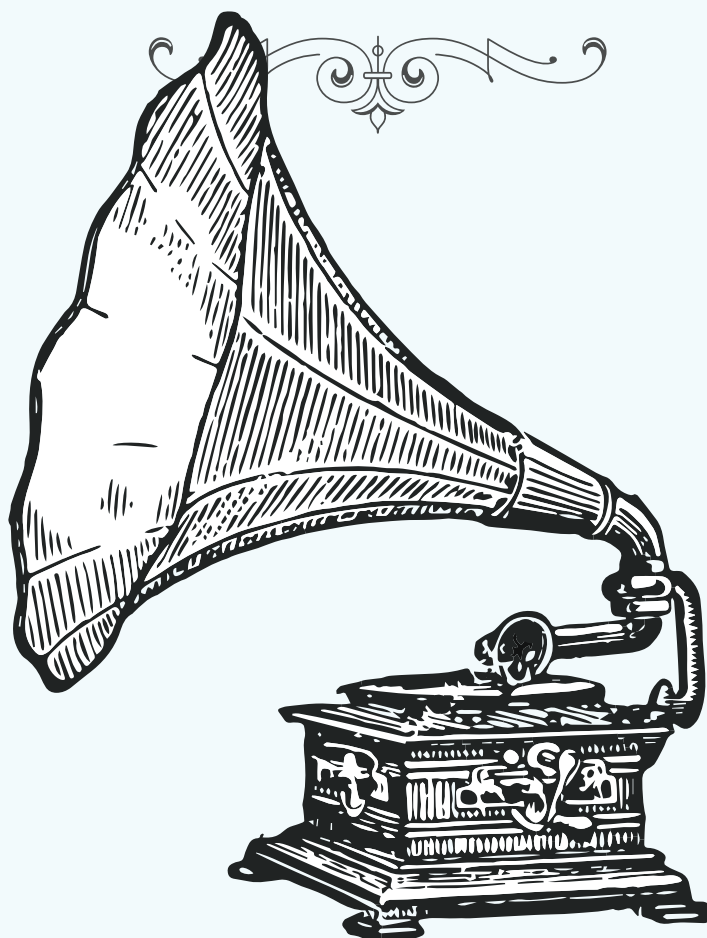
The
HAND-BOOK
OF THE MODERN DEVELOPMENT SPECIALIST



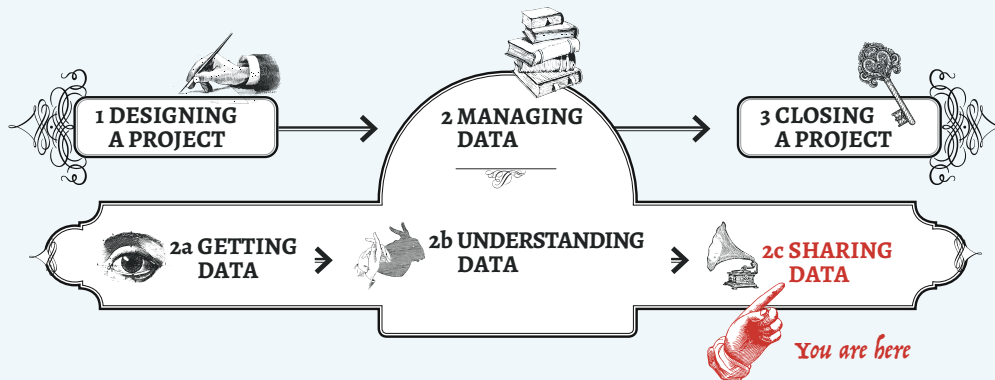
20

SHARING DATA

WHAT TO DO WITH YOUR PROCESSED DATA



THE PROJECT DATA FLOW



TARGET AUDIENCE

Those who are managing how the data will be dealt with once it has been analysed, making decisions on how it will be shared or published, the licenses that will be used, and how it might eventually get presented.

WHEN MIGHT THIS CHAPTER BE USEFUL?

When you're considering you have collected, or the data that you have analysed; is it appropriate to be made public, or shared with a smaller group of allies?

CONTENT SUMMARY

The benefits and limitations of sharing or publishing data are discussed in detail here. Being 'open by default' might be useful in some situations, but when working with sensitive data, it's better to err on the side of prudence before making absolutely sure that no sensitive or potentially harmful data is being released into the open.

The issue of *data ownership* is a tricky one; to whom does the data belong? Those who are represented in the data, or those who have collected the data? There's no easy answer to this question, so thinking it through carefully on a project-by-project basis is necessary to make sure you come to an appropriate answer.

If you do decide to share it, making sure it has an appropriate license which represents well what you want the data to be used for can help clear up any uncertainties or potential misuses of the

data. If you're sharing it with a specific group of actors, a legal or contractual framework might help to clearly outline what the data can and can't be used for by other parties, and different frameworks for establishing these principles are discussed here.

Sharing can happen on different levels; within an organisation, within a closed but controlled set of actors, or - the point of no return - freely shared online. Different pros, cons and considerations for each of these three groups are discussed here, and might be helpful if you're not totally sure which level would be most appropriate for you.

PUBLISHING DATA

If you decide to share it online, there are a number of further things to consider; do you want it to be *open* or *closed*? Chances are that if the data is suitable to be online, it's suitable to be used by others - but this isn't always the case.

If you want to make your data *open* then some tips are included here; like how to give it an *open* license, or make it fit within the *open definition*, or publishing it to the internationally-recognised transparency standard, the International Aid Transparency Initiative.

Before publishing though, you might want to make sure that individuals are taken out of the data - but this isn't always so easy. A section discussing *anonymisation* - and various strategies for *de-identification* of data are included next.

Anonymisation and identification is a particularly difficult responsible data issue, as 'perfect' anonymisation is difficult if not impossible. Releasing personally-identifiable information can be harmful to the people to whom the data pertains, but it can also be hard to remove and make sure that at no point in the future it will be re-attributable.

Some suggested strategies and techniques are offered here, but with the caveat that if you're particularly worried about identification of your data before publishing, it might be most useful to simply get help from those with more experience in anonymisation and de-identification.

PRESENTING DATA

No data is neutral, and normally when you present data you want to convey a specific narrative. Doing this in a way that doesn't skew the data or misrepresent the topics in question is important not just for the reputation of your organisation, but also as an ethical responsibility to the people who are reflected in the data itself.

This section offers a very brief overview of "do's" and "don'ts" when it comes to presenting your data, though diving into the details around ethically and responsibly visualising data lies way beyond the scope of this book. For more on this topic, we suggest referring to the accompanying suggested resources.

USEFUL RESOURCES

'License Chooser' which will help you pick the right license for you:

<https://creativecommons.org/choose/>

Frequently Asked Questions about the Open Data Commons licenses:

<http://opendatacommons.org/faq/licenses/#General>

More information on Open Data Licensing: <http://opendefinition.org/guide/data/>

More information about IATI: <http://www.aidtransparency.net/about>

The UK Anonymisation Network provides consultation and training for NGOs.

Anonymization guide produced by the UK Information Commissioner's Office

http://ico.org.uk/~media/documents/library/Data_Protection/Practical_application/anonymisation-codev2.pdf

Mushon Zer Aviv's short essay "How to lie with data visualisation":

<https://visualisingadvocacy.org/blog/disinformation-visualization-how-lie-datavis>

TacticalTech's Visualising Information for Advocacy guide:

<http://visualisingadvocacy.org>

School of Data's Visualizing Data course: <http://schoolofdata.org>



Licensed under Creative Commons Attribution-ShareAlike 4.0 International License. (CC-BY-SA 4.0)



the engine room

This publication is part of a series found at <https://responsibledata.io>, produced by the engine room Responsible Data Program, 2016.

