Data Sharing Toolkit

CABI

Module 7

Ensuring sustainable access to data

Cheat Sheet

Ensuring sustainable access to data

This cheat sheet gathers all the key points from eLearning on **ensuring sustainable access to data**

Sustainable access to data is critical to ensure that data remains findable, accessible, interoperable and reusable (FAIR) and safeguarded in the long term, allowing a wider group of people to use the data over time and maximising its utility. Sustainable access also supports data ecosystems, giving organisations the confidence to invest in data as a means to develop new products, services and research.

Planning for sustainable access to data should be built in from the very beginning of a grant proposal and should involve any stakeholders that may be implicated when the investment has come to an end. It will also incur a cost, and this should be budgeted for at the start of the project.

What will ensure sustainable access?

Three key factors contribute to sustainable access to data:

- **Policies and processes** including regulations, mandates and internal management plans shape behaviour and hold publishers and data stewards accountable for maintaining sustainable practices.

  - The Foundation has its own Open access policy which requires data underlying any published research results be made accessible and open immediately. Many publishers have similar policies.

  - Data management plans ensure that all relevant aspects requirements for longer term access and data management are considered, and that resources have been put in place.

- **People** act as data stewards, seeing that the policies and plans are followed through. Assigning specific roles ensures it is clear who takes responsibility for datasets so they are accessible and usable in the long term. Roles must be suitably resourced with appropriate time and funds committed to managing and sustaining the data.

- **Technology** enables the curation and control of the data, while appropriate standards ensure data is collected and stored to support future access and use.

  - The Gates Open Research platform provides guidance on preparing data for publication and a non-exhaustive list of repositories approved by the Foundation.

Storing and managing data

There are many ways to store and manage data that will help ensure sustainable access over time. The following questions can help explore the key requirements that will affect the choices made.

What are the legal and contractual obligations?

Governments may require certain commitments or action, with most regulations now requiring a valid legal reason to retain the data. Other policies and mandates related to the investment may require commitments to the publication of data or other conditions, such as the duration for which the data will be made available.

Who are the data users and what are their goals?

Partners, funders, customers or users of a service may specify clear needs to access or use data over a longer period of time They may have their own strategic goals or legal commitments related to providing access to data. The data ecosystem mapping tool helps to map out both existing and potential reusers of data and will help define requirements for sustainability.

What is the scope of the data reuse?

The data may be purely collected and needing to be available to back up research, or it may be suitable for use in digital enabled services and so require ongoing updates and maintenance to ensure relevance. The Gates Open Research platform provides guidance on preparing data for publication and a non-exhaustive list of repositories for data.

Is the chosen solution suitable for the capability of both current and potential users?

Considering the capability of others in your ecosystem is critical to make your data as usable and as interoperable as possible. You can make sure your data can be used by others across different systems and tools by using existing standards and widely used identifiers, documenting your data effectively, and being an active member of the community.

How will sustainable access be funded?

There may be an expectation that the current data holders provide sustained access to the data, or that it will pass to a governmental body. If data is deposited or handed off to a third party, it is important to carefully consider their commitment to sustainable access to data, not only in terms of funding, but any conflicting or commercial interests that may affect access to data.

Where to store data

There are several different types of data repository, each fulfil a different need related to the sustainability of access and set of requirements:

- **Research data repositories**: Discipline-specific or interdisciplinary, broadly allowing anyone to contribute data.

- **Government data platforms**: Governments often have clauses which require anyone working with specific types of data, be it as a government department or third party organisation, to make data available via the official government platform.

- **Curated data repositories**: Provide sustainable access to carefully managed and curated sets of reference data, sustainable through country memberships and donor contributions.

- **Code and data platforms**: Offer management features such as version control, ingest pipelines and validation services.

Where can you find guides and tools to help?

The Data Sharing Toolkit contains a guide to help you identify the most suitable approach to ensuring sustainable access to data:

- Guide: Ensuring sustainable access to data

There is also further external guidance;

- Guide: Developing a data management plan
- Guide: How to create a data inventory
- Guide: Data ecosystem mapping methodology
- Gates Open Research – Data guidelines

Data Sharing Toolkit

CABI Data Sharing Toolkit

ODI open data institute

BILL & MELINDA GATES *foundation*