

Module 5

Data principles & policy within investments

Guide

Data principles & policy within investments

About this guide

This guide is for **program officers** and **grantees** that need to agree how they will access, use and share data in a multi-stakeholder environment, such as among grant delivery partners and consortia. The guide offers a practical checklist for any multi-stakeholder investments, organisations or initiatives funded by the Bill & Melinda Gates Foundation that require access, use of sharing of data. This guide incorporates best practice advice but grantees should note that other funders may require additional or alternative measures.

The meaning of the terms data principles and data policy can change depending on the context. In this guide:

- **Data principles** refer to high level, fundamental aspirations that reflect an organisation's values, and form the basis for other frameworks such as policies, practices, strategies and more.
- A **data policy** refers to the rules framework for the data activities of an organisation that align with the organisation's data principles.

When to use this guide

Start concept | request proposal | **refine proposal** | create agreement | request approval | obtain signatures | **active**

Quick links

- [Data policies help maximise utility and reduce harmful impacts](#)
- [What to include in a data policy for multi-stakeholder investments: a checklist](#)
- [A good data policy will consider the following question](#)
- [Embed and update data policies over time](#)
- [Further resources](#)

This guide is not legal advice. If you are uncertain, seek guidance from a legal professional.

Maximise utility and reduce harmful impacts

The benefit of data sharing can be significant and has been estimated to create social and economic benefits worth up to **4% of GDP**. On the other hand, when data is not used or maintained well, its value is not fully realised and the return on investment is lower. For the Bill & Melinda Gates Foundation, this means fewer benefits for people and communities who need them. That is why the foundation is committed to data sharing and transparency. Maximising the utility of data not only refers to ensuring best practice for collection and sharing, but also means ensuring there is sustainable access to data so that over time as many people as possible can use the data to make better decisions and reduce risk.

Data policies help maximise the utility of data by setting expectations about how data should be accessed, used and shared. They can guide both data best practice and its practical application. Data policies are most effective when they express the needs, opportunities and overall context of the key stakeholders in a given data ecosystem – for example grant delivery partners and the intended beneficiaries of an initiative involving data. To do this they need to be developed openly and with a wide range of inputs from all relevant groups of people.

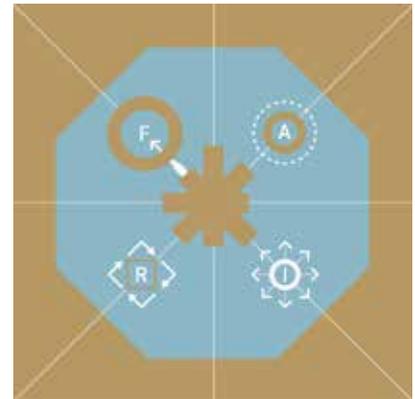
Data policies are tools which help shape how data is used at a project, organisational, regional, national or international level. They can act as checklists when people in an organisation make and monitor decisions involving data. Most importantly, they provide a common language and a shared vision for how data will be used and reused to a consortium of diverse stakeholders, as well as a framework through which to build trust and hold each other accountable.

Examples of what we mean by a data policy:

- New York City's [Open Data Policy and Technical Standards Manual](#)
- Kenya's [National ICT Policy](#)
- The Global Soil Partnership's [Soil Data Policy](#)
- The Bill and Melinda Gates Foundation's [Open Access Policy](#)

What to include: a checklist

Collaboration between actors relies on the ability to share data and information. A data policy guides how collaborators bring together data and information from multiple sources and how multiple actors access and use content. They can offer a guide for organisations with different missions and from different sectors to work together, and to help the private and public sector collaborate.



While data principles can be widely adopted, as the **FAIR principles** have been, a data policy is unique to the group it's written for. It should give a high level outline of data practices and processes, without describing specific behaviours or data release processes, which can be done separately by specialist practitioners. A policy can set out the overall approach while a data management plan (see Module 7) will provide the detail about how the policy is implemented for specific phases and stakeholders within the project.

Agreed data principles like the FAIR principles provide the first stage of development for a data policy. A data policy translates data principles into practices and processes. Data practices describe an organisation's repeated behaviours and culture. Processes set out the steps by which data is accessed, used and shared.

A good data policy will start by setting out the following:

- **The purpose of the policy** should align with the data principles that guide the use of data.
- **The scope of data** should establish the types of data used and covered by the policy.
- **The policy context** will outline relevant legislation, policy and other external guidance.

The next step is to consider the people, processes and metrics. These elements are not siloed and each informs the others. Different elements can be prioritised based on the context of the project, but consulting people, especially those impacted by the policy or the services it is shaping, is fundamental. For instance, a project directly impacting vulnerable groups should prioritise building capabilities for people in those groups in its policy, but a project providing technical infrastructure might prioritise improving a process. The success of both projects, however, will still require meaningful stakeholder engagement.

A good data policy will consider the following questions:

People

- **Who does the policy apply to?**
 - Clearly describe who the policy affects and clearly identify accountabilities of all parties.
 - Co-create the draft policy with key stakeholders and partners to ensure the affected groups are properly engaged.
- **How will the policy engage users and end beneficiaries?**
 - Consider how the internal stakeholders of the project or organisation will work with external stakeholders to help guide the appropriate release of data in useful and usable ways.
 - Guides on [mapping data ecosystems](#) and [understanding personas in agricultural data ecosystems](#) can help do this.

- **How will use of data be equitable, responsible and safeguard against harm?**
 - Evaluate how insights drawn or decisions informed by data have the potential to directly or indirectly impact people and society positively or negatively.
 - Plan how to identify and minimise the risks of harmful impacts from data.
 - The [data ethics canvas](#) may help you identify and address potential ethical issues about data in your project or organisation.
- **How will the policy protect privacy?**
 - Ensure personal data is only released when it should be and recommend steps to support this, for example privacy impact assessments or anonymisation.
 - Comply with local laws like data protection or anti-discrimination legislation.
 - When collaborating with stakeholders across jurisdictions, pay close attention to how project activities may need to adapt. The FAIR data toolbox includes a guide on creating an Agriculture data country profile to help understand the policy and regulatory frameworks to work within.
 - If there are gaps in local privacy laws, ensure privacy and other ethical issues have been considered.
- **How will the policy be enforced?**
 - Consider what organisation will be given enforcement power over the policy, and what the scope of these enforcement powers are.
 - Remind stakeholders of the positive reasons the policy exists by communicating both the benefits of complying and the ramifications of not complying.

Process

- **How will the policy make data findable?**
 - Outline a process for how data will be inventoried, reviewed and released and where these responsibilities lie across the collaborating stakeholders, including which organisations will create or steward the data.
 - Explain how the process will identify data that should not be released, like personal or sensitive data, why this is important, and how it will be audited.
 - Once you have outlined the processes and are ready to move on to more detail, refer to a checklist in Module 7 – How to create a data inventory – and the ‘Plan for data sharing, access and reuse’ section of a checklist in Module 7 – Developing a data management plan.
 - Consider how people will find the data, for example, ensuring any data that is shared is in well-structured, machine-readable formats, with clear metadata, documentation, protocols and languages that can be picked up by search engines.
- **How will the policy make data accessible?**
 - Consider who needs access – the data spectrum for agriculture in this module can help.
 - Ensure data is stored in a place that facilitates sustainable access for the time period required – for the duration of the project or beyond. Setting limits on data storage, especially personal data, reduces the risk of possible harmful impacts from a data leak or the use of outdated data.
 - The guide to ensuring sustainable access to data in the Data Sharing Toolkit and the ‘Storing and securing the data’ section of a checklist in Module 7 – Developing a data management plan.

- Consider who will have rights, or need permission to access specific aspects of the data.
- Consider use of the data when deciding about access. Different use cases will require different types of data access models and technologies – such as data trusts and APIs. Refer to the guide on deciding how to provide access to data in the FAIR data toolbox to help decide on the right type of data access model or technology.
- Ensure local stakeholders are inputting into these access requirements.
- **How will the policy make data interoperable?**
 - Consider what standards are needed to make access, use and sharing of data more likely and research whether or not these standards already exist or if there is a need to develop new ones.
 - Understand what standards for data are commonly used by collaborating stakeholders.
 - Outline your intentions to use existing standards for data that fit your purpose. If there are multiple, try to use the most open and common ones that fit your purpose.
 - Outline how you will engage with the existing standards used by the data community to support an open and trustworthy data ecosystem.
 - For more information on finding, adopting, creating and managing open standards for data, use the [Open Standards for Data](#) guidebook.

- **How will the policy make data reusable?**
 - Establish the ambition for how widely data will be shared. This should include how data will be licensed and how you will ensure that local stakeholders are consulted on this.
 - Highlight the process for agreeing intellectual property rights during data collection, and why this matters. By default, the rights to use and exploit data will normally sit with the organisation collecting it, so it needs to be clear who that is. Set out as early as possible whether these rights need to transfer to another stakeholder after the life of this investment.
 - Establish key approaches to maximise reuse, for example for aggregating and anonymising personal data so it can be shared more widely and accessed by different groups or institutes.
 - Refer to the Data Sharing Toolkit guide How to choose an open data licence and Designing data sharing agreements for more information both in this module.
 - Establish how you will engage stakeholders to ensure they are informing decisions on how widely data will be shared and how it will be licenced for reuse.

Metrics

- **What concrete commitments are being made?**
 - State specifically what the project commits to do regarding access, use and sharing of data, over the life of the policy
- **How will success be measured?**
 - List any metrics the project will use to assess the success of the policy, and say how these measures will be shared.

- See how the [design framework and exemplar metrics for FAIRness](#) suggests measuring your implementation of FAIR principles.
- **How will the policy ensure transparency?**
- Commit to regularly reviewing the policy and the processes it describes informed by feedback from stakeholders, and with clear roles and responsibilities for accountability.

Embed and update data policies over time

Policies provide a framework for ongoing discussion, to work together and hold each other accountable over time. They require continuous maintenance and should evolve over time, as lessons are learned and contexts change.

We recommend making implementing data policy part of daily work, with regular meetings, reminders and reviews, involving as many of the stakeholders as possible that contributed to or used the policy.

It is this repetition and regularity which ensures good data practices are at the core of your work.

Further resources

The following resources in the Data Sharing Toolkit may be helpful when considering what to include in a data policy:

- eLearning
 - Module 5 – Sharing data through data licensing
 - Module 7 – Ensuring sustainable access to data
- Guides
 - Agriculture Data Spectrum
 - Deciding how to provide access to data
 - How to choose an open data licence
 - Ensuring sustainable access to data

External resources:

- The Bill & Melinda Gates Foundation [Open Access Policy](#)
- ODI [Open Standards for Data](#) guidebook
- ODI [Data Ethics Canvas](#)
- CGIAR [open access and data management policy](#)

Data Sharing Toolkit



ACKNOWLEDGEMENTS

This document was authored by the Open Data Institute and CABI as part of a Bill & Melinda Gates Foundation funded investment.

The findings and conclusions contained within are those of the authors and do not necessarily reflect positions or policies of the Bill & Melinda Gates Foundation or CABI.

datasharingtoolkit.org

DOI: [10.21955/gatesopenres.1116762.1](https://doi.org/10.21955/gatesopenres.1116762.1)

cabi.org | theodi.org | gatesfoundation.org

 **CABI Data Sharing Toolkit**



BILL & MELINDA
GATES *foundation*



Except where otherwise noted, content on this site is licensed under a Creative Commons Attribution 4.0 International license.