

Module 5

Agriculture Data Spectrum

Guide

Agriculture Data Spectrum

Helping explain the language of data sharing

Grant-making institutions like the Bill & Melinda Gates Foundation aim to create the widest possible benefits from the research they fund. Open access and open data policies require organisations that receive funding to license their data, code and reports for wider reuse.

It is important for everyone involved in a funded programme to understand how these policies will impact how data is collected, used and shared. This guide helps program officers and grantees to use the Data Spectrum for agriculture when thinking about how widely to share data collected, accessed or used during an investment.

This explainer includes the following sections:

- [What is the Data Spectrum for agriculture?](#)
- [Examples of closed, shared and open data](#)
 - [Closed data](#)
 - [Shared data](#)
 - [Open data](#)
- [The Data Spectrum and FAIR data](#)
- [How to use the Data Spectrum for agriculture](#)
- [Reusing the Data Spectrum for agriculture](#)

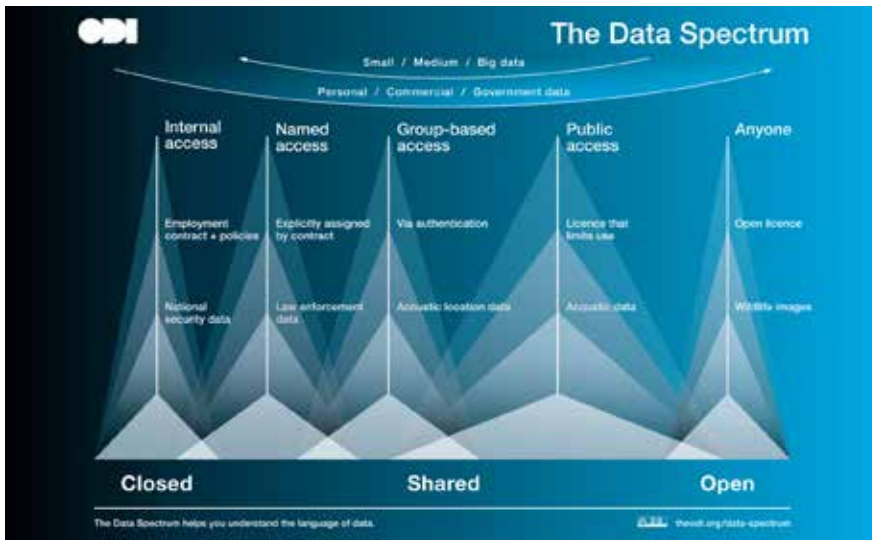


Figure 1: The Data Spectrum: Agriculture¹

What is the Data Spectrum for agriculture?

The Data Spectrum for agriculture (Figure 1) is a tool used to illustrate and understand the language of sharing data. As data moves along the Data Spectrum from left to right an increasing number of people will have access.

The Data Spectrum can be used to help establish whether data from a project is being made available to the largest audience possible, and how sharing might be improved. The Spectrum can prompt helpful questions, such as;

Is data closed when it could be shared to unlock value?

Could data that is currently shared be made more widely available?

If the data cannot be made open because it contains sensitive information about individuals or groups, can it be shared with specific organisations with appropriate safeguards in place?

¹2019, Open Data Institute, 'The Data Spectrum: Agriculture', https://drive.google.com/drive/folders/1BcO98w0e5And-RF2bq-XtHVIRRW1k7z_

Examples of closed, shared and open data

- **Closed data** is data that is held privately within an organisation, such as employment contracts and policies or sales reports.
- **Shared data** is data that is only made available to certain people or groups, such as academic researchers. Data that is shared will typically be made available for specific purposes that are defined by a data sharing agreement, such as survey data or information that is more sensitive or restricted in nature.
- **Open data** is data that is available for anyone to access, use and share. It is published under an open licence that allows it to be used for any purpose. The next section includes examples of open data.

²<https://stateofopendata.od4d.net/chapters/sectors/land-ownership.html>

Some data cannot be made open. How widely data is shared is on a spectrum, from open to shared to closed. Within these three categories there are further subtleties, particularly within ‘shared’ data. The examples below illustrate how the level of access translates to the five sub-areas – internal access, named access, group-based access, public access and anyone – indicated on the spectrum. The **examples included indicate the world as we envision it**, not the current state. For example we know land ownership data is quite closed, but if it was open there is potential to maximise the value that can be gained from it.² We have chosen the examples with the geographies within which the Foundation works in mind, namely Sub-Saharan Africa and south Asia.

Closed data

1. **Internal access.** This is data that is held privately within an organisation for internal access only.

For example: the profit and loss of an individual farm or farmer, and information related to employees contains commercially sensitive and personal data. Access to this data will be kept by the individual farm or farmer and should be restricted to avoid harmful impacts on the business or the individual.

Shared data

2. **Named access.** This is data where access is explicitly granted via a contract only.

For example: data about permits for the application of pesticides will be explicitly granted by a data sharing agreement, which is a contract between two or more parties, containing terms detailing the ways each party can access, use and share the data. Data sharing agreements often include restrictions on how widely the data can be shared.

3. **Group based access.** This is data where access is given to a specific group of individuals or organisations, such as a research group or consortia. Members of the group will be visible to each other, and individuals will need to have their identity verified through an authentication process (i.e. logging in) in order to access and use the data.

For example: The presence of pest or crop disease. Until it's officially registered with the International Pest Protection Convention (IPPC) the possible presence of a pest in a country is likely to be a very sensitive piece of information due to trade implications

e.g.crop exports. Therefore access is likely to be restricted to a smaller number of people before it has been verified.

4. **Public access.** This is data available on the web, where anyone can get access, but the rights to reuse the data are unclear. This can create risks for both the publisher and consumer.

For example: The patents of pesticides, and therefore the ingredients, are **published and available to the public**. This ensures that farmers and agricultural extension officers are able to monitor what chemicals are being applied to crops. However, the patent, and licensing accompanying it, ensure that the formula can not be reused by other manufacturers, protecting the commercial interests of the developer.

Open data

5. **Access to anyone.** This is data available for anyone to use, for any purpose, at any time. When an organisation is publishing open data, it publishes it under an open licence.

For example: Earth observation data is often used for large-scale harvesting of environmental data. **NASA's earth data** is an example of satellite data that is actively promoted for sharing and use.

The Data Spectrum and FAIR data

In order to maximise potential value, data should be findable, accessible, interoperable and reusable (FAIR) and as open as possible.

Data can be FAIR at any point on the Data Spectrum. The **FAIR data principles** emphasise clear and standardised licensing like those published by the **Creative Commons** – but they do not stipulate that data should be closed, shared or open. For instance, sensitive personal data that might only be available to the researchers collecting it, can still benefit from being FAIR, by ensuring they can easily find, access and reuse that data.

How to use the Data Spectrum for agriculture

Where data sits on the spectrum can reflect the current level of openness, the previous level, or an aspirational or potential level. It may be used to provoke discussion, examination or debate. Data can move both ways from where it currently sits on the spectrum. The spectrum can be used in multiple ways depending on your need:

1. **As an educational aid** showing people the choices about data licencing from closed, shared to open and that things like personal data or ‘big data’ can sit anywhere on the spectrum.
2. **As an advocacy tool** using the examples of closed, shared and open data in agriculture to show the type of future desired regarding access to agricultural data.
3. **As a practical tool** helping grantees make conscious and deliberate decisions about how widely to share data. Often data is shared when it could be open, or shared when it should be more closed.

Reusing the Data Spectrum for agriculture

Download the [Data Spectrum for Agriculture here](#).³

³The Data Spectrum: Agriculture,
https://drive.google.com/drive/folders/1BcO98w0e5And-RF2bq-XtHVIRRW1k7z_

The spectrum is openly licensed under the [Creative Commons-Attribution-Share-alike \(CC-BY-SA\)](#) licence.

This means you can:

- Share, copy and redistribute the material in any medium or format for any purpose.
- Adapt, remix, transform and build upon the material for any purpose, including commercially.

As long as you:

- Give appropriate credit (attribution), provide a link to the license and indicate any changes that were made.
- Sharealike if you remix, transform, or build upon the material: you must distribute your contributions under the same license as the original (CC-BY-SA).

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.1116763.1](https://doi.org/10.21955/gatesopenres.1116763.1)

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.