BETTER MEASUREMENT AND MONITORING OF DATA FOR DEVELOPMENT

APRIL 2022
A Stocktaking Report

CONTENTS

OVERVIEW

Performance indexes, maturity models, and other assessment tools provide an analytical framework for evaluating the capacity of government data systems. In this stocktaking report, commissioned for the Data for Development Global Research Hub (D4D.net) with the support of the International Development Research Centre (IDRC), Open Data Watch (ODW) maps twelve widely used performance indexes and assessment tools to the pillars of the Global Data Barometer (GBD) and the stages of Data Value Chain (DVC) to identify similarities and differences in concept and coverage. The results show that none of the indexes or tools currently available provide balanced coverage of government data systems, suggesting that a more comprehensive measure or a combination of complementary indexes and tools are needed to fully capture the functions of the data ecosystem.

PERFORMANCE INDEX
Composite indicators that measure the performance of statistical systems or government functions and produce a numerical score.

ASSESSMENT TOOL
Survey-like processes that document the functioning and management of statistical systems without producing a summary score.

The development of the Global Data Barometer responds to demand expressed at the 2019 Open Government Partnership summit for updated, in-depth, country-level insights on data governance and availability. Over a ten-year it will produce critical longitudinal data and analysis of government data systems in more than 100 countries. The Data Value Chain, first proposed by ODW, describes the steps through which raw data are transformed into high-value indicators that inform policies and impact people's lives. Together they provide complementary frameworks for describing the scope and relevance of the performance indexes and assessment tools discussed here.

In the next section describes the inventory of currently available performance indexes and assessment tools. We use the term performance index to describe composite indicators that measure the performance of statistical systems or government functions and produce a numerical score. Assessment tools are typically survey-like processes that document the functioning and management of statistical systems without producing a summary score. After introducing the indexes and tools, we review their country coverage, maturity, and sustainability.

In following sections, we map the contents of the indexes and tools to the conceptual frameworks of the four GDB core pillars and the stages of the DVC. To gain further insight into the relationships of the quantitative performance indexes, we examine the correlations of their scores and sub-scores. In the final section we summarize the results of the stocktaking with a set of findings and recommendations.

INVENTORY OF INDEXES AND TOOLS

The performance indexes and assessment tools included in this analysis were selected by Open Data Watch after a review of possible candidates and consultations with other index or tool producers. To ensure their relevance and comparability to the GDB, only indexes and tools applicable to government data in multiple sectors were included. Assessment tools that are currently in use and performance indexes that have at least one edition available since 2019 were included in the research inventory. Seven performance indexes provide cross-country comparisons using a numeric score, while five assessment tools enable countries or organizations to conduct assessments of government data and statistical systems based on qualitative and quantitative measures. All indexes and tools included in the inventory seek to measure some element of data governance, availability, quality, openness, or use and impact. The index and tools included in the inventory are shown in **Table 1**.

Table 1: Inventory of performance indexes and assessment tools

NAME	PUBLISHER	COUNTRY COVERAGE	FIRST EDITION	LATEST EDITION
PERFORMANCE INDEXES				
OURdata Index	Organization for Economic Co-operation and Development (OECD)	OECD countries	2015	2020
Open Data Inventory (ODIN)	Open Data Watch	Global	2016	2020
Statistical Performance Index (SPI)	World Bank	Global	2020	2020
Ibrahim Index of African Governance (IIAG)	Mo Ibrahim Foundation	Africa	2007	2020
Worldwide Governance Indicators (WGI)	World Bank	Global	1996	2019
Use of Statistics Index	PARIS21 (OECD)	Global	2019	2020
European Open Data Maturity Assessment (EODMA)	European Data Portal	Europe	2015	2020
ASSESSMENT TOOLS				
Open Data Readiness Assessment	World Bank	NA	2013	2015
Data Quality Assessment Framework	International Monetary Fund (IMF)	NA	2003	2012
Joined-Up Data Maturity Assessment	Global Partnership for Sustainable Development Data (GPSDD)	NA	2019	2020
Open Data Demand Assessment	The GovLab	NA	2018	2018
Open Data Maturity Model	Open Data Institute	NA	2015	2015

The list below describes the selected indexes and tools.

PERFORMANCE INDEXES

OURdata Index: This index measures data availability, data accessibility, and government support for data in 37 Organization for Economic Co-operation and Development (OECD) member and partner countries to identify open data policy achievements and challenges.

Open Data Inventory (ODIN): This index assesses the coverage and openness of official statistics in 187 countries to identify gaps, promote open data policies, improve access, and encourage dialogue between national statistical offices (NSOs) and data users.

Statistical Performance Indicators (SPI): This index assesses the maturity and performance of official statistical systems to improve development outcomes and track progress towards the Sustainable Development Goals (SDGs). The SPI covers 218 countries and territories, but complete data are only available for 174.

Ibrahim Index of African Governance (IIAG): This index measures and monitors governance performance in 54 African countries. It is comprised of four sub-indexes that measure dimensions of governance. The sub-index on foundations for economic opportunity includes measures of the capacity of the statistical system.

Worldwide Governance Indicators (WGI): This index is a composite of other indexes that measure perceptions of the quality of governance in 200 or more countries and territories. There are six WGI indexes. The Voice and Accountability index includes indicators of the reliability of basic economic and financial statistics for 204 countries.

Use of Statistics Index: This index measures the use of official statistics in development plans and documents by government decision makers in 184 countries with a focus on basic use, data disaggregation, monitoring and evaluation, and national policy documents such as national development plans and poverty reduction strategy papers.

European Open Data Maturity Assessment (EODMA): This index assesses the level of open data maturity in 32 European countries through its policies and programs to promote and support open data to government data and to foster high-quality data publication and increase impact.

ASSESSMENT TOOLS

Open Data Readiness Assessment (ODRA): This tool can be used to conduct an action-oriented assessment of the readiness of a government or individual agency to evaluate, design, and implement an open data initiative.

Data Quality Assessment Framework (DQAF): This tool allows countries to conduct self-assessments of five dimensions of data quality to guide country efforts to strengthen statistical systems, to inform preparation of International Monetary Fund (IMF) reports, and to guide data users in evaluating data for their own purposes such as policy analysis, forecasts, and economic performance. There are seven dataset-specific DQAFs, all concerning macroeconomic statistics, and one DQAF on household income and poverty.

Joined-Up Data Maturity Assessment: This tool can be used by national statistical offices (NSOs) and other entities that control or process data in the development sector to assess organizational, human, data, and technological interoperability within a statistical system.

Open Data Demand Assessment: This tool provides open data policy makers and practitioners with an approach for identifying, segmenting, and engaging with demand for open data to empower data champions within public agencies.

Open Data Maturity Model: This tool enables an organization to assess how well it publishes and consumes open data and to identify actions for improvement.

Country coverage

Four of the performance indexes provide global coverage while the other three provide scores for a limited set of countries. In addition to providing scores for countries, many of these indexes also provide aggregate scores for geographic regions, income groups, and a global average score. The three remaining indexes provide scores for a limited set of countries in specific regions, focusing on high-income countries or, in the case of the IIAG, the countries of Africa. The selective, regional coverage of these indexes limits comparisons with other indexes to the subset of countries they have in common.

Country coverage of the global indexes has increased in recent years. For example, in 2015 ODIN provided scores only for 125 low- and middle-income countries before including high-income countries in 2016 for a total of 173; in 2020 ODIN included 187 countries. The Use of Statistics index included 69 countries in 2019 but expanded to 174 in 2020.

Gaps in country coverage of these indexes limits their comparability.

Assessments are needed in low- and middle-income countries as well as in high-income countries.

Most assessment tools do not have a central registry of their application. Because of this and because self-assessments may not have been reported outside the country, it is not possible to give a definitive list of where and how often they have been applied. The World Bank lists 18 countries that have conducted an **Open Data Maturity Assessment** between 2013 and 2020. The IMF includes data quality assessments as part of its Reports on Standards and Codes (ROSC). The IMF's **Dissemination Standard Bulletin Board** lists 83 countries that have had data ROSCs conducted since 2001. Similar statistics for the Joined-Up Data Maturity Assessment, the Open Data Demand Assessment, and the Open Data Maturity Assessment are not available.

Maturity: Characteristics of older and recent indexes and tools

The length of time an index or tool has existed can be an indication of its maturity, as each iteration involves reviews and critiques that provide an opportunity to further refine and hone the index or tool. The oldest index or tool in the inventory is the WGI, first published in 1996. The only index or tool introduced before 2010 was the IIAG (2007) and the IMF's DQAF (2003). The most recently introduced are the SPI (a successor to the World Bank's Statistical Capacity Index) and the Use of Statistics Index.

Most indexes and tools focus on data production, but newer ones have begun to measure data use.

Newer indexes can contribute to the exploration of ways to provide objective measures of data use and impact.

The older indexes and tools tend to focus on limited topics while recent ones have expanded greatly in scope. The oldest indexes, the IIAG and WGI, focus broadly on governance issues, such as government effectiveness, accountability, and corruption. Each includes only a small number of indicators related to data. The DQAF is the oldest among the tools, and it has a specific focus on the quality of macroeconomic data. Indexes and tools developed in mid-2000 or later have a broader scope, including measures such as data availability, quality, policies, or access across a wide spectrum of statistics. In general, most indexes and tools focus on issues related to data production and dissemination, but the newest have begun to develop ways of measuring data use and impact.

Recent **indexes** and **tools** are exploring ways of measuring data use and impact.

The Use of Statistics Index is the only index or tool focused exclusively on the use of statistics. It employs keyword searches to track the use of terms referring to data and statistics in national development plans and documents.

The SPI includes indicators in its first pillar to measure data use by different sets of stakeholders, but so far, direct measurement of data use is only included for dimension 1.5 (data use by international bodies) due to a lack of an established methodology for data use by government, civil society, and academia.

The EODMA seeks to capture examples of re-use and impact. Recognizing that there is no generally accepted methodology for measuring impact, the EODMA questionnaire asks for examples of the re-use of government data in policies, publications, or research and for subjective impressions of the impact of open data in particular settings.[1]

The Open Data Readiness Assessment is designed to address both the supply and demand for open data, but it advises "… to use this tool alongside other tools that focus more deeply on specific areas of interest (e.g., civil society demand for Open Data or technical capacity of the public sector)."

The Open Data Demand Assessment evaluates the demand for open data through a series of questions concerning the impact of open data.

---

[1] The EODMA questionnaire notes that "[A]lthough the impact dimension sets a strong focus on open data re-use cases, the European Data Portal does not consider the availability of re-use examples as direct evidence of impact."

Sustainability: Regular financing for indexes and tools

Sustained assessment programs, whether conducted through performance indexes or assessment tools are likely to have greater impact than one-off efforts. Sustaining production of the performance indexes and their periodic updates requires financial and technical resources to collect fresh source data and manage quality control, particularly when methodologies and country coverage are expanded. Assessment tools also require revisions from time to time to stay abreast of new methods and recommended standards, and financial and technical support may be needed to assist countries carrying out self-assessments.

Sustained production and regular updates will ensure usefulness over time.

Assessments require sponsors that are able and willing to bear the cost of regular updates while making them available as public goods.

Many of the indexes and assessment tools are initiatives of large international organizations such as the IMF, OECD, the European Commission, and the World Bank. In fact, a quarter of all assessments—the WGI, SPI, and ODRA—are funded directly by the World Bank. Because these tools and indexes are sponsored by large and well-funded government organizations, the risk of financial instability should be minimal, but changes in organizational priorities or budgeting pressure can put even well-established programs in jeopardy.

For other measures, sustainability depends on the ability of non-governmental organizations to support them. The IIAG, for example, has been developed by and funded by the Mo Ibrahim Foundation, a non-grant making and non-fundraising organization in Africa with a focus on governance and leadership. ODIN was developed by Open Data Watch, a nonprofit organization that relies on grants for the sustainability of ODIN. Similarly, The GovLab and GPSDD are non-profit organizations funded by a combination of grants and contracts.

MAPPING THE INDEXES AND TOOLS

The GDB intends to provide a comprehensive framework for assessing a country's data ecosystem. The mapping of each index or tool examines the indicators that contribute to their overall assessment and locates them in the relevant GDB pillar. The purpose of the mapping exercise is to document how and where the 12 indexes and tools overlap with the GDB's core pillars and where there may be gaps in their coverage.

A similar mapping has been carried out using the framework of the DVC. The DVC adds a different perspective that describes stages in data's life cycle from production to use. It uses a process-orientated approach to describe value added to data at each stage, and it provides a way to examine which parts of the data lifecycle have been emphasized or overlooked in assessments of data systems.

Mapping to the Global Data Barometer's core pillars

The four pillars of the GDB measure data governance, data capabilities, data availability, and data use and impact. Definitions for each pillar, taken from the **Global Data Barometer's Handbook**, are summarized in **Table 2**.

Table 2: Global Data Barometer core pillars

GLOBAL DATA BAROMETER PILLAR	DEFINITION
DATA GOVERNANCE	This pillar includes factors related to the management of data and the presence of legal and policy frameworks to guide that management (e.g., presence and content of policies on data protection, open data policies, or open data licenses).
DATA CAPABILITIES	This pillar includes factors related to a country's ability to create, manage, and use data effectively, including data producer and data user capabilities (e.g., the implementation or existence of data literacy training for civil servants or non-governmental stakeholders, presence and quality of data portals, or the ability to monitor reuse of data).
DATA AVAILABILITY	This pillar includes factors related to a country's data completeness, sectoral coverage, data quality, and data openness (e.g., the timeliness and frequency of data publication, the interoperability of data, or the implementation of international data standards and classifications).
DATA USE AND IMPACT	This pillar includes factors related to a country's level of data use by various actors and the impact of their data (e.g., evidence of data use by the private sector, evidence of data use by the public sector, or evidence of data use by the academia).

The indicators of each of the indexes and tools were mapped to the four GDB pillars. The mapping is summarized in three levels of emphasis shown in **Table 3**:

Blue represents pillars with **high emphasis**, which occurs when the percentage of indicators included in the index or tool and mapped to a pillar is significantly higher than in other pillars.

Yellow represents **low emphasis,** which indicates only a handful of indicators are mapped to a pillar.

No color represents **no emphasis** when no indicators were mapped to a given pillar.

Indicators that do not fit into any pillar were rare, but most often occurred in assessments that concerned topics outside the scope of the four GDB pillars such as financing for data collection and publication, as well as infrastructure capabilities.

Note that the emphasis on a pillar is not an indication of its significance relative to other indexes or tools. The WGI, for example, has a high emphasis on data availability and none on the other pillars but includes a more limited set of availability indicators than, say, ODIN. Furthermore, the categorization of each indicator according to the four pillars was a subjective process, and therefore the percentages in **Table 3** are meant to provide only a general sense of emphasis rather than an exact weight.

The mapping exercise visualized in **Table 3** shows the indexes and tools with the greatest concentration of indicators in each GDB pillar. None of the mapped indexes and tools provide equal emphasis on all four pillars. The pillar receiving the greatest emphasis is data availability, which has a high emphasis in five indexes and tools and low emphasis in an additional five. Only two indexes and tools have no emphasis on data availability. There is a more moderate concentration on data governance and data capabilities. A different set of three indexes and tools with a high focus and six with a low emphasis is mapped to each of these pillars. Data use and impact is the least emphasized pillar, with five indexes and tools having no emphasis at all.

None of the indexes or tools cover the full scope of the GDB pillars.

Data availability receives greatest emphasis with data use and impact receiving the least emphasis.

Differences are also apparent when comparing performance indexes to diagnostic tools. For instance, indicators on data availability are more common in performance indexes, while indicators that measure data capabilities are more commonly found in the assessment tools. This likely reflects the relative ease of assigning objective scores to track data availability by external reviewers, while the measurement of capabilities requires an internal review. Data use and impact also receives slightly greater focus among indexes, but it remains a particularly challenging area of research that many indexes are still developing. The SPI proposes five dimensions to measure data use, but all but one are still under development according to its **methodology**.

Table 3: Indicator coverage across GDB pillar (% of indicators in each pillar)

NAME	DATA GOVERNANCE	DATA CAPABILITIES	DATA AVAILABILITY	DATA USE AND IMPACT
PERFORMANCE INDEXES				
Open Data Inventory	10%	0%	90%	0%
OURdata Index	42%	38%	17%	4%
Statistical Performance Index	6%	6%	76%	12%
Ibrahim Index of African Governance	20%	20%	60%	0%
Worldwide Governance Indicators	0%	0%	100%	0%
Use of Statistics Index	0%	0%	0%	100%
European Open Data Maturity Assessment	37%	20%	5%	37%
ASSESSMENT TOOLS				
Data Quality Assessment Framework	14%	5%	81%	0%
Open Data Readiness Assessment	22%	72%	6%	0%
Joined-Up Data Maturity Assessment	37%	25%	25%	13%
Open Data Demand Assessment	0%	19%	12%	69%
Open Data Maturity Model	25%	67%	0%	8%

■ High emphasis   ■ Low emphasis   □ No emphasis

The most common and less common indicators used to measure concepts within each GDB pillar are shown in **Table 4**. Examples illustrate the type of indicator that could be included in each category. They are based on indicators from existing indexes and tools but are not shown verbatim. **Annex 1** provides a representative list of examples from specific indexes and tools that correspond with each pillar.

The mapping exercise in **Table 3** shows the least emphasis is on the data use and impact pillar. **Table 4** shows that indicators in all four pillars tend to focus more on data production than data use. For example, under data capabilities, indicators on the capabilities of producers are common, but indicators on data user capabilities are less common.

Table 4: Common and less common indicators grouped by GDB pillars

GLOBAL DATA BAROMETER PILLAR	COMMON INDICATORS	LESS COMMON INDICATORS
DATA GOVERNANCE	**Data management policies** *Ex: Mechanisms are in place to monitor data quality and openness* **Open data policies** *Ex: Existence of an open by default policy and open data initiatives* **Open data licenses** *Ex: Existence of data license or terms of use*	**Data protection/privacy policies** *Ex: Existence of legal and policy framework for the protection of personal privacy* **Feedback mechanisms** *Ex: Existence of formal requirements to consult with users prior to data publication*
DATA CAPABILITIES	**Data producer skills and resources** *Ex: Designated staff for data management and data stewardship* **Government support for reuse** *Ex: Existence and frequency of programs/events designed to promote data reuse among different types of users*	**Data user skills and the enabling environment** *Ex: Strength of research communities' capabilities in data analysis* **Political support** *Ex: Visible political support for open data*
DATA AVAILABILITY	**Data availability** *Ex: Availability of specific indicators* **Data quality** *Ex: Adherence to international guides or use of internationally accepted classification systems*	**Data openness and accessibility** *Ex: Availability of data in machine-readable and non-proprietary formats made available free of charge* **Availability and quality of metadata** *Ex: Comprehensiveness of metadata, use of internally accepted standards for metadata dissemination* **Data disaggregation** *Ex. Availability of indicators by sex, age, disability status, and other characteristics*
DATA USE AND IMPACT	**Evidence of use** *Ex: Examples of data reuse by civil society, national legislature/executive branch, academia, and other users* **Evidence of impact** *Ex: Examples of data reuse that show impact (economic growth, innovation, policy development, improvement of service delivery, increase in institutional transparency and accountability)*	**Mechanisms in place to monitor use or impact** *Ex: Existence of activities or mechanisms to monitor data use/ impact, existence of methodology to measure impact of data*

Mapping to the Data Value Chain

Mapping index and tool indicators across the stages of the DVC provides a different look at the scope of indicators currently in use. The DVC stages, summarized in **Table 5**, describe the steps by which raw data are transformed into high-value indicators and finally to their use in implementing policies that have an impact on people. The mapping of indexes and tools across GDB pillars and mapping across DVC stages complement each other. The DVC focuses on activities directly related to the life cycle of data but does not explicitly consider elements of the enabling environment of the data ecosystem, some of which are covered by GDB pillars. There are, nevertheless, some similarities between DVC stages and the GDB pillars, which may result in the repetition of similar findings. However, mapping the indexes and tools to the DVC allows us to look at the indicators along different dimensions that complement the structure of the GDB.

The mapping of indexes and tools across the GDB pillars and DVC stages complement each other.

The DVC allows us to look at indicators along different dimensions, complementing the structure of the GDB.

Table 5: Stages of the Data Value Chain

DATA VALUE CHAIN STAGE	DEFINITION
DATA COLLECTION	This stage includes indicators concerning policies on and practices of data collection, data quality, adherence to standards and classifications, and data openness. It also includes feedback mechanisms to inform data collection processes and indicators of the capacity to collect data.
DATA PUBLICATION	This stage includes indicators concerning the amount and type of data and metadata published, along with how the data are published (file formats, download options, data portal functionality).
DATA UPTAKE	This stage includes indicators concerning explicit actions taken by the data producers to increase the use of data. It includes open data licenses, measures to increase data accessibility, data literacy programs, hackathons, and other similar activities.
DATA IMPACT	This stage includes indicators that measure evidence of use or reuse of data for policy decision making, business, or project creation (ex: application development), academic, news or other reports, or other uses.

The mapping exercise for the DVC stages uses the same methodology as the mapping done for the GDB pillars, with blue reflecting high emphasis, yellow reflecting low emphasis, and no color reflecting no emphasis. While the majority of indicators within the indexes and tools could be mapped across the DVC stages, indicators related to enabling factors such as governance, financing, producer capabilities, infrastructure, and feedback loops do not relate directly to the stages of the DVC and were not included.

As **Table 6** shows, data publication receives the most emphasis across the indexes and tools, which coincides with the emphasis on data availability in the GDB. Data uptake, which relates to connecting users to data and active promotion of use, receives significant focus among the assessment tools, but less among the performance indexes, which reflects the relative ease of including a focus on capacity as part of internal assessments. Data collection receives less emphasis by both indexes and tools. As with the GDB pillars, there is a lack of emphasis on data impact. The mapping to the Data Impact DVC stage closely matches the Data Use and Impact pillar of the GDB.

Table 6: Indicator coverage across DVC stages (% of indicators in each stage)

NAME	DATA COLLECTION	DATA PUBLICATION	DATA UPTAKE	DATA IMPACT
PERFORMANCE INDEXES				
Open Data Inventory	0%	87%	13%	0%
OURdata Index	28%	24%	43%	5%
Statistical Performance Index	69%	17%	3%	11%
Ibrahim Index of African Governance	17%	67%	17%	0%
Worldwide Governance Indicators	0%	100%	0%	0%
Use of Statistics Index	0%	0%	0%	100%
European Open Data Maturity Assessment	9%	40%	15%	37%
ASSESSMENT TOOLS				
Data Quality Assessment Framework	41%	55%	4%	0%
Open Data Readiness Assessment	9%	5%	62%	24%
Joined-Up Data Maturity Assessment	19%	45%	36%	0%
Open Data Demand Assessment	0%	8%	0%	92%
Open Data Maturity Model	9%	36%	55%	0%

As was done in the mapping to the GDB pillars, indicators were further organized into categories within the DVC stages and summarized in **Table 7.** This analysis highlights the most and least common indicators, revealing potential overlaps and gaps. Examples illustrate the type of indicator that could be included in each category, but do not reflect the wording of any specific indicator from the indexes and tools. The examples are grouped in subcategories that are not used by the DVC but are descriptive of the types of indicators found in the indexes and tools. For example, data collection is divided into five subcategories: data quality measures, stakeholder engagement, data sources, open data policies, and data privacy policy and legislation. **Annex 2** provides a representative list of examples from specific indexes and tools that correspond with each stage of the DVC.

Table 7: Common and less common indicators grouped by DVC stages

DATA VALUE CHAIN STAGE	COMMON INDICATORS	LESS COMMON INDICATORS
DATA COLLECTION	**Data quality measures** *Ex: Source data are routinely assessed, e.g., for coverage, sample error, response error, and non-sampling error* **Stakeholder engagement** *Ex: Various open data events are organized by a mix of actors (public and private sector, civil society, and academic bodies) throughout the country to foster the exchange on the open data topic* **Data Sources** *Ex: Measures of the extent to which sources being used enable the necessary statistical indicators to be generated*	**Open data policies** *Ex: Open Data policies and strategies are in place at national level* **Data privacy policy and legislation** *Ex: Compliance with privacy legislation; the development of data inventories in the public bodies at national, regional, and local levels is defined as priority in the national policy and/or strategy*
DATA PUBLICATION	**Dissemination mechanisms, standards, and activities** *Ex: Portal visibility is enhanced by organizing/attending info sessions and/or events to promote the national portal* **Accessibility of data** *Ex: Data accessible free of charge and in open formats on the central/federal data portal.* **Availability of data** *Ex: How many data sets are available?*	**Existence of metadata** *Ex: Documentation of concepts, scope, classifications, basis of recording, data sources, and statistical techniques is available* **Quality of published data** *Ex: Statistics are consistent within the dataset.*

DATA VALUE CHAIN STAGE	COMMON INDICATORS	LESS COMMON INDICATORS
DATA UPTAKE	**Data promotion and engagement** *Ex: Existence of formal partnerships with businesses and the civil society to support data re-use* **Open data plans and policies** *Ex: The national open data strategy incentivizes the re-use of open data by both the public and private sectors and access to real-time data.*	**Capabilities of users** *Ex: The organization promotes the availability of third-party learning resources and tools, and training activities for civil servants working with data are in place*
DATA IMPACT	**Evidence of impact** *Ex: Conducted or financed research on socio economic impact of open data.* **Use and re-use of data** *Ex: Various re-use examples exist that show the impact of open data on enable better policy and decision making processes.*	**Demand for data** *Ex: What is the extent of intra- and inter-government actual demand and latent demand for data?*

Indicator mapping: Summary

The previous two sections have examined the contents of the performance indexes and assessment tools by mapping their indicators to the pillars of the GDB and the stages of the DVC. The purpose of the mapping was to identify concepts that are well-measured by the indicators used by the indexes and tools and, conversely, concepts that are not well measured or for which few indicators have been found. During the mapping exercise, two pillars of the GDB were found to align closely with two stages of the DVC: Data Availability in the GDB and the Publication stage of the DVC and Data Use and Impact in the GDB and Impact in the DVC. This report will discuss the mapping of these pillars and stages jointly and the remaining pillars and stages separately.

DATA GOVERNANCE (GDB)

The most common indicators found among the indexes and tools are related to data management policies, open data policies, and open data licenses. Data management policies vary the most in how they are measured, while indicators of open data policies and licenses are more standardized. However, few indexes or tools included indicators on data protection or privacy policies or on policies concerning formalized feedback mechanisms for users. Given the concern for data protection and privacy policies across the world, particularly in conversations about open data and data governance, this is a critical area that represents a

significant gap in how the indexes and tools measure data governance. Another gap relates to whether opportunities for feedback are available to users. Some assessment tools ask about feedback, but many fail to measure the existence of formal feedback mechanisms such as features on data portals that allow users to submit direct feedback. This type of formal, institutional mechanism is best practice and ensures a sustainable process of engagement between data producers and users.

Data governance is a broadly defined concept with many activities also relating to other categories.

Few indexes or tools included indicators strictly related to data policies or management practices such as protection or privacy.

DATA CAPABILITIES (GDB)

This pillar represents a broad collection of activities, some of which could be included in other pillars. Government support for re-use of data also appears in the Impact and re-use pillar; data management functions may be included under data governance; and data creation activities may be included in the data collection and publication stages of the DVC and the data availability pillar of the GDB.

Measures of internal producer capabilities are more available than measures of user skills.

Quantifiable measures of user skills are not included among existing indexes.

The assessment tools put greater emphasis on data capabilities. They often focus on assessing the skills and resources of data producers and whether there is government support for re-use of data. Evidence of government support for data re-use includes data promotion initiatives targeting businesses and civil society and data literacy programs in government. Indicators of the skills and of data users and measuring political support for open data are less well covered by the indexes and tools. Recent publications recognize the importance of human capabilities as both **barriers** to and **determinants** of data use. To better integrate users within the data ecosystem, data capabilities should be measured from both the data producer and user perspective.

COLLECTION (DVC)

Data collection activities, which include processes related to the identification, collection, and processing of observations to produce high-quality datasets, are not explicitly included in the GDB pillars, although they might be considered part of data capabilities. Most of the indexes and tools examined in this study do not focus on this stage. Those that do have indicators that measure engagement with stakeholders, correct use of sources in data collection, and issues related to data quality. Some indexes also focused on legislation that regulates privacy and openness at the collection stage.

Most indexes and tools focused on the outcomes of data collection rather than data collection itself.

Survey design, adherence to standards, data compilation, and timeliness are all relevant to data collection.

DATA AVAILABILITY (GDB) AND PUBLICATION (DVC)

This pillar is similar in concept to the publication stage of the DVC, which involves the analysis, release, and dissemination of data. Although the datasets considered by each index or tool differed, most were concerned with the availability of development data. But there is a lack of attention to disaggregated data, with only two indexes including an explicit requirement for disaggregation (see the box **Elevating a focus on gender data** for further discussion). There are also gaps related to the availability and quality of metadata. The adherence to best practices for openness and accessibility in data dissemination was also inconsistently addressed.

Data availability receive attention from most indexes and tools, but few assess disaggregation.

Disaggregation is needed to measure differences in gender, age, ethnicity, and other important variables.

BOX: Elevating a Focus on Gender Data

Data that are disaggregated by sex or that reflect gender-related issues allow decision makers to develop better policies and initiatives that improve lives and help achieve gender equality. Sponsors of indexes and tools should ensure that a gender focus is included in all dimensions. However, only four of the indexes reviewed here have at least one indicator related to gender. And none of the diagnostic tools have an explicit focus on gender.

Gender-related indicators included in performance indexes:

OPEN DATA INVENTORY
- Are indicators and sex-disaggregation available in ten gender-relevant data categories?

OURDATA INDEX
- Existence of requirements to provide timely access to disaggregated data, although disaggregation by sex is not specified.

STATISTICAL PERFORMANCE INDICATORS
- Labor force participation rate by sex and age
- Availability of data to measure SDG Goal 5: Gender Equality

USE OF STATISTICS INDEX
- Occurrence of the word "gender" and other related terms in government policy documents.

Many indexes and tools include indicators that may measure sex-disaggregation indirectly. For example, the SPI covers the availability of data to measure SDGs beyond Goal 5, and while many indicators needed to measure these goals require sex disaggregation, it is not made explicit, and the extent of sex-disaggregation is not included in their assessments.

UPTAKE (DVC)

Uptake involves connecting users to the data through intermediaries and encouraging use. Most indicators related to this topic focus on producer activities to promote uptake or engagement with government data. Some also take note of official plans and policies. Legislation that promotes data use is also monitored to some extent. However, as noted under the GDB capability pillar, there are few measures of data literacy or the ability of users in government or in civil society to understand and make decisions using data.[2]

DATA USE AND IMPACT (GDB) AND IMPACT (DVC)

Data use and impact is measured by indicators related to the level of data use by actors and the impact of their use of data. Use or re-use is more readily measured than impact. The Use of Statistics Index measures the occurrence of statistical terms and indicators in national policy documents as a proxy for data use by decision makers. Other common indicators in this pillar addressed whether or not mechanisms were in place to measure use, such as counting downloads on a data portal. The SPI intends to include indicators that measure data use by different stakeholders, but exact methodologies are still in development. Some assessment tools ask whether a methodology is in place to measure the extent of impact, but it remains challenging to develop objective, quantitative measures of the impact of data. For example, the EODMA asks whether governments have methodologies in place for measuring impact and then uses subjective reports on the impact of data use on government processes, public awareness, or research. This falls somewhat short of quantifying the benefits realized from using data, although increasing use is presumably an indication that data has beneficial impacts.

Most indexes and tools rely on indicators measuring data use policies and programs or anecdotal reports of data use.

The PARIS21 Use of Statistics index is an exception to this, directly measuring the occurrences of statistical terms and indicators in national policy documents.

A closely related problem is measuring the demand for data. The demand for data reflects the benefits they bring, but because open data are a public good, it is difficult to get data users to reveal the value they receive from using data. The ODDA assessment tool attempts to gauge the demand for open data by first identifying a significant public problem, the data needed to solve that problem, and actors that can use the data. ODDA's systemic approach follows a series of steps that have parallels in the GDB pillars and the DVC stages. They begin with an assessment of data quality, data governance, and data availability and then engage with stakeholders who will use the data, producing valuable outcomes. Valuing those outcomes remains the core problem of quantifying benefits and measuring impact.

---

[2] The PARIS21 index of data literacy was not included in this study because it was last produced in 2016.

Quantitative comparison of performance indexes

The seven performance indexes produce scores and rankings that can be compared with one another. As a test of their similarities and differences, we computed correlation coefficients using the index scores and, where available, the sub-indexes from which they are composed. Indexes that measure similar concepts may be expected to produce similar scores (allowing for differences in scaling), and therefore exhibit positive correlations. Weak correlations do not imply that the indexes are deficient but only that they are measuring different concepts. However, correlations may also be affected by confounding variables, such as income levels or geographic location, that have a similar effect on all indexes.

Table 8 shows the pairwise correlations of the overall index values. Because the WGI does not produce an overall score from its six indexes, only the value of the WGI index of voice and accountability, which includes data-related indicators, was included. Comparisons between the seven indexes are limited to the countries they have in common in the most recent year for which data are available. Four of the performance indexes are global in scope: ODIN, SPI, Use of Statistics, and the WGI. The remaining three are limited to a region (European Open Data Maturity Assessment and the Ibrahim Index of African Governance), or international organization (OECD's OURData Index). Gaps in country coverage further reduce the countries shared between pairs of indexes.

The highest correlation is found between the two governance measures, the IIAG and WGI. Although these indexes use different methodologies to compute their scores, they are both composites of many closely related indicators and indexes, some of which appear in both indexes. The next highest correlation is between the SPI and ODIN. Both indexes measure aspects of data governance, openness, and the availability of official statistics. The SPI includes two indicators derived from ODIN. ODIN and SPI also show strong correlations with the IIAG overall index in African countries and the WGI index of Voice and Accountability. This suggests a relationship between the performance of the general functions of government and the performance of the official statistical system.

Correlation between indexes of data availability suggest a consensus in how coverage, openness, and capacity are measured.

However, similar measures are not widely available for other types of government data.

The EODMA is less strongly correlated with the other indexes. It may be measuring a different concept than the SPI and ODIN and the two governance indexes, perhaps one that is less closely linked to the capacity of official statistics systems. It may also be the case that among the subset of countries included in the EODMA, there is less of association between the adoption of open data policies and the outcomes measured by the other indexes. Finally, PARIS21's Use of Statistics index and the OECD's OURData index give quite different signals for the countries to which they apply. The Use of Statistics index is negatively correlated with SPI, OURData, EODMA, and WGI, and has positive but near-zero correlations with ODIN and IIAG. OURData is negatively correlated with the SPI, ODIN, and the WGI. Correlations between IIAG and OURData and EODMA cannot be computed because the latter two indexes do not include African countries.

Table 8: Correlations between overall index values (%)

	SPI Overall 2019	ODIN Overall 2020/21	PARIS21 Use of Statistics 2019	OURdata Overall 2019	European Open Data Maturity Assessment Overall 2020	World Governance Indicators Voice and Accountability 2019	Index of African Governance Overall 2019
SPI Overall 2019	100						
ODIN Overall 2020/21	79.1	100					
PARIS21 Use of Statistics 2019	-2.4	1.9	100				
OURdata Overall 2019	-0.7	-10.0	-6.9	100			
European Open Data Maturity Assessment Overall 2020	41.5	45.1	-23.5	30.6	100		
World Governance Indicators Voice and Accountability 2019	54.2	46.0	-2.2	-36.2	39.1	100	
Index of African Governance Overall 2019	66.9	59.8	3.5	NA	NA	87.9	100

Looking at their areas of emphasis within the GDB pillars in **Table 3**, the four indexes with the greatest emphasis on data availability—SPI, ODIN, WGI and IIAG—are the most strongly correlated. EODMA and OURData, which have similar profiles and place the greatest emphasis on governance are moderately correlated, although EODMA is somewhat more strongly correlated with SPI, ODIN, and WGI while OURData, which measures adherence to the International Open Data Charter, exhibits weakly negative correlations with the rest of the indexes. This suggests that measurements of policies or commitments as reported by governments may not be good predictors of outcomes.

The PARIS21 Use of Statistics index is an outlier among indexes.

It is the only index directly measuring data use, and the anomalous correlation results merit further investigation.

Among the global indicators, the Use of Statistics—with its exclusive emphasis on data use—is an outlier. The original 2019 report on this index also found the index to be negatively correlated with broad measures of statistical capacity (the World Bank's old SCI) and development (UNDP's Human Development Index) (PARIS21 2021). This may be a problem of sample selection. The index is computed by text mining two types of documents: national development plans (NDPs) and national poverty reduction plans (NPRPs). As a result, it may give higher scores to NDPs and NPRPs from poorer countries that have benefited from more intensive guidance from donors or consultants and that are more likely to use the language in the text-mining vocabulary. Although the anomalous results of the Use of Statics index require further investigation, the text mining algorithm employed by PARIS21 may have useful applications in other contexts.

The relationships between the performance indexes were further investigated at the sub-index level. A mapping of indicators between the sub-indexes was used to identify similarities and differences along with correlations between the subindexes. A discussion of these results and the full correlation matrix is available in a supplementary report.

FINDINGS AND RECOMMENDATIONS

This report has compared the scope and contents of selected performance indexes and assessment tools with the framework of the GDB pillars and the stages of DVC. The mapping exercise has identified concepts within these frameworks that are well measured by current indexes and tools and some that are less well measured. A correlation analysis has revealed other similarities and differences across the performance indexes. The findings suggest several general recommendations.

The mapping of the twelve indexes and tools to the GDB pillars shows that none have a strong emphasis across all stages of the GDB. Data availability is the primary focus of most indexes, while the assessment tools place greater emphasis on data capabilities. Data impact and use is least measured. Employing the DVC framework, the performance indexes focus on data publication, while the assessment tools put greater emphasis on measures of data uptake. The least well-measured pillar is data use and impact.

This final section highlights the principal findings from this review and makes recommendations to address them. These recommendations are not intended solely for the development of the GDB but address more generally the strengths and weaknesses of the performance indexes and assessment tools currently used to monitor the production and use of government data. Taken together, the indexes and tools offer a multidimensional view that can be used to identify needs, set policies, and improve the performance of national data systems. Collaboration between the producers of indexes and tools, statistical agencies, and data users can make them even more useful.

Measuring GDB pillars

DEFINING DATA GOVERNANCE	Data governance is a broadly defined concept. Many activities included under data governance may also be included as elements of data capabilities, availability, and use. Few indexes or tools included indicators of data protection or privacy policies.
Recommendation:	*Indexes seeking to measure governance should adopt a unique definition of the domain of data governance and seek indicators that provide evidence of good policies and data management practices, including data protection and privacy measured from both a producer and a user point of view.*
GREATER FOCUS ON USER CAPABILITIES	Data capability receives relatively little attention from the performance indexes but somewhat greater attention from readiness or maturity assessments such as ODRA and ODMM. Indicators of internal capabilities for data management and programs to encourage data use are more readily available than measures of user skills.

Recommendation: *The readiness and maturity assessments offer useful measures of data capabilities with a particular focus on user skills and uptake. A broad-based index measure could set itself apart by developing quantifiable measures of user capabilities. Doing so would contribute to the better integration of users into the data ecosystem.*

INCREASE ATTENTION ON DATA DISAGGREGATION

Data availability and the corresponding publication stage of the DVC are headline topics that receive attention from many indexes and tools that draw attention to data gaps. But only a few fully assess the availability of disaggregated indicators needed to measure differences in gender, age, ethnicity, or other characteristics of vulnerable populations.

Recommendation: *Measures of data availability provided by indexes and assessment tools should go beyond counts of aggregate indicators to include assessments of the availability of all relevant disaggregations. Greater attention should also be given to the availability of adequate metadata and to demonstrated adherence to open data practices.*

MEASURING DATA USE AND IMPACT

Older indexes tend to emphasize indicators that measure data production while more recent indexes have begun to measure data use and impact. Data use and impact has received greater emphasis in performance indexes applied to high-income countries. However, these measures often rely on indicators that report the existence of policies or programs to encourage data use or on anecdotal reports of data used in policy and planning documents, but they do not directly measure data use or quantify the results. An exception is the PARIS21 Use of Statistics index that was found to be weakly and negatively correlated with most other indexes.

Recommendation: *Further research should be done to evaluate policies and programs that monitor data use and its impacts. Can these methods be applied more generally to produce quantified measures of data use and impact that can be applied to countries across the world? These efforts are still in their infancy. They need encouragement and rigorous testing.*

STARTING WITH DATA COLLECTION

Data collection—the identification, collection, and processing of raw data—is a complex process that requires capable, well-governed, and adequately funded statistical organizations. The indexes and tools reviewed here are more concerned with the outcome of the data collection stage than data collection activities.

Recommendation:	Data collection through census, surveys, and administrative records is the starting point for producing high quality data. A comprehensive assessment of the data ecosystem should include measures of the data collection stage, including survey design, adherence to standard definitions and classifications, frequency of data collection, good practices in data compilation, and timeliness of publication.

Structure of indexes and tools

OBJECTIVE, VERIFIABLE MEASURES NEEDED	Some indexes and many of the assessment tools rely on self-reported indicators of the adherence to policies or subjective assessments of current practices. Objective measures of the implementation and outcomes of policies and programs are less often available.

Recommendation:	Self-assessments are important tools for internal evaluations but are less useful as a yardstick measure for comparison with other organizations or countries, particularly when they are based on qualitative or subjective assessments. Reliance on self-reports—through surveys or interviews—may introduce biases from self-interested reporters or through non-response. Indexes or tools used to make comparisons between countries or programs should be based on objective indicators than can be reliably measured over time. Verifiable indicators provide incentives to make productive changes and reduce incentives to "game the system."

INDEXES OF DATA AVAILABILITY ARE STRONGLY CORRELATED	Strong correlations between ODIN and the SPI and the broad-based indexes of government performance—the WGI and IIAG—suggest a consensus on the measurement of the coverage, openness, and capacity of official statistical systems. However, similar measures are not widely available for other types of government data.

Recommendation:	Measures that encompass all data produced by governments should consider extending the ODIN and SPI methodology to produced datasets outside the national statistical system to provide a comprehensive measure of the availability and openness of public data.

BALANCE DATA PRODUCTION AND DATA USE

The indicators used by performance indexes and assessment tools are more likely to reflect the activities of data producers than data users.

Recommendation:

The GDB should balance its four pillars by including indicators that provide robust measures of their concepts from both a producer and user viewpoint. Although ex-post weights can be used to provide a numerical balance of the pillar results, they do not compensate for information that goes unmeasured because of missing or inadequate indicators.

IMPROVING COUNTRY COVERAGE

Gaps in country coverage and infrequent or irregular updating limit the usefulness and comparability of indexes between countries and over time.

Recommendation:

Assessments of the data ecosystem are needed in low- and middle-income countries as much as in high-income countries to guide their development. Indicators used to measure performance should be available for all countries at regular intervals. These assessments require sponsors willing to bear their cost while making them freely available as public goods.

STRENGTHENING COLLABORATION

None of the indexes or tools cover the full scope of the GDB pillars or the stages of the DVC, but taken together and with recommended changes, they can provide an informative view of the current state of data systems and guidance for their development.

Recommendation:

This report has benefited from the work of the sponsors of the indexes and tools discussed here, including a productive webinar that discussed their purpose and role. Continuing this collaboration with country representatives and other stakeholders would provide a mechanism for the further development of these measures and efficient use of resources for implementation and sustainability. As with all efforts to increase the use and impact of data, a data collaborative dedicated to improving the measurement of data systems should include both the producers of tools and indexes, their users, and the intended beneficiaries of their use.

REFERENCES

Dang, H.-A., Pullinger, J., Serajuddin, U., & Stacy, B. (2021). *Statistical Performance Indicators and Index: New Tool to Measure Country Statistical Capacity*. Retrieved from **https://documents1.worldbank.org/curated/en/440191616164007723/pdf/Statistical-Performance-Indicators-and-Index-A-New-Tool-to-Measure-Country-Statistical-Capacity.pdf**.

European Data Portal. (2020). *Measuring Open Data Maturity*, sixth edition. Retrieved from **https://data.europa.eu/sites/default/files/method-paper_insights-report_n6_2020.pdf**.

Lafortune, G., & Ubaldi, B. (2017). *OECD OUEData Index 2017: Methodology and Results*. Paris: OECD. Retrieved from **https://www.oecd-ilibrary.org/governance/oecd-2017-ourdata-index_2807d3c8-en**.

PARIS21. (2020). *About Use of statistics index*. Retrieved December 3, 2021, from PARIS21 Statistical Capacity Monitor: **https://statisticalcapacitymonitor.org/indicator/127**.

PARIS21. (2021). *Measuring References to Statistics in National Policy Documents*. Paris: OECD. Retrieved from **https://paris21.org/sites/default/files/2021-05/PARIS21-paper_Measuring%20References%20to%20Statistics.pdf**.

World Bank. (2020). *Worldwide Governance Indicators*. Washington, DC: World Bank. Retrieved from **https://info.worldbank.org/governance/wgi/Home/Documents**.

ANNEX I.
EXAMPLE INDICATORS MAPPED ACROSS THE GDB PILLARS

PILLARS	EXAMPLE INDICATORS
DATA GOVERNANCE	• Existence of requirements to provide timely access to disaggregated data. (OURData Index) • The national open data strategy incentivizes the re-use of open data by both the public and private sectors. (European Open Data Maturity Assessment) • An organization operates an effective data sharing policy that provides guidance on the various ways in which data sharing should take place, from publication under an open license, through to the use of data sharing or processing agreements. (Joined-Up Data Maturity Assessment) • What is the legal and policy framework for the protection of personal privacy? (Open Data Readiness Assessment) • A regular exchange between the data stewards as well as data publishers and re-users is ensured. (European Open Data Maturity Assessment)
DATA CAPABILITIES	• Guidelines and/or tools are available to assist publishers in choosing the right type of license for their data. (European Open Data Maturity Assessment) • Performance incentives for public officials that implement open by default policies. (OURData Index) • To what extent is there an academic or research community which trains people with technical skills or has capabilities in data analysis? (Open Data Readiness Assessment) • How does the wider political context of the country help or hinder Open Data? (Open Data Readiness Assessment)
DATA AVAILABILITY	• Are the data published with comprehensive metadata? (Open Data Inventory) • Reliability of basic economic and financial statistics (e.g., national accounts, price indexes, foreign trade, currency, and credit). (Worldwide Governance Indicators) • Statistics are presented in a way that facilitates proper interpretation and meaningful comparisons (layout and clarity of text, tables, and charts). (Data Quality Assessment Framework) • Availability of high value datasets (as identified in the G8 Open data Charter). (OURData Index) • How many indicators and disaggregations are available across 22 data categories? (Open Data Inventory)
DATA USE AND IMPACT	• To what extent do policy documents incorporate and use data for decision making? (Use of Statistics Index) • A methodology to measure the impact of open data is in place or first steps in this direction are taken. (European Open Data Maturity Assessment) • Re-use of open data is monitored at national level via, for example, the national portal. (European Open Data Maturity Assessment)

ANNEX II.
EXAMPLE INDICATORS MAPPED ACROSS THE DVC STAGES

PILLARS	EXAMPLE INDICATORS
DATA COLLECTION	• Quality of data releases measured through adherence to SDDS/e-GDDS standards. (Statistical Performance Indicators) • Existence of formal requirements to consult stakeholders for data release. (OURData Index) • Percentage of harvested sources from total existing sources is known. (European Open Data Maturity Assessment) • What is the legal and policy framework for data security, data archiving and digital preservation? (Open Data Readiness Assessment) • Measures of the extent to which sources being used enable the necessary statistical indicators to be generated. (Statistical Performance Indicators)
DATA PUBLICATION	• Analytics tools are used to derive insights into users' behavior and needs. These insights are embedded into the portal update cycles. (European Open Data Maturity Assessment) • Data accessible free of charge and in open formats on the central/federal data portal. (OURData Index) • What data is already made available outside government - either free or for a fee - and on what conditions? (Open Data Readiness Assessment) • Documentation on concepts, scope, classifications, basis of recording, data sources, and statistical techniques is available, and differences from internationally accepted standards, guidelines, or good practices are annotated. (Data Quality Assessment Framework) • Externally published documentation is reviewed by key internal stakeholders before release. (Open Data Maturity Model)
DATA UPTAKE	• Existence of formal partnerships with businesses and the civil society to support data re-use. (OURData Index) • Do data have an open license or terms of use? (Open Data Inventory) • Which potential infomediaries (such as data journalists) are able to help translate Open Data into meaningful information for the public? What actions are needed to develop or enhance these parts of the Open Data Ecosystem? (Open Data Readiness Assessment)
DATA IMPACT	• A methodology to measure the impact of open data is in place or first steps in this direction are taken. (European Open Data Maturity Assessment) • Assessment or previous plans: How much analysis has been undertaken in the report on previous plans, showing use of data over time. (Use of Statistics Index) • The relevance and practical utility of existing statistics in meeting users' needs are monitored. (Data Quality Assessment Framework)