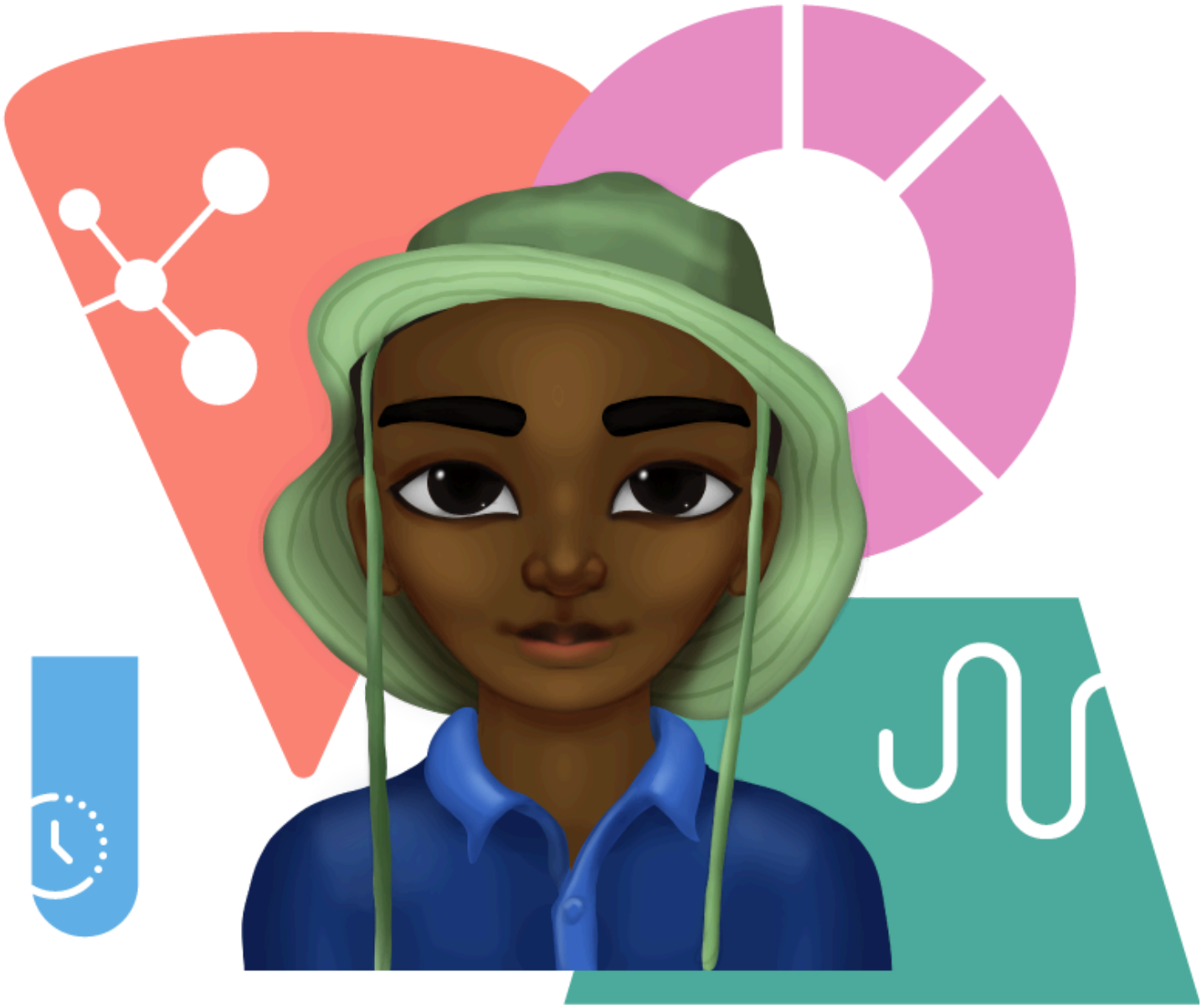


3.4 Compiling the inventory

Gathering information to start putting your inventory together and deciding how best make it accessible to others

Why should I do this?

After planning your inventory, identifying outputs and inputs, and establishing metadata and attributes, this step helps you bring everything together.

This activity provides a structured opportunity to review, refine and finalize the inventory as a shared resource for the investment. This final step also helps ensure that all relevant stakeholders are aligned and that the inventory is prepared for publication and use beyond the project.

This activity will also prepare you for developing a comprehensive data management and access plan (DMAP) later.

In this activity you will:

Consolidate the inventory by reviewing and finalizing the input and output data assets together with the metadata attributes.

Decide on the most appropriate format for your inventory, whether a simple spreadsheet or a searchable online catalog, based on the needs of the project.

Determine where and how the inventory will be published, considering factors like accessibility, security, and the potential for public access.

1) If you are a Program Officer (PO), you may want to share this page directly with your grantee, so they can act on it.

2) Use the workbook (and supporting factsheet) for Step 3 here. We recommend using the same document throughout this step, so you have a single document that captures all your workings.

3) This activity involves pulling together insights gathered from previous steps, such as data identification, contextual information, and the quality and privacy checks conducted to consolidate your data inventory.

4) Consolidate your data inventory:

Confirm completeness, and that all necessary data assets identified in Steps 3.2 and 3.3 have been included in the final inventory.

Involve key stakeholders, including data custodians and other relevant parties, in a final review to address any gaps or inaccuracies.

5) Next steps:

Decide on the publication plan

Define access levels

Plan for regular updates

Decide on the publication plan

Choose a suitable platform for publishing the inventory, considering access controls, transparency needs, and the target audience.

For small projects, a periodically updated spreadsheet may be all you need. For projects with a large volume of data, a database is more appropriate.

If one of your goals is enhancing data discovery and sharing, you could consider transforming the inventory into an online data catalog that enables searching, browsing, and linking to various data assets.

There are many online resources available for publishing your data inventory, including:

CKAN (Comprehensive Knowledge Archive Network)

Data Hub

Dataverse

More information on data catalogs is covered in Step 6.

Define access levels

Establish clear guidelines on who can access the published inventory, and any security measures required to protect sensitive information.

One way to do this is to determine different levels of access based on stakeholder roles and responsibilities. For example:

Full access can be given to data custodians and administrators who need to manage, update and audit the data inventory.

Read-only access can be given to stakeholders who only need to view and reference the inventory, such as project managers, analysts and decision-makers.

Restricted access can be given for sensitive or confidential data, where only designated individuals or teams can access certain sections or datasets based on their role and authorization.

Plan for regular updates

Decide on regular updates (e.g., quarterly) and event-based updates for key changes.

Automate where possible. Use automated prompts and scheduled checks to streamline updates.

Integrate into workflows. Make updates part of routine tasks.

---

6) Refer to the investment type examples to help you complete your plan.

Laying the groundwork for the Data Management and Access Plan (DMAP)

This activity focuses on refining and validating the data inventory, metadata and roles, serving as a practical foundation for the DMAP in Step 5.2.

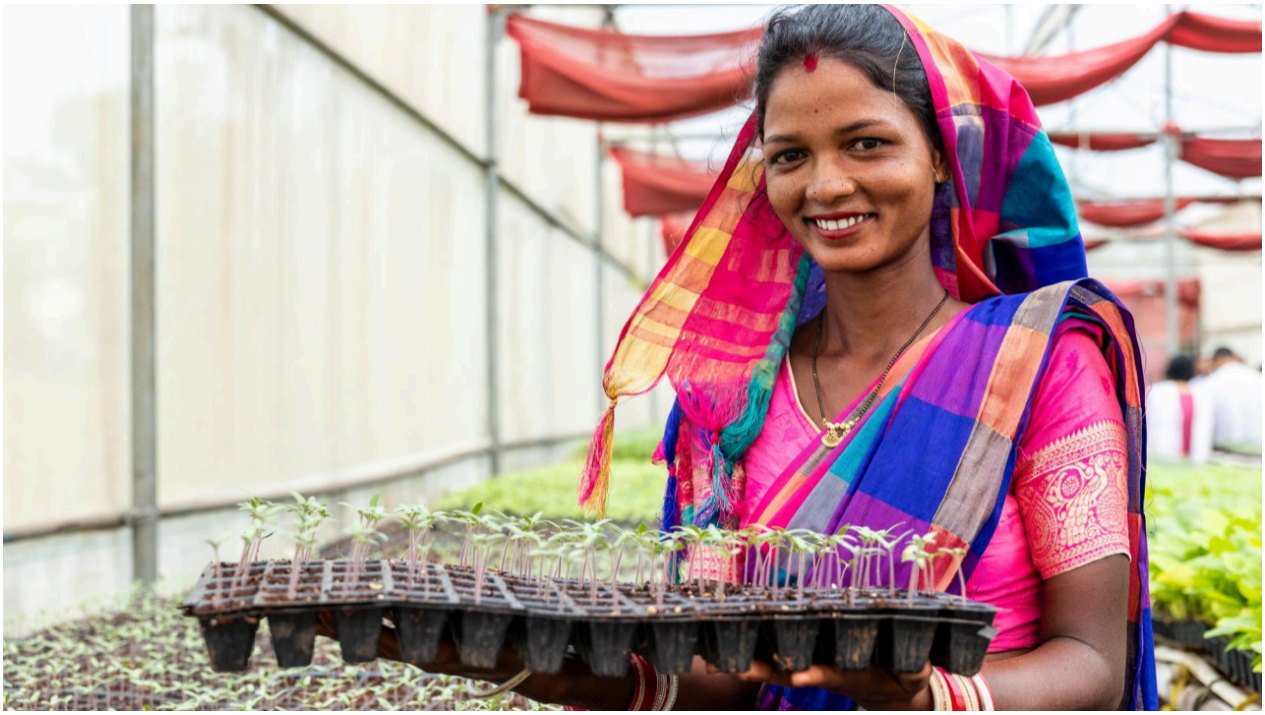
The personas identified as you were planning the inventory will be important when you formally articulate roles and responsibilities in the DMAP.

By reviewing the metadata now, the emphasis is on ensuring completeness and consistency. The DMAP will later specify long-term standards and procedures for maintaining and updating metadata.

Initial decisions made now on inventory publication will be formalized in the DMAP and include protocols for monitoring and updates.

Investment types

Overview

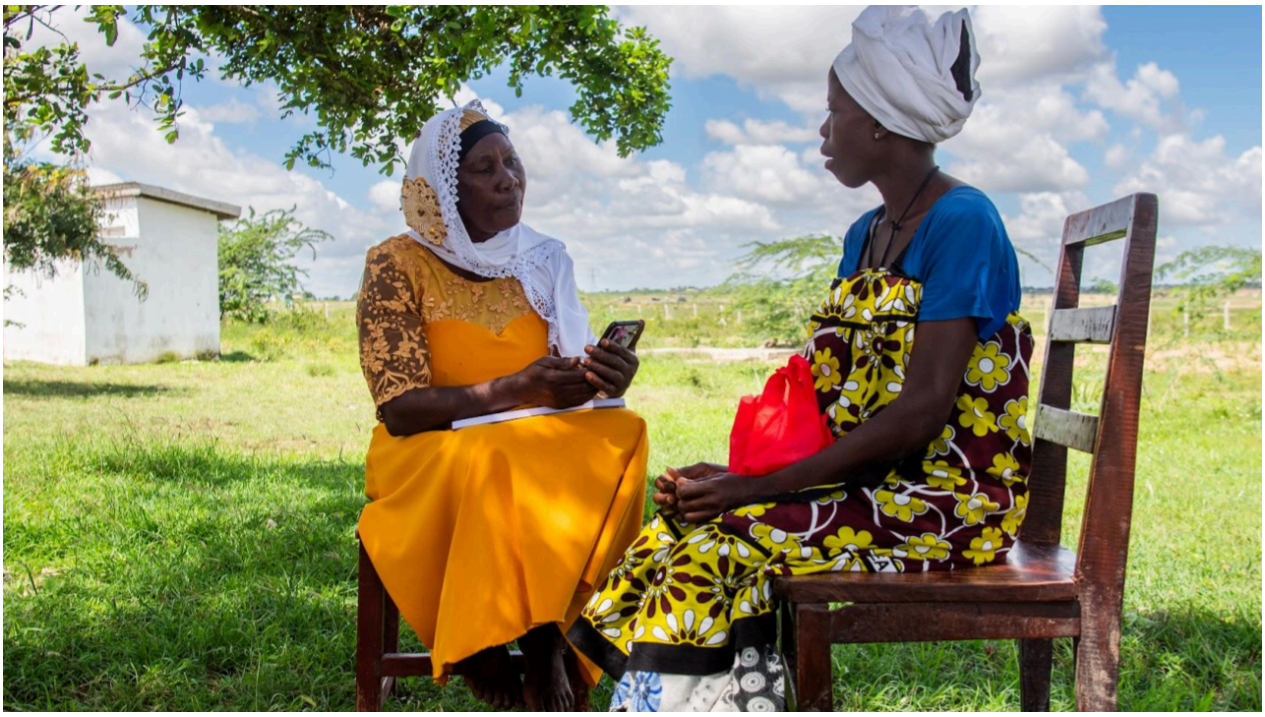


©Gates Archive/Mansi Midha

Every investment project is unique

The application of the six steps will vary accordingly. To provide examples that align with your project, common characteristics of AgDev investments were researched and three 'investment types' were developed.

AgriConnect: a digital solutions investment



©Gates Archive/Alissa Everett

AgriConnect: Finalizing and publishing a data inventory

AgriConnect is working to develop a data repository platform for smallholder farmers, focusing on access to agricultural inputs, financing options, and market insights. In Activity 3.4, AgriConnect's team consolidates their data inventory by:

Finalizing Data Assets

They review collected data on smallholder farms, soil health, crop yields, financing options, and market prices. Each dataset is assigned metadata such as data source, collection date, geographic area, and accessibility level (public or restricted).

Choosing a publishing platform

AgriConnect decides to use CKAN for its open-source capabilities and ease of customization, which suits their goal of making data easily accessible to smallholder farmers and partners.

Defining access levels

Data custodians, including project leads and partners, receive full access to manage and update

the repository, while external stakeholders and farmers have read-only access to relevant datasets. Sensitive data, such as farmers' financial information, is restricted.

Setting up regular updates

AgriConnect plans to update the inventory quarterly, integrating updates into regular data collection cycles and automating metadata prompts in CKAN for new entries.

AgroThrive: a policy and advocacy investment

©Gates Archive/Thomas Omondi

AgroThrive: Compiling a structured data inventory

AgroThrive aims to inform national agricultural policies with evidence-based data, focusing on credit, financing, infrastructure, and climate resilience. In this phase, they compile the inventory as follows:

Finalizing data assets

The team reviews data assets collected on climate resilience programs, credit and financing for smallholder farmers, and infrastructure projects. They ensure that each data asset is paired with metadata that describes its relevance to policy, data owner, collection date, and confidentiality level.

Choosing a publishing platform

AgroThrive selects DataHub, allowing easy integration of external datasets and compatibility with government systems. This supports their goal of making data accessible to government liaisons

and policy analysts.

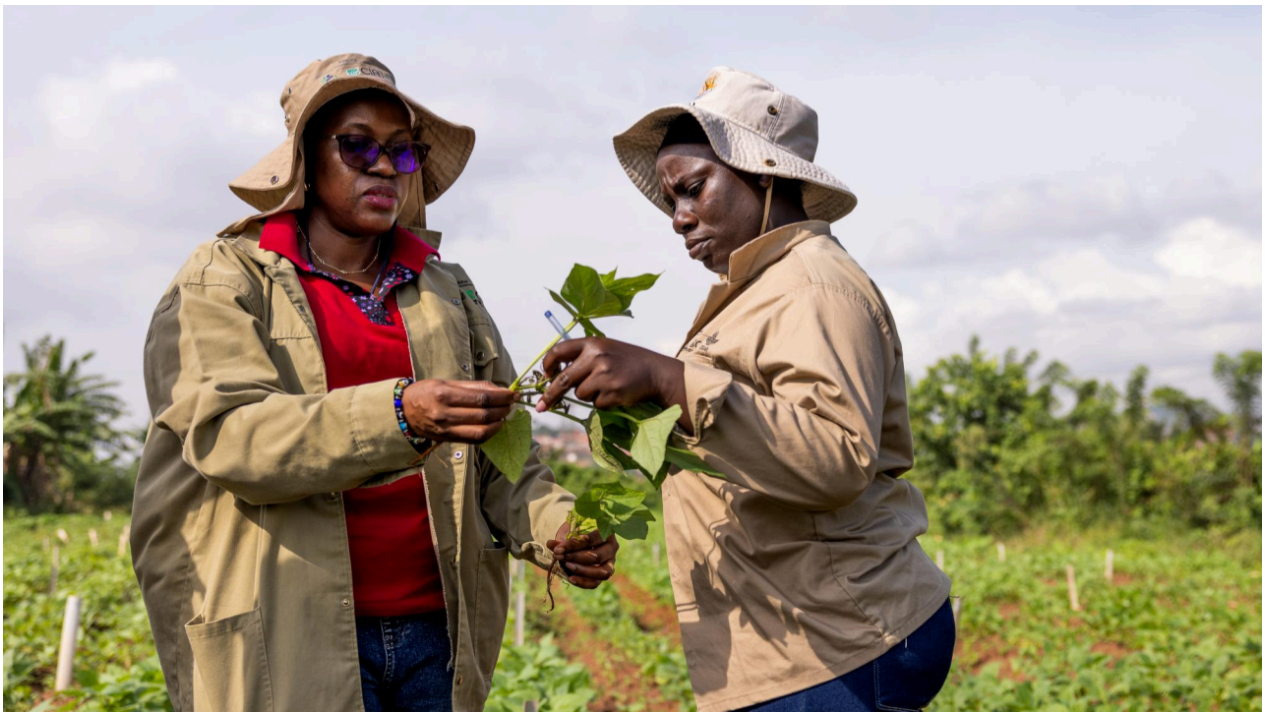
Defining access levels

Full access is granted to data custodians and administrators, while read-only access is provided to policy analysts and government liaisons. Highly sensitive data on individual smallholders' credit profiles is given restricted access to protect privacy.

Setting up regular updates

AgroThrive commits to quarterly updates and event-based updates for new infrastructure data, integrating updates with their ongoing policy analysis cycle.

NGBT: a field research investment



©Gates Archive/Esther Mbabazi

NGBT: Finalizing a structured data inventory

NGBT (Next-Gen Barley Trials) is developing a climate-resistant, nutrient-rich barley varietal to address food insecurity. Activity 3.4 for NGBT involves the following:

Finalizing data assets

The team consolidates data on barley crop trials, soil health metrics, and climate impact assessments. Metadata includes research phase, geographic trial location, season, and access level (public, partner-only, or restricted).

Choosing a publishing platform

NGBT opts for Dataverse, ideal for managing and sharing research datasets with academic

collaborators. Dataverse's integration features help them share findings responsibly with other research institutions.

Defining access levels

Research partners and lead scientists have full access, while general stakeholders, including climate scientists and agronomists, have read-only access to anonymized trial data. Sensitive genetic data is restricted to specific team members.

Setting up regular updates

NGBT schedules semi-annual updates, synchronizing them with research publication milestones and automating metadata entry for new trial data as it's uploaded.

Over a number of years working in digital agriculture and agronomy, I've seen how a lot of good evidence, collected by researchers, gets lost.

Martin Parr, Director, Data Policy & Practice, CABI

Learn more

FAIR Process Framework has been developed by the Enabling Data Access (EDA) project team at CABI and is funded by the Bill & Melinda Gates Foundation to support the foundation's Open Access Policy. The FAIR Process Framework is a tool to assist partners in developing data access and management plans (DMAPs) that incorporate FAIR and responsible data practices.