Data Sharing Toolkit

CABI

Module 3

Reusing data from third-party sources

Cheat Sheet

Reusing data from third-party sources

This cheat sheet gathers all the key points you should remember from eLearning on reusing data from third-party sources.

Reusing third party data

Digitally-enabled services rely on sustainable access to relevant data, provided in a way that enables integration, analysis and reuse.
Such services can require data from multiple third parties, such as geographic data, population and socioeconomic data or sample data from previous projects.

Building on top of existing data saves significant time in collecting data that already exists. Encouraging reuse can help to reduce costs and increase returns (ROI) from previous investments in data collection and platforms. To get the most benefit, the data needs to be findable, accessible, interoperable and reusable (FAIR). The most common barrier to effectively reusing third party data is a lack of clear policies and processes in place to support its sharing and reuse.

By default, the creator of anything that involves an 'intellectual effort' in its creation has exclusive rights to that creation. In order for anyone else to reuse it they must be given permission from the rights holder. Therefore when reusing or sharing data that has been collected by someone else,

it is important to ensure that permissions to access, use and share data from those you are obtaining data from are clear.

Obtaining permission to reuse third party data

It is important to have explicit **data sharing agreements** or **data licenses** to use third-party content, and to understand relevant legal restrictions or permissions. The permissions and limitations of these licenses will determine how 'open' the data is – to what extent the data can be shared.

If datasets are not open, then some negotiation may be required with the rights holder to see what can be done to ensure third party data can be used legally. This could include helping data holders to address any concerns they have related to data sharing or to set up a data sharing agreement.

The Data Sharing Toolkit contains a guide on considering data rights and permissions in investments to help with this. Other things to consider when sourcing third party data:

Compatibility with existing policies

Even when you have established what permissions you have to reuse a single dataset, existing in-country policies or funder policies (such as the Bill & Melinda Gates Foundation open access policy) might set out incompatible terms and conditions, for example requiring data to be available for commercial reuse or explicitly blocking it. Concessions may have to be sought from one or more of the parties involved.
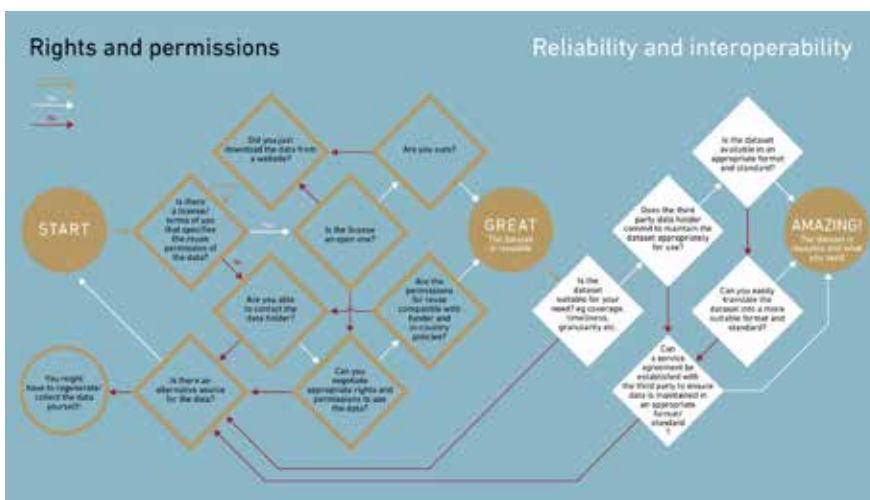
Reusability of data from third party sources

Beyond obtaining rights to reuse the data, you will need to check the practical reusability of the data. This may involve:

- Ensuring third party data is using **standards** compatible with your needs. How is the data collected, represented and exchanged?

- Checking **coverage.** Does the data actually relate to what you require?

- The **timeliness** of the data: Does it cover the right period, is it up to date?

- The **granularity** of the data. Is it as precise and inclusive of enough detail?

The Data Sharing Toolkit includes a case study on harmonising data from third party sources as well as guide on anonymisation.

This flowchart may help you consider whether you have the rights and permissions to use third party data in the way you wish.

Understanding the role of third parties within a data ecosystem

In order to ensure the appropriate mechanisms to support the access to and sharing of data are in place and enforced, it is essential to work with the stakeholders who hold data or who help shape data sharing policies and processes from early on in proposal development.

Representing data ecosystems visually can be particularly useful when contexts are complex, not well understood or not yet fully developed. There are likely to be many third party stakeholders and creating a map of a data ecosystem can help to identify the key people, the relationships between them and the different roles they play.

Where can you find guides and tools to help?

The Data Sharing Toolkit contains key resources that can help when working with data from third party sources:

- Guide: Considering rights & permissions for data in grant proposals
- Explainer: Agriculture Data Spectrum
- Case study: Harmonisation of data from different sources
- Understanding personas in agricultural ecosystems

There are a number of other external resources which may be helpful;

- Guide: Data Ecosystem mapping methodology
- You may wish to access specialist legal support

Data Sharing Toolkit

CABI Data Sharing Toolkit

ODI open data institute

BILL & MELINDA GATES *foundation*