

# Opening data with trust: Reflections on a holistic approach to data governance at IFPRI

**Opening data with trust: Reflections on a holistic approach to data governance at IFPRI**  
 Indira Yerramareddy, Nilam Prasai, Erica Saito, Leigh Plummer, Mulugeta Bayeh, Ryan Miller, Xinyuan Shang and Yunchul Jung  
 International Food Policy Research Institute, Washington DC, United States

**Research data governance**

Over the years, IFPRI has improved the management of research datasets and associated with our mission to offer data, as global public goods, to the research community and policy-makers. This has been achieved through a holistic approach to data governance, involving us in the work of various international organizations. This holistic approach to data governance is based on the following principles:

- Ensure that research data is available to the research community for the research process. In recent years, this research has made significant progress in terms of data availability and quality. Data privacy, data reuse and reproducibility research, to address the challenges generated by the age of the research community, have been implemented, and the broader community (IFPRI researchers, donors, policymakers, and the broader community) IFPRI published a Data Governance Strategy 2020 to improve overall research quality and transparency, privacy, and security, and promote the reuse of research data.

**Measuring success**

IFPRI regularly tracks both of key performance indicators (KPIs) for measuring the success of its research data governance approach. Some of the key metrics are shared below:

1. Number of datasets published

Number of datasets published by IFPRI over the years

Year	Number of datasets published
2015	100
2016	150
2017	200
2018	250
2019	300
2020	350

2. Number of datasets downloaded

Number of IFPRI datasets downloaded over the years

Year	Number of datasets downloaded
2015	1000
2016	1500
2017	2000
2018	2500
2019	3000
2020	3500

3. Number of datasets cited
4. Citations attributed to publications
5. IFPRI score for publications created by IFPRI researchers

**FAIR principles for data and publications**

**Winning the trust**

Trust is an important factor for the adoption of the Open Data ecosystem through policies, tools, and education. IFPRI focuses on trust in the following ways:

- The researchers about sharing data and materials by our researchers, using standard tools
- Levels of open sharing of openness
- Concern about the inclusion of the privacy of the research subjects
- Concern about data privacy related to research data sets
- Concern that someone will publish a copy paper before the researchers are able to write and publish their data
- Concern about the individual and institutional reputation in case of a data breach

IFPRI is approaching to winning trust for sharing data to enhance the research ecosystem for sharing data. IFPRI has a clear strategy to build trust in the research ecosystem, including its policies, procedures, and system components. Some specific measures include:

- Established Research Data Governance and Management Committee

**Research data management: evolution**

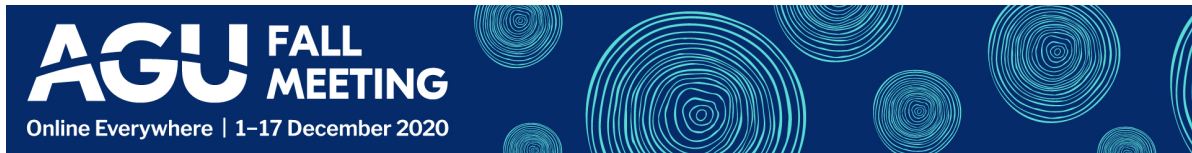
Research data management evolution in IFPRI

ABSTRACT REFERENCES CONTACT AUTHOR PRINT GET POSTER

Indira Yerramareddy, Nilam Prasai, Erica Saito, Leigh Plummer, Mulugeta Bayeh, Ryan Miller, Xinyuan Shang and Yunchul Jung

International Food Policy Research Institute, Washington DC, United States

PRESENTED AT:



## RESEARCH DATA GOVERNANCE

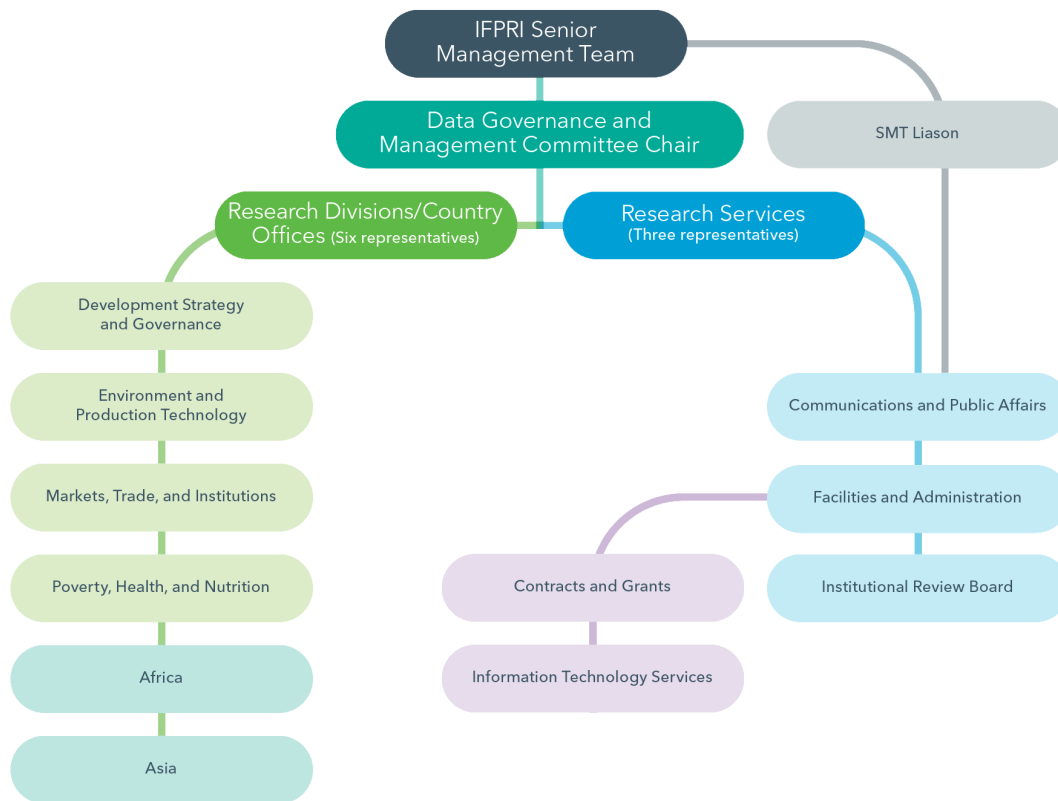
Over the years, IFPRI has improved the management of research datasets and compliance with our mandate to offer them as global public goods. IFPRI first started sharing data through CD-ROM in 1990s and developed data policy in 2000, making us an early innovator in data sharing among research institutes. However, offering data as freely and openly accessible public goods remained mostly voluntary for the research projects. In recent years, the research community has made significant progress in developing policies and standards in open data, data privacy, data security, and reproducible research. In order to meet the challenges presented by changes in the research environment from donors, competitors, and the broader community, IFPRI established a Data Governance Team (DGT) to improve overall research quality and streamline policies, processes, and protocols for managing research data. The DGT has helped to:

- Save resources and time
- Manage risk
- Address conflicting situations
- Assure the quality of data
- Comply with donor's and legal requirements
- Comply with FAIR principles (Findable, Accessible, Interoperable, and Reusable)
- Support open and reproducible research

### Research data governance structure

In 2017, IFPRI established a term-limited Data Governance Team (DGT) to review data management practices at IFPRI. The team commissioned a review by an external consultant, which suggested establishing a permanent data governance body. As a result, the institute established the Data Governance and Management Committee (DGMC). The structure of the committee is shown below.

## RESEARCH DATA GOVERNANCE STRUCTURE



## Research data governance approach

Research data management was handled mainly on three fronts:

1. Established new policies and revised existing policies to support research data management;
2. Established and improved best practices, processes, and infrastructure for research data management;
3. Strengthened capacity of research and services staff for research data management.

# WINNING THE TRUST

Trust is an important factor for the adoption of the Open Data movement alongside policies, tools, and infrastructure. IFPRI remains no exception to this.

The reservations about sharing data and materials by our researchers mainly stemmed from:

- Lack of understanding of openness;
- Concern about the violation of the privacy of the research subjects;
- Concern about others may misuse or misinterpret the data;
- Concern that someone will publish a new paper before the researchers are able to write and publish from their data;
- Concern about the individual and institutional reputation in case of a data breach.

## IFPRI's approach to winning trust for sharing data

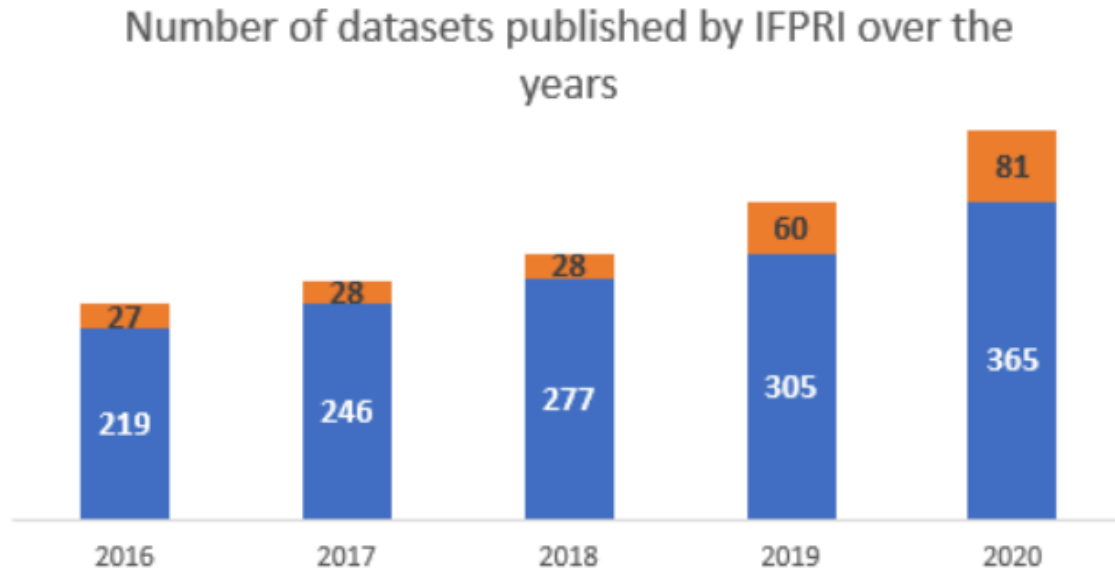
To enhance the trust of researchers for sharing data, IFPRI took a holistic approach including raising awareness about data sharing to policies, protocol, procedures, and system improvement. Some specific measures include:

- Established Research Data Governance and Management Committee with representation of researchers to allow them to share concerns and issues.
- Raised awareness with research staff about the degree of openness: "open as possible and closed as necessary."
- Incorporated research data into the institutional risk register. Including research data in the risk register led the institute to assess the key risks associated with the improper management of research data and assess the risk environment. This helped IFPRI to plan, monitor, and track the identified risk related to research data management.
- Instituted the protocol for the second review of dataset beside research team to ensure the anonymity of research subjects.
- Developed data sharing agreement for sharing confidential data and for sharing unpublished data thereby safeguarding IFPRI researchers to some extent.
- Developed "terms of use" agreement for the data repository to ensure fair use of its resources.
- Encouraged attributing authorship to data documentation, which can be included as a separate product for performance appraisal. Corporate attribution being IFPRI's established practice.
- Developed guidelines for research data management.
- Worked with researchers and addressed their needs for publishing data including publishing only minimum dataset associated with their publication or to make only metadata publicly accessible.

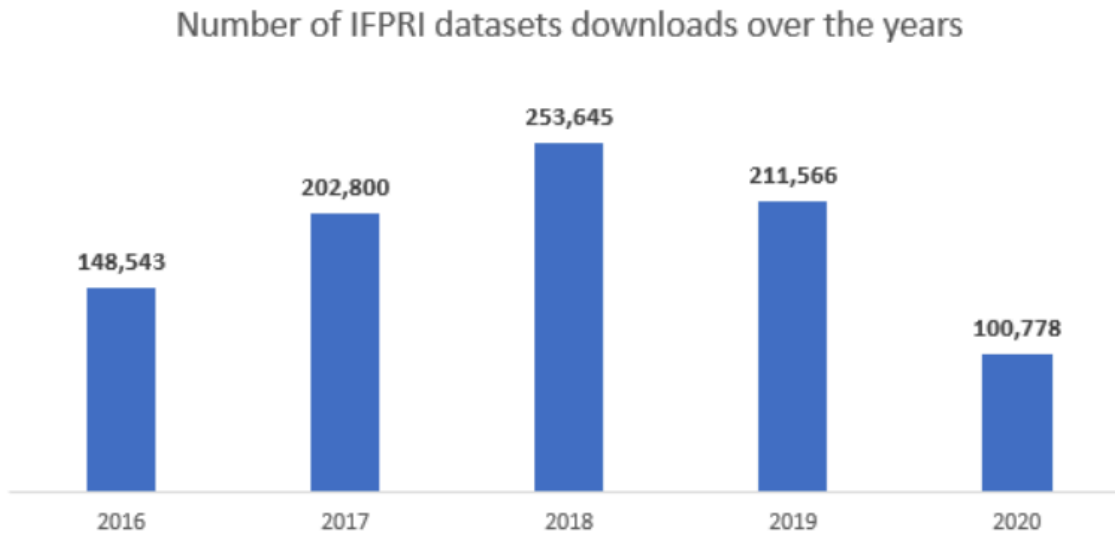
## MEASURING SUCCESS

IFPRI regularly keeps track of key performance indicators (KPIs) for measuring the success of its research data governance approach. Some of the key metrics are shared below.

### 1. Number of datasets published\*



### 2. Number of datasets downloads\*



### 3. Number of datasets cited

### 4. Datasets interlinked with publications

5. FAIR score for published datasets (provided by Global Agricultural Research Data Innovation Acceleration Network (<https://gardian.bigdata.cgiar.org>) (GARDIAN), which is the CGIAR flagship data harvester)

### 6. Number of data papers

### 7. Datasets that are open

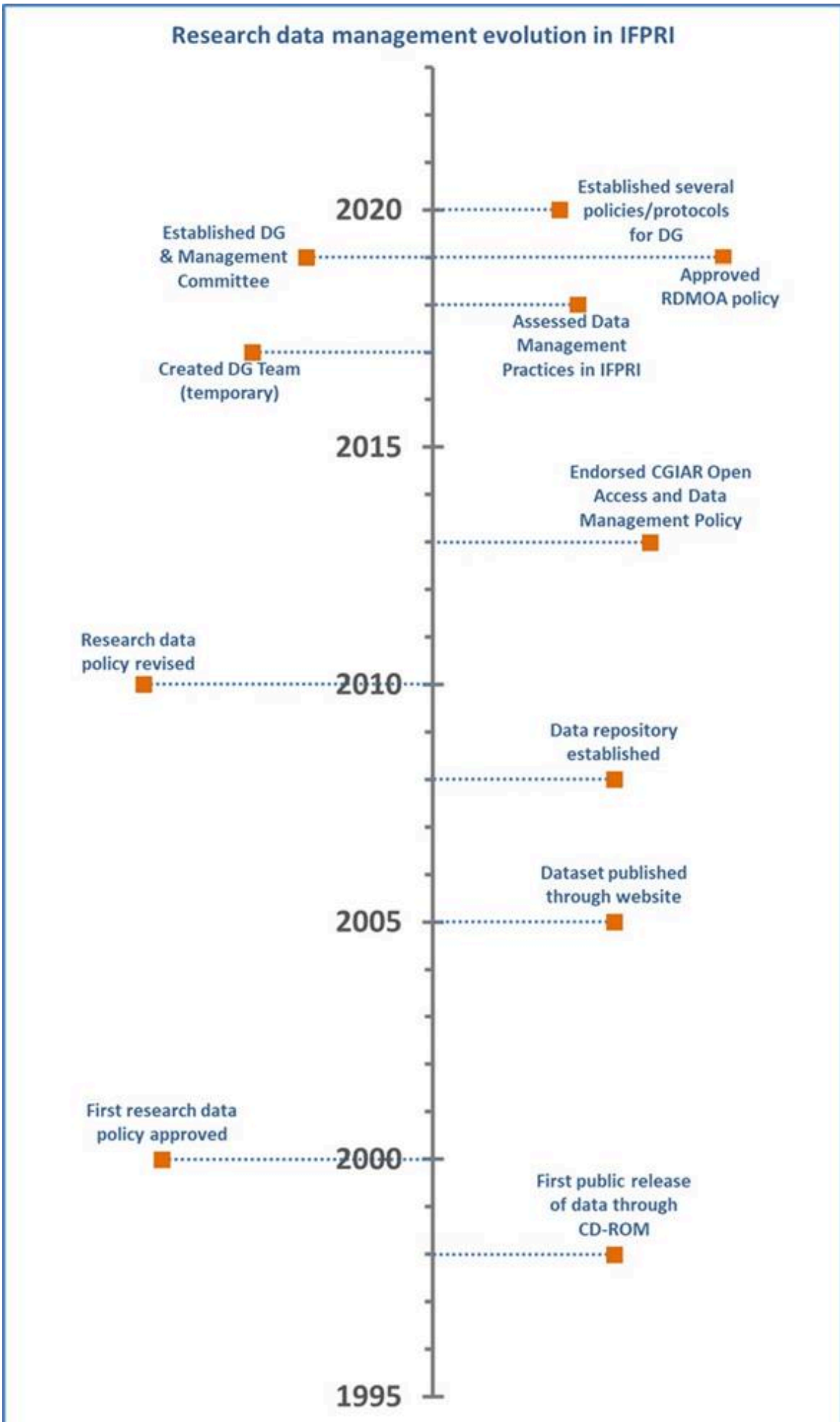
8. Datasets having minimum metadata
9. Incidence of the research data breach or privacy violation in the last three years
10. Requests for developing a data management plan
11. Number of ad hoc data requests
12. Number of staff hours responding to data requests

\*As of October 2020

# FAIR PRINCIPLES FOR DATA AND PUBLICATIONS

[VIDEO] <https://www.youtube.com/embed/PefS2Cst79E?rel=0&fs=1&modestbranding=1&rel=0&showinfo=0>

# RESEARCH DATA MANAGEMENT EVOLUTION





## ABSTRACT

IFPRI's experience building overall data governance for Research Data Management (RDM) has helped us to build trust among IFPRI staff for opening their data. We will highlight how policies, protocols and processes are equally important components for opening data in addition to infrastructure and digital repositories.

The mandate of IFPRI is to provide research-based policy solutions to sustainably reduce poverty and end hunger and malnutrition in developing countries. In pursuit of that mandate, IFPRI generates data ranging from household surveys to process-produced data and makes those data available and easily accessible. IFPRI has moved progressively toward full FAIR compliance beginning with publicly available datasets hosted on the website in 1999, a published data policy in 2000, establishing a data repository through Dataverse in 2007, and a Data Governance Team (DGT) in 2017. Revisions to the data policy in 2005 required researchers to publish the datasets generated during research projects within two years of data collection/finalization. However, the intention of the institute did not lead to automatic compliance by research staff. The data policy was not observed consistently, and publishing datasets remained voluntary. Researchers repeatedly expressed hesitations for opening data widely.

To overcome resistance from research staff, respond to open data movements, regulations, and policies from donors and governments, and identify gaps in RDM within the institute, IFPRI established a term-limited DGT in 2017. The first undertaking of this team was launching a review of the data management practices by an external consultant. Based on the recommendations from the review, IFPRI established a permanent data governance body, Data Governance and Management Committee (DGMC). DGMC facilitates the implementation of other recommendations from the review and address staff concerns. IFPRI has established new policies, processes, incentives and systems for RDM. As a result, IFPRI has 420 well-documented open datasets, and researchers are more confident and trust the system.

## REFERENCES

IFPRI Dataverse (<https://dataverse.harvard.edu/dataverse/IFPRI>)

IFPRI Research Data Management Website (<https://rdm.ifpri.info/>)

IFPRI Research Data Management Open Access (RDMOA)

(<http://ebrary.ifpri.org/utills/getfile/collection/p15738coll2/id/133308/filename/133517.pdf>)

CGIAR Open Access and Data Management (OAOD) Policy

(<https://cgspace.cgiar.org/bitstream/handle/10947/4488/Open%20Access%20Data%20Management%20Policy.pdf?sequence=1&isAllowed=y>)