



World Bank Group Policy

Development Data Quality Policy

Bank Access to Information Policy Designation

Public

Catalogue Number

SEC1.01-POL.104

Effective

October 30, 2024

Content

This document sets out key principles for quality assurance, control and improvement of Bank Development Data.

Applicable to

IBRD, IFC, MIGA, ICSID, IDA

Issuer

Executive Directors of, IBRD, IDA, IFC, MIGA

Sponsor

Senior Vice President & Chief Economist, DECVP

SECTION I – PURPOSE AND APPLICATION

- a. This Policy establishes a framework (the “**Data Quality Framework**”) that governs the production, acquisition, processing, storage and dissemination of Bank Development Data (as defined herein). Recognizing the importance of data in delivering on the Bank’s commitment to better anticipate, design, and deliver solutions and lasting impact for clients, the purpose of the Data Quality Framework is to establish a system for quality assurance, control and improvement, all with a view to ensuring that Bank Development Data can be trusted and are useful for development purposes. For the avoidance of doubt, it is noted that the Data Quality Framework does not intend to regulate Bank Corporate Data (as defined herein). In addition, this Policy aims to ensure that Third Party Data are assessed for suitability to the extent (i) disseminated by the Bank or (ii) used as an input to the production of Bank Development Data.
- b. Data quality is a relative concept in that it refers to the degree to which a Data Product’s inherent characteristics align with specified requirements designed to ensure a Data Product’s fitness for purpose in supporting intended tasks or objectives. Also, data quality is a multi-dimensional concept that comprises, among other things, the following attributes: relevance, accuracy, reliability, coherence, comparability, timeliness, accessibility, usability, interpretability, discoverability, and visibility.
- c. The purpose of this Policy is not to define a uniform standard against which to measure substantive quality of Bank Development Data. Instead, this Policy sets out a number of high level behavioral, procedural, technical and governance principles that apply to the various stages of the Data Management Lifecycle, which the Bank considers a key prerequisite for trustworthy and useful Bank Development Data. Recognizing the above-referenced relative and multidimensional nature of data quality, this Policy will be complemented by detailed procedures and practical guidance for staff involved in any stage of the Data Management Lifecycle.
- d. This Policy applies to the Bank.

SECTION II – DEFINITIONS

As used in this Policy, the following capitalized terms or acronyms have the meanings set out below:

1. **Access to Information Policy**: in the case of IBRD and IDA, the Policy governing public accessibility of information in the Bank’s possession, bearing catalogue number EXC4.01-POL.01 (as updated, amended, or restated from time to time), and in the case of MIGA, the Access to Information Policy of the Multilateral Investment Guarantee Agency (as updated, amended, or restated from time to time), and in the case of IFC, the International Finance Corporation Access to Information Policy originally dated January 1, 2012 (as updated, amended, or restated from time to time), as applicable.
2. **API or Application Programming Interface**: a form of software interface that allows an application to access the features or data of, and interact with an IT operating system or application.

3. **Artificial Intelligence:** a machine-based system that, for explicit or implicit objectives, infers, from the input it receives, how to generate outputs such as predictions, content, recommendations, or decisions that can influence physical or virtual environments.
4. **Bank:** IBRD, IDA, IFC, MIGA and ICSID.
5. **Bank Corporate Data:** any Data Product that strictly pertains to the internal organization, business functions and administration of the Bank.
6. **Bank Development Data:** Any Data Product that (i) is not Third Party Data, and (ii) is produced or disseminated by the Bank or by the Bank in collaboration with a third party for the purpose of facilitating development or otherwise relating to any development activity of the Bank (which for the avoidance of doubt excludes Bank Corporate Data).
7. **Chief Statistician:** the Chief Statistician of the Bank.
8. **Data Curation:** the process of (i) acquisition or collection, (ii) verification and processing, (iii) protection, documentation and packaging, and (iv) storage, preservation, maintenance, archiving or disposal of a Data Product (collectively).
9. **Data Management Lifecycle:** the various steps that data may go through in a typical statistical business process, which may occur in different orders (i.e. linear or circular) and can be broken down as follows: (i) acquisition or collection, (ii) verification and processing, (iii) protection, documentation and packaging, (iv) storage, preservation, maintenance, archiving or disposal, (v) dissemination, (vi) evaluation (including usage monitoring).
10. **Data Product:** any numerical data (whether qualitative or quantitative) provided in whatever form.
11. **Data Process:** a process designed to facilitate any stage of the Data Management Lifecycle.
12. **Data Quality Framework:** has the meaning ascribed thereto in paragraph (a), Section I hereof.
13. **Data Set:** an organized collection of Bank Development Data that is either (a) resulting from the same data collection or production activity, or (b) organized as a single package, and stored in a machine-readable format (including, without limitation, time series and indicators databases, survey or census microdata files, sensor readings, transaction records, satellite imagery, and geographic data files).
14. **Data Stewardship Network:** a network comprising Staff in each relevant Unit, tasked to support the development, promotion, implementation and continued delivery by such Unit of the Data Quality Framework.
15. **Metadata:** information that describes, explains, locates or otherwise facilitates the retrieval, use or management of a Data Product.
16. **Metadata Standards:** a set of rules, guidelines, conventions and concepts designed to ensure that Metadata are structured, comprehensive and machine-readable, and that otherwise facilitates usability of Metadata.

- 17. **Statistical and Data Management Standards:** a set of rules, principles and procedures for the production (including, among other things, data collection), processing and dissemination of data, and designed to achieve the uniform treatment of data.
- 18. **Third Party Data:** a Data Product that has not been produced by or on behalf of the Bank.
- 19. **Unit:** an organizational unit of the Bank.

SECTION III – SCOPE

1. Data Quality Principles

a. Impartiality and Independence

Bank Development Data shall be produced professionally, impartially and free from inappropriate influence.

b. Responsible Data Management

The Bank shall ensure that its activities that are part of the Data Management Life Cycle are designed and conducted responsibly, taking into account potential ethical, legal and information technology risks.

c. Verifiability

Bank Development Data shall be verifiable. The Bank shall ensure that, among other things, its Data Processes are designed and documented accordingly.

d. Transparency

The Bank shall openly and promptly communicate on any identified material quality issues with respect to Bank Development Data.

e. Relevance

The Bank aims to achieve that all Bank Development Data are relevant to likely end users. The Bank implements consultation and feedback mechanisms to (i) understand specific needs and interests for Bank Development Data, and (ii) prioritize or adapt the Bank’s data activities and allocate resources accordingly.

f. Coherence and Comparability

Bank Development Data shall aim to incorporate standard concepts, definitions, classifications, and methods that enable (i) coherence, and (ii) comparability over time and between geographies.

g. Statistical and Data Management Standards

- a. The Bank shall for purposes of the production of Bank Development Data apply statistical methods that (a) comply with Statistical and Data Management Standards (to the extent applicable), or (b) otherwise ensure compliance with this Policy.
- b. The Bank shall continuously improve Statistical and Data Management Standards applied by the Bank. For such purpose, the Bank collaborates with international, regional and national statistical organizations, and the academia.

h. Access, Dissemination and Storage

- a. The Bank strives to maximize public and timely access to, and use of Bank Development Data on an equitable and non-discriminatory basis.
- b. To the extent reasonably possible and appropriate taking into account potential ethical, legal and confidentiality considerations, Bank Development Data shall be made available under an open data license.
- c. Acknowledging the various needs and preferences of a wide and diverse community of potential users of Bank Development Data, the Bank adopts multiple modes of data and metadata dissemination (including, without limitation, API's).
- d. Bank Development Data disseminated by the Bank shall be made available together with their related Metadata.
- e. The Bank shall maintain a central, searchable and publicly accessible data catalogue for Bank Development Data and ensure that Staff (i) document Data Sets, and (ii) deposit related Metadata accordingly.
- f. Any dissemination by the Bank of Bank Development Data shall comply with applicable Bank policies, including the Access to Information Policy.

i. Efficient Data Collection

Any decision by the Bank to collect data or have data collected on the Bank's behalf (including, without limitation, a decision on the selection of data sources and the design of data collection methods and instruments) shall take into account, among other things, (a) cost-benefit, (b) the potential burden on respondents, and (c) subject to article 1(j), section III hereof, availability of suitable Third Party Data.

j. Assessment of Third Party Data

In cases where the Bank contemplates the dissemination by the Bank of Third Party Data, or the use of Third Party Data as an input to the production of Bank Development Data, the applicable Unit shall document its assessment of the suitability of such Third Party Data for the intended purpose, addressing, among other things, the following:

- a. whether quality of the Third Party Data is sufficient taking into account the principles that apply to Bank Development Data as per this Policy;

- b. whether use of the Third Party Data is appropriate from a legal, ethical and technical perspective.

k. Metadata Standards

The Bank's processes and systems used for purposes of Data Curation and dissemination of Development Data shall comply with the Metadata Standards.

l. Inter-operability

- a. The Bank aims to achieve maximum inter-operability of its data systems. For such purpose the Bank (i) applies an 'API-first' approach (meaning that all of a platform's functionality is accessible through APIs and available for other systems to leverage), and (ii) otherwise ensures that its data systems are built making use of common tools.
- b. The Bank ensures that its data and subject-matter specialists and IT experts closely collaborate in the development of data management and dissemination platforms.

m. Responsible Innovation

- a. Recognizing the importance of responsible innovation, the Bank shall ensure that the adoption by the Bank of new technologies (including, without limitation, Artificial Intelligence) in the Data Management Lifecycle shall comply with the Data Quality Framework.
- b. The Bank shall ensure that subject matter specialists (including statisticians and data scientists) are closely involved in the design, deployment and evaluation of such technologies as referred to in paragraph (a) of this article 1(m).

2. Roles and Responsibilities in Relation to the Data Quality Framework

- a. Acknowledging that the Bank's data activities are decentralized by design, (i) data quality assurance is a shared responsibility of all Staff involved in the production of Bank Development Data, and (ii) all Staff involved in the production of Bank Development Data shall act in accordance with this Data Quality Framework. Accordingly, each Unit shall allocate such resources as may be necessary or appropriate to sustain its Bank Development Data-related activities in accordance with the Data Quality Framework.
- b. The Chief Statistician is assigned to support and procure the application of this Policy within the Bank. For such purpose, the Chief Statistician shall, among other things:
 - a. Develop the Data Quality Framework, including, without limitation, by issuing Directives, Procedures and Guidance in connection with the same, as may be deemed necessary or appropriate.
 - b. Be consulted and advise on the development and adoption of Statistical and Data Management Standards.

- c. Define the Bank's Metadata Standards and provide guidance on the implementation of the same.
- d. Advise on and promote, the responsible innovation and adoption by the Bank of new technologies (including Artificial Intelligence) to the extent relating to Data Curation and dissemination of Bank Development Data.
- e. Define and support the implementation of the Data Stewardship Network.
- f. Represent the Bank (with full power of substitution) in deliberative fora on data and statistics.

SECTION IV – EXCEPTION

N/A

SECTION V – WAIVER

Only the Board may grant a Waiver of a provision of this Policy.

SECTION VI – OTHER PROVISIONS

Management will keep the Board updated on the gradual implementation of the Policy through the issuance of detailed practical guidelines over a period of two (2) years as of the date hereof.

SECTION VII – TEMPORARY PROVISIONS

N/A

SECTION VIII – EFFECTIVE DATE

This Policy is effective as of the date on its cover page. For the avoidance of doubt, the Policy is not intended to apply retro-actively.

SECTION IX – ISSUER

The Issuer of this Policy is as stated in the cover page.

SECTION X – SPONSOR

The Sponsor of this Policy is as stated in the cover page.

SECTION XI – RELATED DOCUMENTS

1. WBG Personal Data Privacy Policy
2. Bank Access to Information Policy
3. IFC Access to Information Policy
4. MIGA Access to Information Policy
5. WBG Information Security Procedure Framework
6. WBG Information Classification and Control Directive
7. Bank Procedure on Development Data Set Acquisition, Archiving and Dissemination
8. Principles of Staff Employment
9. Staff Rules

Questions regarding this Policy should be addressed to the Sponsor.