IMF POLICY PAPER

OVERARCHING STRATEGY ON DATA AND STATISTICS AT THE FUND IN THE DIGITAL AGE

March 2018

IMF staff regularly produces papers proposing new IMF policies, exploring options for reform, or reviewing existing IMF policies and operations. The following documents have been released and are included in this package:

- A **Press Release** summarizing the views of the Executive Board as expressed during its March 9, 2018 consideration of the staff report.

- The **Staff Report**, prepared by IMF staff and completed on February 9, 2018 for the Executive Board's consideration on March 9, 2018.

Electronic copies of IMF Policy Papers
are available to the public from
http://www.imf.org/external/pp/ppindex.aspx

International Monetary Fund
Washington, D.C.

IMF Executive Board Supports New Strategy for Data and Statistics in the Digital Age

On March 9, 2018, the Executive Board of the International Monetary Fund discussed the report, *Overarching Strategy on Data and Statistics at the Fund in the Digital Age,* supporting its forward-looking approach to gathering, processing, and sharing economic data and statistics.

The paper – which lays out the Fund's first-ever overarching strategy on data and statistics – comes at a critical time, amid a fast-changing data landscape, new data needs for evolving surveillance priorities, and calls from across the membership for assistance in further improving the quality of economic data.

"The strategy focuses on three "I"s: integration, innovation, and intelligence," said Louis Marc Ducharme, chair of the working group behind the report and head of the IMF Statistics Department. "*Integration* by bringing into alignment fragmented initiatives. *Innovation* by working to take advantage of Big Data. And *intelligence* by leveraging artificial intelligence for analyzing data and statistics."

The strategy encompasses six priorities: (i) agility in the identification of data needs; (ii) building the global data commons—an integrated network of member country websites publishing data essential for surveillance on a pre-announced schedule; (iii) supporting the use of Big Data and other innovations; (iv) securing seamless access and sharing of data within the Fund; (v) promoting the production of data that are comparable across countries; and (vi) addressing weaknesses in official data.

Executive Board Assessment[1]

Executive Directors welcomed the overarching strategy for data and statistics at the Fund. They noted that the strategy, which recognizes data as a strategic institutional asset, is both ambitious and pragmatic in setting a course to enable the Fund and its members to better respond to the challenges and opportunities of digitalization. The approach builds on the findings of the 2016 IEO report on data at the Fund, and is an important step forward in strengthening the linkage between data issues and other upcoming policy reviews.

---

[1] An explanation of any qualifiers used in the summing up can be found here: http://www.imf.org/external/np/sec/misc/qualifiers.htm

Directors noted that while generating new opportunities, the new data landscape also poses challenges. They concurred that, with surveillance priorities evolving, the need to analyze larger and more heterogenous amounts of data will require expanding the skills of staff. Directors emphasized that the expansion of the digital economy requires updating the methodologies and technology underpinning macroeconomic statistics. They agreed that the conduct of multilateral and macro-financial surveillance requires more cross-country comparable data and integrating new data sources.

Directors agreed with a major objective of the overarching strategy to integrate work streams on data provision by member countries, international statistical standards, capacity development, and data management. They supported the emphasis on innovation, including the potential for leveraging artificial intelligence in analyzing data and supporting member countries in the production of official statistics. Close engagement with the authorities on the benefits and potential resource and capacity requirements to implement the strategy was encouraged.

Directors welcomed the six strategic priorities, most notably, an integrated approach to prioritizing the Fund's evolving data needs. They supported the vision for establishing a global data common—an integrated cloud-based network of country websites publishing key data needed by the Fund and markets to monitor economic conditions and policies. Establishment of this framework is expected to decrease the burden of the authorities on data dissemination and provision, and publication of such data would serve as a public good supporting data transparency for the benefit of policy makers, investors, and the public. In this regard, Directors encouraged staff to explore possible synergies with other international organizations that receive and disseminate similar data.

Directors saw merit in exploring the use of Big Data to support earlier detection of risks and to complement the compilation of official statistics. While the effective use of Big Data is subject to several challenges, it also offers the potential for better understanding market developments, risks, and vulnerabilities. Directors noted that, in the absence of internationally accepted standards, exercising quality assurance will be necessary to ensure the sound use of Big Data, taking into account reliability and privacy concerns. In this connection, they recognized that the Fund could play a key role in facilitating peer learning across the membership.

On data management, Directors agreed that the nature of Fund activities requires accommodation of data diversity. Nonetheless, they saw scope to reduce duplication and increase efficiency. They agreed that the new governance structure, including the decision to move the Economic Data Team to the Statistics Department, will help simplify organization, increase synergies, and pool expertise. Directors attached high importance to promoting the compilation of data that are comparable across countries to support robust multi-country analyses and projections. Enhancing staff incentives with respect to data management functions could also strengthen implementation.

Directors agreed on the importance of addressing data weaknesses to strengthen surveillance and emphasized that the Fund should continue to build statistical capacity across the membership. They welcomed plans for a more structured and transparent assessment of data adequacy for surveillance, which should help the authorities and the Fund better prioritize capacity development and promote a meaningful dialogue on data issues. A few Directors cautioned against creating rankings of countries based on these assessments.

Directors indicated the need to implement the reforms in a budget-neutral manner, with trade-offs and re-allocation of resources carefully considered, as needed. They looked forward to receiving progress reports on the implementation of the strategy in due course.

INTERNATIONAL MONETARY FUND

OVERARCHING STRATEGY ON DATA AND STATISTICS AT THE FUND IN THE DIGITAL AGE

EXECUTIVE SUMMARY

The first data and statistics strategy for the Fund comes at a critical time. A fast-changing data landscape, new data needs for evolving surveillance priorities, and persisting data weaknesses across the membership pose challenges and opportunities for the Fund and its members. The challenges emerging from the digital revolution include an unprecedented amount of new data and measurement questions on growth, productivity, inflation, and welfare. Newly available granular and high-frequency (big) data offer the potential for more timely detection of vulnerabilities. In the wake of the crisis, Fund surveillance requires greater cross-country data comparability; staff and authorities face the complexity of integrating new data sources and closing data gaps, while working to address the weaknesses noted by the [IEO Report](#) (Behind the Scenes with Data at the IMF) in 2016.

The overarching strategy is to move toward an ecosystem of data and statistics that enables the Fund and its members to better meet the evolving data needs in a digital world. It integrates Fund-wide work streams on data provision to the Fund for surveillance purposes, international statistical standards, capacity development, and data management under a common institutional objective. It seeks seamless access and sharing of data within the Fund, enabling cloud-based data dissemination to support data provision by member countries (e.g., the "global data commons"), closing data gaps with new sources including Big Data, and improving assessments of data adequacy for surveillance to help better prioritize capacity development. The Fund also will work with policymakers to understand the implications of the digital economy and digital data for the macroeconomic statistics, including new measures of welfare beyond GDP.

The key elements of the overarching strategy are: integration, innovation, and intelligence: *integration* by bringing into alignment initiatives that are currently fragmented and unifying the data management function; *innovation* by working to take advantage of Big Data for higher-frequency monitoring, and deploying new technologies to close data gaps and meet surveillance needs, including through machine-to-machine data dissemination—particularly benefiting low-income countries; and *intelligence* by leveraging artificial intelligence for analyzing data and statistics.

The strategy encompasses six priorities:

- Agility in the identification of data needs;

- Building the global data commons—an integrated network of official websites publishing data essential for surveillance on a pre-announced schedule;

- Supporting the use of Big Data and other innovations;

- Securing seamless access and sharing of data within the Fund;

- Promoting the production of data that are comparable across countries; and

- Addressing weaknesses in official data.

To address data weaknesses—a crucial recommendation of the IEO—the Fund will continue to build statistical capacity across the membership, including with donor support. The new multi-donor vehicles (Data for Decisions and the Financial Sector Stability Fund) have been designed to ensure that capacity development is demand-driven and further aligned with surveillance priorities, including by supporting low-income countries and fragile states. The Fund will lead the work to promote cloud-based data dissemination through the broader implementation of the data standards initiatives and support its members in the use of non-traditional data sources (such as Big Data), further developing and applying methodological knowledge and experience for integrating data sources and for quality assurance.

The Fund needs the right skills to implement the overarching strategy. This will require fostering innovation, training, and retooling. Outside experts can supplement in-house expertise and bring new ideas. Access to data scientists will be needed to work on Big Data, and the forthcoming HR strategy presents an opportunity to find ways to bridge a talent gap.

To commence implementation of the overall strategy, Management has decided to:

- Reform the governance arrangements for data and statistics by broadening the purview of the responsible structure to cover all related work streams, including data management; and

- Move the Economic Data Team (EDT) from the Office of the Managing Director to the Statistics Department, thereby simplifying organization and facilitating wider synergies.

The strategy will aim to be broadly cost-neutral in the medium term, as implementation will involve reallocation of resources and exploitation of efficiencies. As the strategy fundamentally calls for integrating and reorienting long-standing frameworks and initiatives, the costs imputable to its vision are envisaged to be modest. Eventual actions to address the broader challenges of the digital world, beyond the scope of this paper, will emerge as thinking in this area is further developed.

CONTENTS

Glossary

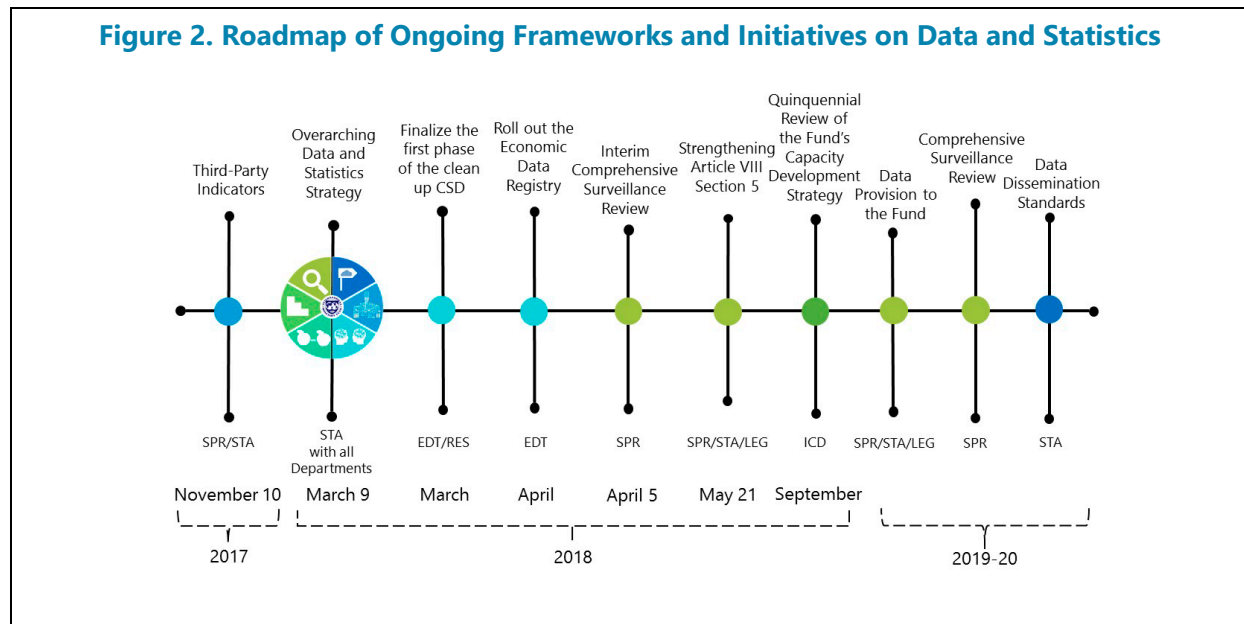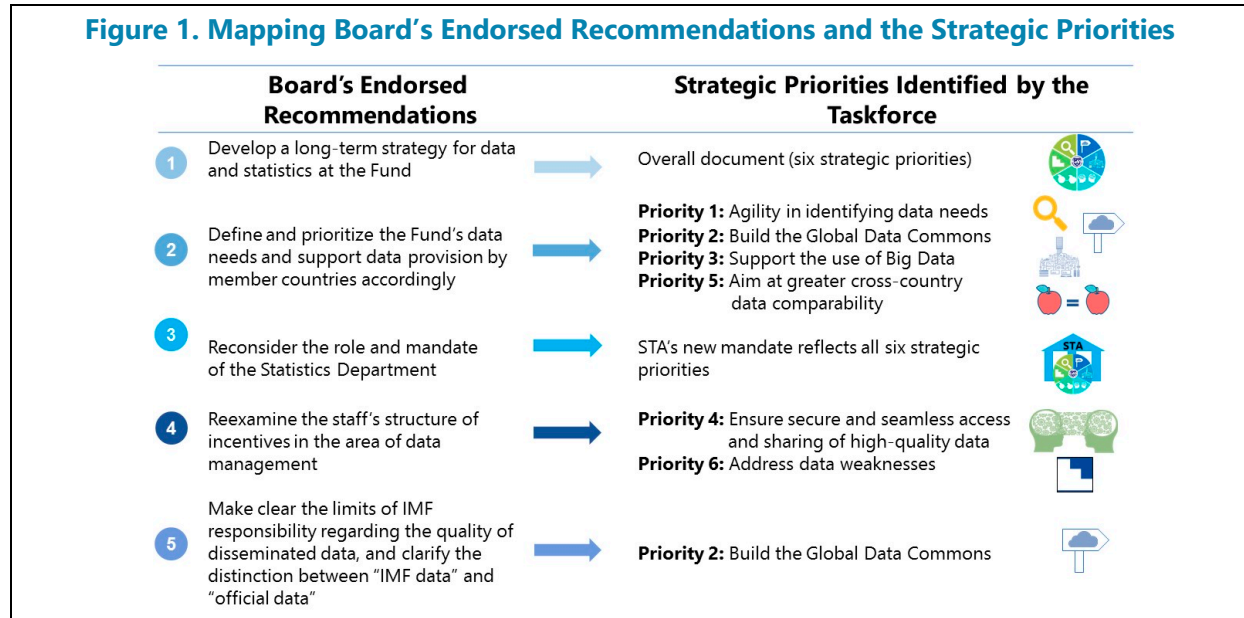|	
BIS	Bank for International Settlements
BOP	Balance of Payment Statistics
CSR	Comprehensive Surveillance Review
D4D	Data for Decisions
EDGG	Economic Data Governance Group
EDMI	Economic Data Management Initiative
EDSC	Economic Data Steering Committee
EDT	Economic Data Team
e-GDDS	Enhanced General Data Dissemination System
FAD	Fiscal Affairs Department
FSSF	Financial Sector Stability Fund
KMAC	Knowledge Management Advisory Committee
IADB	Inter-American Development Bank
IEO	Independent Evaluation Office
IFS	International Financial Statistics
iLab	Innovation Lab
LIC	Low-Income Country
MIP	Management Implementation Plan
NSDP	National Summary Data Page
OECD	Organization for Economic Co-operation and Development
ROSC	Report on Standards and Codes
SDDS	Special Data Dissemination Standard
SDDS Plus	Special Data Dissemination Standard Plus
SDMX	Statistical Data and Metadata Exchange
STA	Statistics Department
TCIRS	Table of Common Indicators Required for Surveillance
TSR	Triennial Surveillance Review
UFR	Use of Fund Resources
UN	United Nations
WEO	World Economic Outlook

INTRODUCTION AND OVERALL APPROACH

1.   Context. In November 2016, following the formulation of a Management Implementation Plan (MIP) to address the recommendations of the March 2016 IEO report—[Behind the Scenes with Data at the IMF](#), Management asked the Director of STA to chair an interdepartmental Taskforce to craft an overarching strategy on data and statistics.[1]

2.   This overarching strategy is intended to respond to the evolving data needs of the Fund and its members. It integrates multiple and inter-related domains covering work streams on data provision by member countries, international statistical standards, capacity development, and data management within the common institutional objective of ensuring that data are treated as a strategic institutional asset. The strategy also covers innovation, including the potential for leveraging artificial intelligence in analyzing data and statistics.[2] It maps the Board-endorsed recommendations as part of an integrated vision, with six strategic priorities (Figure 1).

3.   This strategy aims to guide future work on data and statistics, building on existing policy frameworks and ongoing initiatives (Figure 2). It sets the foundation, for example, for a well-integrated approach for treating data matters in the upcoming papers on Strengthening Article VIII, Section 5 (2018), the Review of the Fund's Capacity Development Strategy (2018), the Comprehensive Surveillance Review (2019), the Review of Data Provision to the Fund for Surveillance Purposes (2019), and the Review of the Fund's Data Dissemination Initiatives (2020). It takes account of the rapidly-evolving data landscape and the importance of technology to enable the Fund and its members to meet evolving data needs in an increasingly digitalized economy.

4.   Data issues at the Fund. The IEO evaluation found that data provision from member countries has improved markedly over time, allowing the IMF, to a large extent, to keep abreast of the growing complexity and interconnectedness of the world economy. However, the evaluation concluded that data problems have, at times, adversely affected the Fund's surveillance and lending activities. In the aftermath of crises, data often have been put at the forefront, prompting important changes in the Fund's approach to data. Yet, once crises subside, data issues typically receive less attention and are viewed as mere support activities to strategic operations. The roots of these problems are diverse, ranging from external (e.g., country capacity constraints) to internal (e.g., lack of appropriate staff incentives, entrenched work practices). While recognized for decades, most of these problems are now cast in a new light due to the unrelenting proliferation of data sources, technological advances, and the surging demand for multilateral and financial surveillance and

---

[1] Staff engaged also with external shareholders during a Hackathon on the Strategy on Data and Statistics—a crowdsourcing retreat held at end-February 2017. The Hackathon participants (about 80)—including Taskforce members, private-sector representatives (e.g., Google, Haver, Bloomberg, IBM), IFIs (BIS, IADB, Eurostat, and the World Bank Group), and academia—were asked to identify innovative solutions to the Fund's data and statistics challenges as input to the development of the overarching strategy.

[2] Artificial intelligence is a data-intensive technology that emulates human performance by machine learning and data mining to automate execution of routine tasks and extract information from large amount of data.

cross-country analyses. This long-term overarching strategy takes a holistic approach to addressing the fragmentation of the multiple frameworks and initiatives on data and statistics, and sets as a fundamental common objective the recognition of data as a strategic institutional asset.

Figure 1. Mapping Board's Endorsed Recommendations and the Strategic Priorities

Board's Endorsed Recommendations	Strategic Priorities Identified by the Taskforce
1 — Develop a long-term strategy for data and statistics at the Fund	Overall document (six strategic priorities)
2 — Define and prioritize the Fund's data needs and support data provision by member countries accordingly	**Priority 1:** Agility in identifying data needs **Priority 2:** Build the Global Data Commons **Priority 3:** Support the use of Big Data **Priority 5:** Aim at greater cross-country data comparability
3 — Reconsider the role and mandate of the Statistics Department	STA's new mandate reflects all six strategic priorities
4 — Reexamine the staff's structure of incentives in the area of data management	**Priority 4:** Ensure secure and seamless access and sharing of high-quality data **Priority 6:** Address data weaknesses
5 — Make clear the limits of IMF responsibility regarding the quality of disseminated data, and clarify the distinction between "IMF data" and "official data"	**Priority 2:** Build the Global Data Commons

Figure 2. Roadmap of Ongoing Frameworks and Initiatives on Data and Statistics

HOLISTIC VISION FOR DATA AND STATISTICS: INTEGRATION, INNOVATION, AND INTELLIGENCE

5.   **The strategy seeks to foster integration, innovation, and intelligence.** It addresses the problems with the status quo and offers an integrated vision that incorporates the opportunities and challenges emerging from the fast-changing data landscape. The goal is to move to a new ecosystem of data and statistics that allows secure and seamless access and sharing of data within the Fund, leverages data dissemination to support countries' provision of data, closes data gaps with new data sources (including Big Data), and promotes more transparent and candid assessments of data adequacy for surveillance (Box 1) to better support the membership and guide prioritization of capacity development.

6.   ***From Fragmentation to Integration.*** The current Fund universe of data and statistics is multi-dimensional and fragmented. While aware of the silo mentality in data management practices, the staff typically has only partial understanding of the multiple streams of work on data and statistics associated with the governing policy frameworks, several of which are periodically reviewed by the Board. As initiatives are generally not well integrated, the resulting fragmentation hinders efficiency, the exploitation of synergies, and the effectiveness of the Fund in discharging its duties, including surveillance. To overcome this fragmentation, the strategy points to potential synergies among different initiatives (e.g., tightening the interaction between the comprehensive surveillance review, capacity development strategy review, the review of data provision to the Fund, the G20 data gap initiatives, and the review of the Fund's data standards initiatives—Box 2). It also identifies gaps in the current legal frameworks and offers a high-level sense of direction on prospective changes.

7.   ***From Integration to Innovation.*** Staff has sought to craft a strategy that is resilient to change—sufficiently flexible to adapt to a changing environment to better prepare the Fund for the future. As the digital and IT revolution is generating vast amounts of data and new ways to distill signals, shifts are taking place in both the supply (official, third-party, commercial) and demand of data necessary for sound policymaking. Enhanced data availability could strengthen the Fund's core operations by exploiting new methods and Big Data for earlier risk detection and closing data gaps.

8.   ***From Innovation to Intelligence.*** Big Data and artificial intelligence (AI) will allow extraction of signals, including through data mining, to elicit patterns and inform economic analysis.[3] While Big Data can be an asset that is difficult to exploit, AI may unlock its value, including through machine learning. Big Data and AI have the potential to help the Fund better tailor policy advice to member needs. This will require investment, as reaping the full potential of data (text) mining requires a platform for AI and cloud data storage. Enhancing the Fund's cyber resilience will also be important.

---

[3] Data (text) mining is the process of discovering meaningful data correlations, patterns, and trends by sifting through large amounts of data (text) in repositories through pattern recognition technologies and statistical and mathematical techniques.

A. Six Strategic Priorities for Data and Statistics

9.   The daily volume of data now created globally is unprecedented. Estimates are for worldwide digital data to increase by 45 times by 2020 relative to 2009. Most come from the digital economy, as the number of internet users has grown by more than a billion over the last five years. The amount of information produced by machines will account for about 10 percent of global data, with one-third to be stored in the cloud.

10.   While generating new opportunities, the new data landscape is also posing challenges. With surveillance priorities evolving, the need to analyze more heterogenous and larger amounts of data will require expanding staff skills. Core Fund activities require economists to work intensively and extensively with official and other data (commercial and non-structured), including to provide micro-foundations to policy advice. However, absence of internationally-accepted standards governing third-party indicators recently raised concerns, and copyright issues pose a challenge. Also, consideration of an open data policy to increase transparency further requires addressing questions about the confidentiality expected of a trusted advisor.[4] The expansion of the digital economy requires sharpening and updating methodologies as new metrics are needed—beyond GDP—to better capture welfare. The increasing role and complexity of multilateral and macro-financial surveillance have generated new data demands, including integrating new data sources (e.g., micro data, balance sheets, bank and firm-level data). Big Data is subject to several challenges, but provides untapped opportunities to better understand markets, risks and vulnerabilities.[5] While capacity development has helped close data gaps, new needs require use of commercial data, especially in flagships, with the consequent increase in the budget for commercial databases.

11.   Against this background, the strategy seeks to address the fragmentation of the data and statistics landscape in a framework consisting of six strategic priorities, which reflect a considered view of *where* the universe of data and statistics is heading. Addressing each priority involves recommendations (i.e., *what* the Fund should do to meet the new data challenges and tap new opportunities) and concrete actions (i.e., *how* the strategy can be implemented—Figure 3). The six strategic priorities are to:

1.   be agile in identifying data needs for effective surveillance;

2.   build the global data commons—a cloud-based network of official websites publishing data essential for surveillance according to a pre-announced schedule;

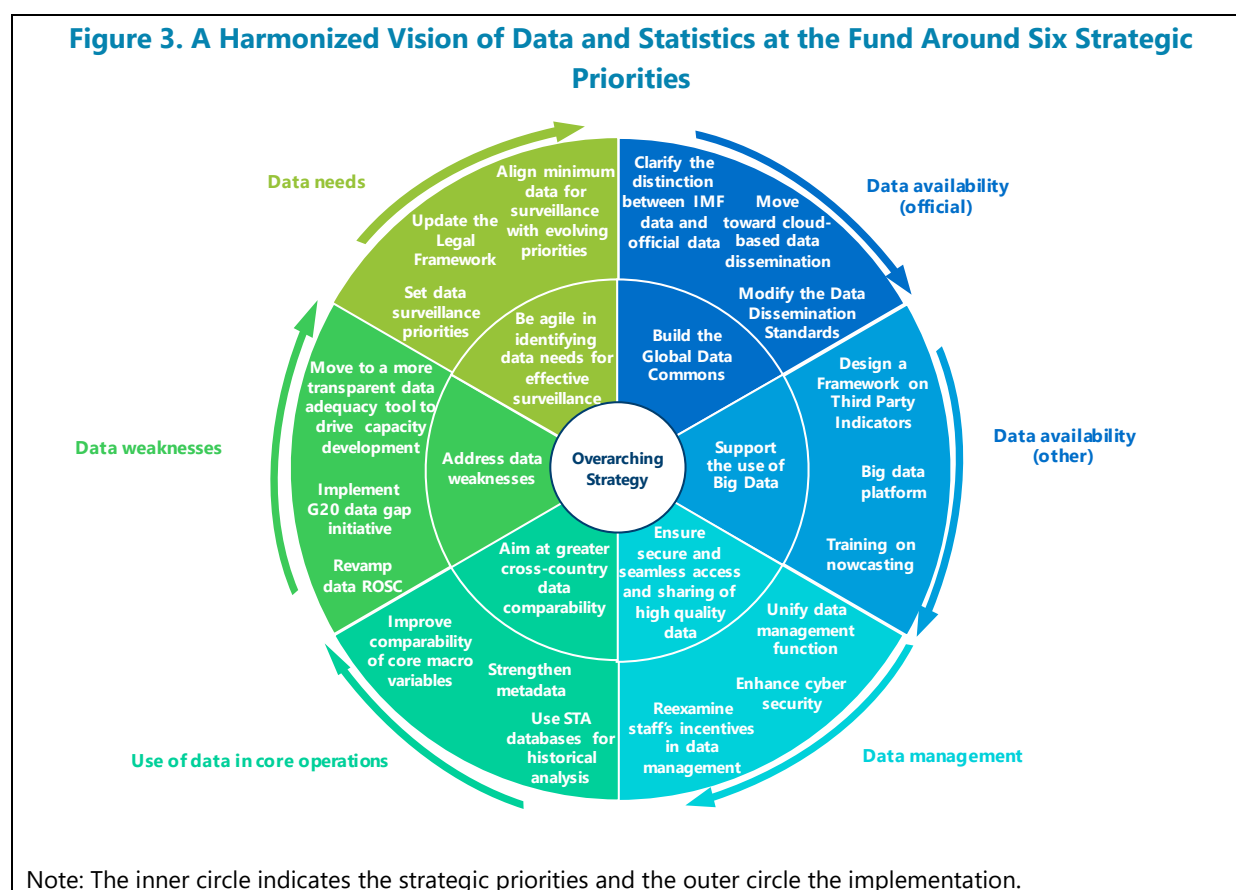3.   support the use of Big Data in Fund operations;

---

[4] An open data policy implies easy, universal access to most of the Fund's operational data and the data underlying research and other publications.

[5] Challenges include: methodological, such as selection bias; analytical, in extracting information from noise; and data management, in processing and storing the large amount of data.

4.   ensure secure and seamless access and sharing of high-quality data within the Fund;

5.   aim at greater cross-country data comparability to better integrate bilateral and multilateral surveillance; and

6.   address data weaknesses by integrating surveillance priorities with capacity development.

12.   The strategic priorities correspond to six dimensions of data and statistics at the Fund. These dimensions are elements of a dynamic and organic process that involves all Fund departments. The process includes: the data needs of the Fund; which data are available and how they are made available; whether the data are provided to the Fund by members or come from commercial sources; how staff manages the data; how staff uses data in surveillance and other operations; and how data weaknesses are identified and addressed. The different parts of this framework interact with each other but share one goal: at each stage of the data value chain, the data are treated as an institutional asset of the Fund.

Figure 3. A Harmonized Vision of Data and Statistics at the Fund Around Six Strategic Priorities

Note: The inner circle indicates the strategic priorities and the outer circle the implementation.

STRATEGIC PRIORITY 1: AGILITY IN IDENTIFYING DATA NEEDS

A. Framing the Issue: A Changing Data Landscape

13.   Data and statistics needs continue to evolve. A key lesson from the global financial crisis is that in a highly-interconnected world, Fund surveillance must be agile in adapting to emerging challenges. Recent triennial surveillance reviews (TSRs) have focused on integrating and deepening the analyses of risks and spillovers in bilateral and multilateral surveillance. To support this work, the 2014 TSR called for further efforts to address data gaps, including through implementation of the G-20 Data Gaps Initiative. New analytical tools, along with a widening scope of risk analysis, have further expanded data needs, such as for balance sheet and firm-level data.

B. Looking Ahead

14.   An integrated approach to the frameworks that prioritize the Fund's evolving data needs for surveillance is warranted. The Comprehensive Surveillance Reviews (CSR)—which entail consultation with internal and external shareholders—and interim reports will continue to guide the evolving surveillance priorities. The 2018 Interim Surveillance Review will be an opportunity to consider how other strategic work at the Fund—including on data and statistics—can help position the Fund to continue to adapt to new challenges. Data and statistical needs will be set and operationalized in the Reviews of Data Provision to the Fund for Surveillance Purposes (the next is scheduled for 2019). The review will also consider how to better use publicly available data and how to leverage data produced by other institutions. Work on statistical methodologies, dissemination standards, and capacity development will be aligned with these priorities to ensure effective support for Fund operations. To strengthen the integration of capacity development in surveillance, staff will also propose, at the time of the upcoming Review of Data Provision to the Fund for Surveillance Purposes, a new diagnostic tool to promote a more transparent and structured assessment of data adequacy for surveillance (Section on data weaknesses—Priority 6).

15.   Given the evolving challenges and data needs, the data and statistics frameworks should be more evolutionary, flexible, and resilient to changes in the data landscape. The list of minimum data required for Fund's core activities should evolve with emerging data needs.[6] This could be addressed by aligning the data standards initiatives (review scheduled for 2020) more closely with the information requirements for surveillance, and harmonizing the Table of Common Indicators Required for Surveillance (TCIRS) with the list of the minimum data required for the Fund's core activities.

---

[6] Some data categories listed as minimum required in Article VIII, Section 5, have been superseded by innovation in statistical methodology (e.g., CPI in lieu of the retail price index).

16.   The forthcoming Board paper on Strengthening Article VIII, Section 5, will consider some of these issues, clarifying members' requirements in terms of data provision—including cases of delayed provision of data to the Fund and what constitutes data provision to the Fund (direct transmittal and/or publication).[7] Traditionally, members have provided the required information directly to staff. With advances in technology and the general move toward transparency, however, some members publish information required by the Fund on government websites (MOF, central bank, statistical agency). While this can be considered "provision" of information, as the data are accessible to Fund staff, it poses complications in that staff may not be aware of the availability of the information or its date of publication. Staff will consider how it could be possible to adapt its operational approach under the Article VIII, Section 5 framework to this reality. Staff will also explore how data dissemination standards could be further used to support members in providing data to the Fund and to the global community (Priority 2).

17.   Staff will continue to focus the methodological work on evolving data priorities. Such work currently aims to help policymakers deepen their understanding of the digital economy (in line with the Global Policy Agenda), and to apply digital data to the compilation of macroeconomic statistics. Expanding data sources, innovation, and technical progress often complicate statistical measurement. Staff will update statistical methodologies to ensure that the range of statistical frameworks (national accounts and prices, government finance statistics, external sector statistics, and monetary and financial statistics) consistently account for digitalization and other innovations. In a digital world, GDP may be too narrow an indicator, measuring only the production of marketable goods and services. With some member countries already using alternative indicators of prosperity and well-being, new metrics (beyond GDP) that take account of distribution and nonmarket aspects of economic welfare may be needed to enhance surveillance.

18.   In line with Fund-endorsed strategic priorities, the Fund will continue to support mainstreaming macro-financial surveillance, including balance-sheet analysis. To this end, staff will mainstream and strengthen the balance sheet tool featuring assets and liabilities across the five key economic sectors of the economy.

STRATEGIC PRIORITY 2: THE GLOBAL DATA COMMONS

A. Framing the Issue

19.   Data dissemination is a public good that promotes transparency, lowers borrowing costs, and should be harnessed to better support data provision to the Fund. There is scope for Fund efforts that encourage country publication of key data to better support members in the provision of data to the Fund for surveillance purposes, thereby streamlining the reporting burden

---

[7] The Articles do not specify the periodicity and timeliness for the reporting of data categories under Article VIII, Section 5. Members are required to compile information on a regular basis in as up-to-date a manner as possible and to provide the Fund with such data as soon as it becomes available.

on the authorities. During the discussion of the MIP, the Board cited the need to lessen the burden of new data requests and encouraged members to adopt international dissemination standards.

B. Looking Ahead

20.   The strategy calls for an integrated network of country websites under the IMF's data dissemination initiatives—the *global data commons*. Member commitments under the data standards initiatives entail regular publication of selected data—according to an agreed advance release calendar—to give certainty to investors, rating agencies, and markets. While nearly 100 countries maintain such websites in keeping with commitments to implement the IMF's data dissemination initiatives, fully integrating a universal network will require upgrading the technology underpinning the websites in many countries as well as expanding the number of countries committing to publish the data under the Special Data Dissemination Standard (SDDS), SDDS Plus, and the enhanced General Data Dissemination System (e-GDDS).

21.   To this end, staff will help countries deploy the cloud-based technology and "machine-to-machine transmission" already used by those adhering to the SDDS Plus (14) and implementing the e-GDDS (26). This will entail establishing open-data platforms (and a specialized data transmission system—SDMX) that permit voluntarily moving from the current "push" system for data reporting (in which countries send their data to the Fund and other parties) to a "pull" system (where countries upload their data to a single web-based repository, allowing all users—including the Fund—to draw on the data). Completing the global data commons will require upgrading the technology underpinning the official website (National Summary Data Page—NSDP) in all SDDS countries and implementing the recommendations of the e-GDDS in about 80 countries. Such an upgrade is not expected to have major resource implications for the authorities.

22.   The next review of the data standards initiative should facilitate the upgrading of technologies supporting data dissemination to underpin the global data commons. To implement this vision, the Fund will need to redouble efforts to achieve universal implementation of the data dissemination initiative, which will also facilitate direct transmission of data, upon release, to country teams. While all members will benefit from the global data commons, there will be even greater gains for low-income countries, and small and fragile states where streamlining data requests would ease capacity constraints while also allowing the Fund to conduct real time surveillance beyond the time of the Article IV consultation. Construction of the global data commons should take about five years, as staff works with authorities to accelerate the implementation of the e-GDDS and upgrade the SDDS infrastructure. This network will publish—as a public good—data for surveillance in a predictable and disciplined manner for the benefit of the authorities, the Fund, rating agencies, and the markets. Moreover, it will allow countries voluntarily to transmit confidential data—including on UFR—directly to the Fund through a secure channel.

23.   The global data commons will promote publication of the data used by the Fund for surveillance. Country provision of data under Articles IV and VIII will continue to be an essential obligation of Fund membership. But as the world economy has changed since the 1940s, it now makes more sense to promote *publication* of such data, including because the Fund is no longer alone in monitoring and assessing country economic conditions and policies. Thus, the Fund needs

to dynamically align requirements under the data dissemination initiatives with those set in the reviews of data provision to the Fund for surveillance purposes. The Tenth Review of Data Standards, scheduled for 2020, will seek to achieve this objective. The variables under the e-GDDS should continue to be aligned with the TCIRS, and the SDDS and SDDS Plus should be aligned with the evolving data needs as reflected in other strategic Fund priorities, as envisaged in the MIP.

24.   **With greater emphasis on dissemination, the distinction between IMF data and official data merits clarification (in line with recommendation 5 of the MIP).** In this connection, it bears clarifying that the Fund does not "endorse" the data that appear in its publications, databases, or the Dissemination Standards Bulletin Board. While the WEO and IFS flagships and underlying databases distinguish between IMF staff estimates and official data, Article IV reports often do not. This will need to be rectified, and a disclaimer added to the Dissemination Standards Bulletin Board.

STRATEGIC PRIORITY 3: SUPPORTING THE USE OF BIG DATA THROUGH STATISTICAL INNOVATION

A. Framing the Issue

25.   **In endorsing the MIP, the Board asked that the strategy place the Fund at the forefront of data and statistical innovation.** Over the last year, the Fund has led innovative work on new statistical methodologies to support surveillance and policymaking. This work included Big Data, the digital economy (in line with the Global Policy Agenda), and work on the use of third-party indicators—in keeping with the notion that statistical innovation and non-official data can bridge data gaps. As the Fund moves to consider macro-criticality beyond the traditional areas covered by macroeconomic and financial statistics, innovations may facilitate analyses through use of commercial data, including Big Data. Enabling the use of Big Data in surveillance would sharpen the Fund's ability to deepen risk analysis to support its members in line with the recommendations of the 2014 Triennial Surveillance Review. Big Data is increasingly used by country authorities and multilateral agencies to provide real-time, high-frequency information (Box 3).

26.   **Big Data can benefit member countries and improve policymaking by**:

- acquiring real-time information about traditional and emerging macro-relevant issues, with more timely and comprehensive signals on emerging risks;
- closing data gaps;
- improving member and Fund staff estimates of economic and financial variables;
- helping identify progress in attaining the Sustainable Development Goals in low-income countries (LICs); and
- improving the timeliness of official statistics, especially in LICs.

B. Looking Ahead

27.   **New technology and tailored methodological advice could allow LICs to leapfrog in their statistical development efforts.** Statistical systems must be adapted to reflect changing

needs and technological progress, and new data sources afford the potential for improving data compilation. Big Data, for example, will complement traditional data sources and help statistical agencies overcome the limitations of traditional surveys—such as low and declining response rates, overburdening respondents, and lack of flexibility. Use of Big Data can improve the accuracy and timeliness of estimates relative to traditional survey-based data. For example, use of mobile phone data on travel services could be more cost-effective than border surveys for external sector statistics, a potentially valuable option for countries with heavy reliance on tourism and remittances.

28. **Supporting the use of Big Data in Fund operations requires statistical expertise in integrating data sources.** Such integration is to be carried out by collaborating and sharing expertise with other institutions. The recent Staff Discussion Note on [Big Data](#) (SDN/17/06) explained that statistical techniques can be used to extract signals from unstructured data for policy analysis, and that the statistical implications of Big Data are likely to be substantial, including for data quality.

29. **Quality assurance is needed for sound use of Big Data.** The absence of internationally accepted standards governing third-party indicators recently raised concerns, requiring statistical guidance on how area and functional departments should use these indicators. Concerns relate mainly to potential surveillance risks arising from weaknesses in the underlying methodology that would affect the credibility of derived country rankings and the soundness of policy advice. The potential for reputational risk is associated with the misperception of endorsement of third party indicators. To mitigate this risk, the Board endorsed recently a new framework based on the IMF's Data Quality Assessment Framework.

30. **Over the medium term, a new policy should guide the integration of different data sources**. It should guide the procurement and use of Big Data, set quality controls, and train Fund staff on the new data sources—all while building partnerships with member countries, other IFIs, and the private sector. The implications of advances in technology, including new data sources for surveillance will be discussed as part of the 2019 Comprehensive Surveillance Review.

31. **Staff work over the medium term will address the following:**

- **First, enhancing learning/stock-taking and knowledge sharing on the use of Big Data in analyses.** This includes learning from country authorities (especially leading statistical offices) and facilitating peer learning across the membership.[8]

- **Second, addressing skill gaps.** The need to absorb and analyze a large amount of data using new tools will require *re-tooling* staff skills through training on techniques, tools, applications and on the methodological implications of integrating new data sources. The forthcoming HR strategy presents an opportunity to assess emerging skill needs, including for those of data scientists, and to strike a balance in recruitment decisions that combines statistician, methodologist, fungible economist, data scientist, mathematician, and IT staff. Enhancing

---

[8] To this end, the Fund is partnering with Statistics Netherlands and the United Nations (UN) Statistical Division on Big Data, including to take stock of countries' experimentation and emerging best practices.

training on communication and interpersonal skills—to complement analytical skills—will be essential for fostering partnerships among professionals with different specializations.

- **And third, supporting the use of Big Data in Fund present operations by:**

  ✓ *Integrating data sources*. Big Data will not be a substitute for official statistics but rather a complement, improving the production of official statistics and enriching the Fund's databases;

  ✓ *Undertaking quality control*. In view of the methodological challenges arising from the sample bias intrinsic in Big Data, the Fund will need to assess fitness for purpose, in the spirit of the recent work to guide use of third-party indicators;

  ✓ *Tailoring TA*. In building capacity, the Fund could sponsor visits of staff of member countries' leading statistical offices that pioneer innovation to countries interested in learning to use Big Data to produce official statistics. Best practices learned by pioneering compilers can be used to customize TA by the Fund, which also will leverage its ongoing TA work on administrative data; and

  ✓ *Innovating by building partnerships*. The Fund should develop new partnerships with national authorities, other international agencies, and global private firms, including to strengthen data sources as requested by Directors during the informal brief on the SDN on Big Data.

STRATEGIC PRIORITY 4: SECURE SEAMLESS ACCESS AND SHARING OF HIGH-QUALITY DATA

A. Framing the Issue: Current Ecosystem of Data and Statistics

32.   The current ecosystem of data and statistics at the Fund is complex and fragmented, complicating data management. With different and clearly specified mandates, each Fund department group (area, functional, and support) strives to gather data with the characteristics that best support its specific operations. Thus, multiple paths of data collection are necessary, with operational data flowing from the authorities to most country desks, while official data complying with the more stringent methodological requirements for cross-country comparability flowing (at a slower-than-desired pace) to STA. The trend toward greater use of commercial data with *new species* (Big Data and third-party indicators) that fall outside the well-established methodological guidance followed by official data poses new challenges, including budgetary and legal (such as copyrights).

33.   The complexity of the current eco-system is largely organic and reflects the nature of Fund activities, which implies tolerance for data diversity.

- *Operational data*: Fund surveillance is based on a policy dialogue tailored to the specific circumstances and domestic realities of each member country. Country teams obtain the data utilized by the authorities for policy purposes ("operational data") which, once inside the Fund, is used to produce macroeconomic scenarios involving five-year projections that are shared internally for the preparation of the World Economic Outlook (WEO) and for populating the Common Surveillance Database. Owing to considerable differences in countries' statistical capacity, institutional organization, policy arrangements, and other idiosyncrasies, the historical

series that underpin projection scenarios are not always comparable across countries (see Priority 5 below on the need to achieve greater cross-country comparability), with the methods used by country teams to produce five-year projections also different across countries.

- *STA databases:* Under the Articles of Agreement (Article VIII, Section 5(c)), the Fund is to act as a "*center for the collection and exchange of information on monetary and financial problems.*" In this context, the Fund—largely through STA—collects and disseminates data, helps build members' statistical capacity, and establishes international best practice for the compilation and dissemination of macroeconomic and financial statistics. As progress in improving statistical data takes time, particularly in countries with lower capacity and income level, compilation of higher-quality data proceeds slowly, which is reflected in less-than-universal coverage by STA databases. Moreover, data published in STA databases are less timely than those provided to country teams (which obtain data often before public release). Staff often are unaware of the difference in collection goals: timely, country-specific data for the policy dialogue and real-time surveillance (operational data) versus compilation of data that are cross-country comparable (specialized STA databases).

34.   Nevertheless, there is scope to reduce duplication and increase efficiency. Fragmentation of the data management function has hindered exploitation of economies of scale, leading to some redundancy between operational databases and some STA databases, which need to be addressed. Inadequate sharing of data across the Fund at times has resulted in duplication, with more data seen to substitute for the effective use of available but difficult-to-access data. Multiple platforms also have complicated effective Fund-wide data management.

35.   The Fund has taken important steps to improve data management. The 2011 Economic Data Management Initiative (EDMI) Report underscored the fragmented approaches to data management. In response, a new data management governance was established in 2012, creating the Economic Data Steering Committee (EDSC) and the Economic Data Governance Group (EDGG) as key anchors of the reform process. The EDSC was tasked with delivering strategic direction on data management matters, with member departments given a voice in data management decisions. The EDGG is responsible for operational issues, and facilitation for these two structures is provided by the Economic Data Team (EDT)—a unit accountable to the Office of the Managing Director.

36.   The EDSC has guided the design and launch of initiatives to improve data management in area departments and to improve the sharing of operational data across the Fund. Country databases have been transitioned to a structured format and aggregated into the Common Surveillance Database, which serves as a repository of all desk operational data. Nevertheless, problems in data management have persisted, partly because of inadequate staff incentives for proper data management.

B. Looking Ahead

37.   A more integrated ecosystem of data and statistics at the Fund is needed. It should facilitate seamless, frictionless access and sharing of high-quality data, enhancing the Fund's ability to conduct surveillance and offer sound policy advice. Management's decision in January 2018 to give STA primary responsibility for data management by integrating the office of the Economic Data

Team into STA is intended to achieve a more integrated approach to data management. Building on the gains and progress achieved so far, this integration will facilitate identification of opportunities to streamline, simplify, and economize on data collection. It also will help identify new opportunities to integrate data sources to support Fund analyses and projections.[9]

38.   This increasingly integrated data ecosystem must be supported by a cultural change with incentives for staff that ensure improved data management, and by strengthening data sharing and security. This will require:

- Cultivating a corporate culture that recognizes data as an institutional asset, attaching institutional priority to the thorough knowledge of country data and good data management. Flexibility will be essential for teams to balance competing operational priorities in a manner consistent with a reasonable workload and the more urgent surveillance imperatives;

- Reshaping incentives and training to more effectively meet current and future data management needs, including by giving more weight to data management skills in the APR process as well as enhancing knowledge transfer on data;

- Continuing to strengthen Fund-wide data sharing by further improving the Common Surveillance Database and rolling out the Economic Data Registry as a single access point for data sharing of *internal and external databases.* The Registry is a business intelligence tool and a search engine to become the Fund's primary tool for searching and accessing internal and external (including commercial) data;

- Adopting an open data policy consistent with the Fund's confidentiality policy, affording electronic access to official data and metadata underpinning tables in Article IV reports and IMF Working Papers; and

- Making further progress in protecting the Fund's more sensitive data and strengthening cyber security.

39.   Digitalization and technology offer opportunities to improve data collection and publication:

- *Timeliness*. Improved timeliness will be achieved by leveraging SDMX technology to "pull" data directly from national websites, as soon as released. Thus, databases are expected to accommodate more data revisions in the future, as country compilers correct classification and other methodological errors detected by Fund experts;

- *Moving to digital dissemination.* Continued technological progress eventually should permit the orderly discontinuation of the hard-copy version of flagship data publications and a move to tailored digital dissemination. Wider dissemination of data-pulling technologies eventually

---

[9] About 60 country teams already use monetary data from STA databases and STA data are used for benchmarking BOP data and government debt data supplied by desks to the WEO, with FAD also using STA data to complement desk data on tax revenue and functional expenditure.

should allow commercial users (re-sellers of IFS data such as Haver and Bloomberg) and others (including in academia) to pull the data from the global data commons; and

- *Facilitating access.* This will require considering partnerships with the private sector to secure more efficient processing of data processing and distribution (e.g., with Haver and Bloomberg).

STRATEGIC PRIORITY 5. GREATER CROSS-COUNTRY DATA COMPARABILITY

A. Framing the Issue

40.   The increasing importance of multilateral surveillance and cross-country analysis implies a greater need for data that are comparable across countries. As stated in the IEO report, *"Multilateral surveillance (and cross-country analysis, more generally) poses a special challenge for data as it is predicated on data being comparable across countries."* An evaluative survey of external data users by the IEO indicated that almost 90 percent believed IMF data are comparable across countries—a misperception that could pose a reputational risk.

41.   Achieving *full* cross-country data comparability across all series and databases is to remain a dynamic aspiration, including because rates of speed of statistical development vary. The IEO noted that "*The fundamental data dilemma for multilateral surveillance is international comparability versus country specificity" (paragraph 26 of the IEO report).* Country desks generally use data as reported by the authorities (operational data) to ensure that staff and the authorities "speak the same language," reflecting domestic policy and institutional realities. The data are then integrated into the WEO and Common Surveillance Database, generally following national definitions, and methodologies that also could differ due to varying levels of statistical capacity. Even when operational data meet the latest international standards, coverage and scope may differ owing to differences in countries' institutional arrangements. On the other hand, official data reported to STA are more cross-country comparable, although less timely and covering fewer countries. According to the IEO: "*The IMF's work on methodology and capacity development in the area of statistics has gone a long way to strengthen comparability. This is particularly true for STA's databases, with an emphasis on data that meet methodological standards. But the main sources of data for much of multilateral surveillance are area departments, where data are more likely to conform to country specificities."*

B. Looking Ahead

42.   The strategy calls for addressing the strong demand for cross-country data comparability identified in the Interim Surveillance Review. As staff and external users may not be fully aware of comparability nuances, staff will strengthen metadata across databases to help users identify potential cross-country differences in methodology or coverage of variables. Specialized STA databases with comparable data will continue to support cross-country analytical work, with the Fund to promote further the adoption of internationally-agreed methodologies to allow production of statistics by the authorities that are inherently cross-country comparable.

43.   Staff will focus on focus on working with member countries to improve comparability for at least a set of key macroeconomic variables. In the next review of the data standards initiatives, planned for 2020, countries should be encouraged to provide more comparable data as part of their commitment to the SDDS/e-GDDS. The same review could assess countries' progress in producing comparable data for at least a set of variables required for effective Fund surveillance.

STRATEGIC PRIORITY 6: ADDRESSING DATA WEAKNESSES

A. Framing the Issue

44.   Data shortcomings and gaps across the membership hamper policymaking, Fund surveillance and program work, including projections and policy advice. Reliable, timely and internationally comparable macroeconomic and financial statistics are essential for policymakers and IFIs as well as for private users. To promote production of high-quality data, the Fund's capacity development work has grown by more than 70 percent over the last six years, with funding from bilateral and multilateral partners.[10] In discussing the MIP, the Board called for strengthening incentives for a more candid assessment of data deficiencies, following the IEO's finding that country teams—in some cases—fall short of a candid assessment of severe data weaknesses in order to avoid undermining the basis of their analyses and relations with the authorities.

B. Looking Ahead

45.   Addressing data weaknesses hindering surveillance requires a more candid assessment of data adequacy for surveillance, a key responsibility of country teams. To this end, staff will develop a tool to guide country teams to prepare the assessment in a transparent and structured manner. By using an assessment tool to reveal data weakness as part of the Article IV consultation, country teams will be better able to discuss with the authorities the implied uncertainties for the analysis and policy advice. A short questionnaire will help highlight how data weaknesses affect the quality of analysis, thereby facilitating the prioritization of TA, training, and the more comprehensive statistical quality assessment conducted by data ROSC missions, if warranted.[11] Use of the tool should mitigate the tension between frankness about data weaknesses and traction of the policy advice, with an account of the discussion with the authorities to be reflected in the outlook and risk sections of the staff report when data weaknesses impose considerable uncertainty in staff analyses, projections, and policy recommendations.

---

[10] Current statistical capacity development priorities include closing data gaps, improving data quality, and broadening data dissemination to help detect economic vulnerabilities and improve economic decision-making. See Annex II of 2018 Quinquennial Review of the Fund's Capacity Development Strategy Review—Concept Note.

[11] To increase effectiveness, planned changes to the data ROSC include shifting to a thematic approach (by sector) and focusing on operational data.

46.   Guided by a more candid assessment of data weaknesses hindering surveillance, capacity development will be better targeted through a more informed discussion with the authorities. Capacity development will be facilitated by two new multi-partner vehicle trust funds (i.e., the Financial Sector Stability Fund (FSSF) and the Data for Decisions Fund—D4D) established to help finance expanded technical assistance to support financial sector stability and address data weakness and gaps, especially in low-and middle-income countries, including fragile states.

47.   Capacity development will remain central to helping countries improve data quality. Statistical capacity development will continue to serve several objectives. At the most fundamental level, some countries require assistance in building statistical infrastructure to produce basic macroeconomic statistics, particularly in the aftermath of a conflict, an economic transition, or joining the Fund. The Fund will continue to lead the work on the G20 Data Gaps Initiative—jointly with the Financial Stability Board—with a focus on monitoring the implementation of Phase Two.

48.   Using AI/text mining, country teams will be able to use cross-country data embedded in TA reports to facilitate comparator studies involving regional and other groupings. Work is already underway to provide staff better access to data on Fund-wide capacity development (CD), including delivery by country, topic, and region. This information is already available to staff via the CD data portal and is communicated to the Executive Board and the public in the budget outturn paper and in the IMF Annual Report.[12] Data on results is collected in the Fund's Results-Based Management software, CD-PORT, which has been used for all CD activities since May 2017, although data will only be available over time as CD yields the intended changes on the ground.

GOVERNANCE, IMPLEMENTATION, AND RESOURCE IMPLICATIONS

49.   The governance framework approved by Management in January 2018 will ensure the effective implementation of the strategy over a five-year horizon. A new Standing Committee on Data and Statistics—comprising department heads and Management—will succeed the EDSC with a broader purview, and a mandate to oversee implementation through 2023. Starting work in FY2019, the Standing Committee will decide on sequencing and implementation pace, seeking to ensure integration of what has been a fragmented universe of data and statistics. It will promote the necessary consultation and coordination of all aspects of the data and statistics work streams, and promote an integrated approach to the anchoring frameworks for bilateral and multilateral surveillance. As part of the new governance framework, EDT will move to the Statistics Department.

50.   The strategy aims to be broadly cost-neutral in the medium term, with implementation involving reallocation of resources and exploitation of efficiencies. As the

---

[12] Reviewing and augmenting internal access to TA reports—an important Fund knowledge asset—is fully aligned with the priorities of the Knowledge Management work stream. However, some pilot work is needed to understand better how to best surface and share the underlying economic data.

strategy fundamentally calls for integrating and re-orienting long-standing frameworks and initiatives, the costs imputable to its vision are envisaged to be modest. Costs will come from the planned stepping up of support for member countries' efforts to achieve cloud-based data dissemination and support for the use of Big Data, including through a platform for running Big Data experiments and storing and disseminating results. Demand for new skills is expected to be met through re-training of current staff and recruitment of some new staff with specialized skills (e.g., data scientists).

51.   The elements of the strategy reflecting the new vision can be financed through reallocation of structural resources and some transitional resources. The indicative estimates of the costs expected for the more significant elements of the strategy for the FY2019-21 medium-term budget are shown in Table 1. Pending resolution of these uncertainties, the elements of the strategy reflecting the new vision are likely to require transitional resources of about US$3.9-6.5 million over the next three years, while structural needs would amount to US$2.6-4.3 million. Many of the envisaged activities can likely be accommodated through a reallocation of resources. Over the medium term, the decline in time spent on data collection is expected to generate savings to support the expansion of the global data commons. Eventual actions to address the broader challenges of the digital world, beyond the scope of this paper, will emerge as thinking in this area is further developed.

ISSUES FOR DISCUSSION

- Do Directors agree with the six strategic priorities?

- Do Directors have specific views on the relative importance of the elements of the overarching strategy?

- Do Directors agree that the changing data and technology landscape affords an opportunity to expand support to the membership in publishing data essential for monitoring economic conditions and policies?

- Do Directors agree that a more candid and transparent assessment of data adequacy for surveillance is necessary?

- Do Directors think that the overarching strategy will help mitigate the risk posed by data weaknesses and associated projections?

Table 1. Data and Statistics Strategy:
Indicative Gross Resource Needs and Savings for FY2019–21[1]
(In millions of U.S. dollars, unless otherwise noted)

	Transitional	Structural	Implementation Timeline
INDICATIVE GROSS ADMINISTRATIVE NEEDS	**3.9-6.5**	**2.6-4.3**	
Strategic Priority 2: Global data commons			
Support members establish cloud-based open data for publication (SDMX)	1.1-1.8		FY19-FY21
Strategic Priority 3: Support the use of Big Data			
Big Data Platform License	0.6-0.9		FY19-ongoing
Knowledge Exchange for best practices	0.2-0.4	0.5-0.8	FY19-ongoing
Sponsor peer-to-peer learning for sharing of best practices on Big Data	0.5-0.8		FY19-ongoing
Training courses on tools and application of Big Data for Fund surveillance	0.5-0.8		FY19
Strategic Priority 4: Ensure secure and seamless sharing of data			
Roll out the Economic Data Registry	TBD once development is completed (April 2018)		
Enhance IT infrastructure to allow secure data sharing	0.5-0.8		FY19
Strategic Priority 5: Increase cross-country data comparability			
Create a unit or group in STA to focus on fostering cross-country comparability of key data		0.7-1.1	FY19-ongoing
Strategic Priority 6: Address data weaknesses by integrating surveillance priorities with capacity development			
Design and roll out of data adequacy tool	0.7-1.1		FY19-ongoing
Redesign and roll out the data ROSC	0.1	0.5-0.8	FY19-ongoing
Cross-cutting needs			
Retooling of staff, ongoing training		0.1-0.1	FY19-FY21
Recruitment of staff with specialist skills (data scientists)		0.9-1.6	FY19-FY21
INDICATIVE GROSS ADMINISTRATIVE SAVINGS			
Reallocation of resources and efficiency savings	TBD; the aim is to achieve savings to meet structural costs		FY19-FY21

Source: IMF staff estimates.

1/ Excludes capital expenses related to economic data management projects already in train, including the Big Data Platform.

Box 1. A Proposed Tool for Assessing Data Adequacy for Surveillance

The IEO found an upward bias in the assessment of data adequacy for surveillance. Staff has designed a tool that ensures a structured and transparent approach to determining fitness for purpose for key variables in each of the macroeconomic datasets essential to the work of teams. The tool would help promote a better-informed and more targeted dialogue with the authorities on data issues and guide capacity development priorities. The assessment will be firmly based on eliciting the implications of data weaknesses on the robustness and credibility of analyses and policy advice.
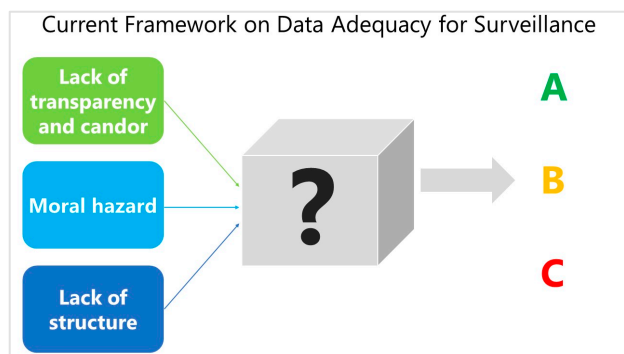
The new tool includes a rating system based on a questionnaire. The questionnaire would be appended to the Statistical Information Appendix of staff reports. It would cover five sectors: 1) national accounts, 2) price indices, 3) government finance statistics, 4) external sector statistics, and 5) monetary and financial statistics. For each sector, there will be five-six questions eliciting the impact of data weaknesses on the robustness of analyses and the soundness of policy advice. The proposed assessment tool offers several desirable features:

Surveillance-driven. The assessment of data adequacy for surveillance will continue to led by country teams, with input from the authorities. Given the evolving nature of economic conditions and Fund analyses, country teams would have the option to adjust the basic questionnaire to fit their circumstances (with a short explanation included in the Statistical Appendix). Moreover, assigning a rating will continue to be the country team's responsibility—taking into account Fund policies and statistical knowledge. The rating would reflect the degree to which data weaknesses limit mission staff's ability to conduct analyses and quantify the effects of policy options, thereby compromising the validity of the conclusions.
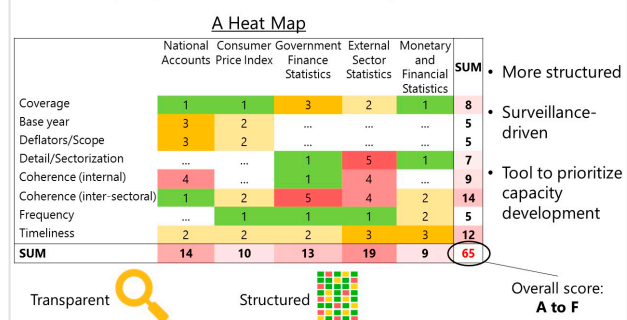
Heatmap presentation. The questionnaire responses could be presented in a heatmap that would highlight the main areas of data weaknesses hampering surveillance.

Enhancing policy dialogue around data and traction. The proposed tool should help elevate data issues to the policy level, as teams will need to explain to the authorities the weaknesses underpinning specific ratings. The team will thus need to engage the authorities in a more structured exchange about recent statistical developments, the scope for necessary data improvement, and, when warranted, targeting the increase in resources to address data weaknesses and upgrade statistical systems. As warranted, the staff report will need to provide an account of the discussion with the authorities.

Current Framework on Data Adequacy for Surveillance

Data Adequacy for Surveillance – Proposed Assessment Tool

A Heat Map

	National Accounts	Consumer Price Index	Government Finance Statistics	External Sector Statistics	Monetary and Financial Statistics	SUM
Coverage	1	1	3	2	1	8
Base year	3	2	...	...	...	5
Deflators/Scope	3	2	...	...	...	5
Detail/Sectorization	...	...	1	5	1	7
Coherence (internal)	4	...	1	4	...	9
Coherence (inter-sectoral)	1	2	5	4	2	14
Frequency	...	1	1	1	2	5
Timeliness	2	2	2	3	3	12
SUM	14	10	13	19	9	65

- More structured
- Surveillance-driven
- Tool to prioritize capacity development

Transparent   Structured   Overall score: A to F

Tool to prioritize capacity development. The tool would help operationalize synergies between surveillance and capacity development. By highlighting areas where data weaknesses are policy-critical the tool will help country teams and the authorities better prioritize TA and training.

Box 2. Data Dissemination Initiatives: A Public Good Supporting Surveillance

In the mid-1990s, after the Mexican financial crisis, the Fund introduced the Data Standards Initiatives with voluntary participation by members. The initiatives seek to promote transparency by encouraging publication of economic and financial statistics and improvement in countries' statistical systems. The initiatives, periodically evaluated by the Executive Board to adjust to the evolving needs of the membership, comprise the SDDS, the SDDS Plus, and the e-GDDS.

The SDDS, established in 1996, was to guide countries with access to international capital markets (mainly emerging markets and advanced economies) in publishing economic and financial data with prescribed coverage, periodicity, and timeliness—on a national data summary page and according to an advance release calendar. Flexibility options under the SDDS were introduced to allow for the dissemination of up to two data categories with periodicity and/or timeliness less than prescribed to accommodate country-specific circumstances. The SDDS Plus was established in 2012, building on the SDDS by focusing on members with systemically important financial sectors. The SDDS Plus builds on the progress achieved under the SDDS by introducing more rigorous requirements for disseminating a broader set of data (nine additional datasets) with no flexibility options. The e-GDDS is meant for countries with lower capacity, especially low-income countries, and seeks to help them improve data transparency and governance by publishing the key macroeconomic and financial data used for surveillance.

Data dissemination brings a broad range of benefits for the country and the global community. Timely economic and financial data is a global public good, critical for the proper functioning of financial markets and global financial stability.

Transparency promotes interactions with the outside world, as noted by Fischer (2002). Transparency enables a better-informed debate by governments and various stakeholders in-country and abroad about the formulation, analysis, and results of policy-making. It highlights the basis, trade-offs, and risks that feed into decisions, thereby instilling credibility to the authorities' policies. In addition to fostering accountability with domestic stakeholders—mainly parliament, civil society and the public at large—making more and better data available has also helped mitigate perceived risks by domestic and international investors. In disseminating information, providers must interact by listening to and absorbing the reactions of outsiders. In this way, presenting such information also enhances the effectiveness of Fund surveillance. The data dissemination standards can also help countries meet other needs, such as reducing data reporting burdens, lowering borrowing costs,[1] and helping identify critical data gaps in the context of Article IV consultations.

Public access to data can promote data quality by enabling public scrutiny of accuracy and reliability, and the soundness of methodologies—all of which support policy analysis and decision making. Benefits also accrue from international comparisons and peer review of data.

[1]/A recent IMF staff Working Paper (WP/17/74) suggests that greater data transparency—promoted through the IMF data standard initiatives—may result in a 15 percent reduction in the spreads on emerging market sovereign bonds three months after the improvements are made, with this effect tending to increase over time.

Box 3. Use of Big Data in Member Countries

Big Data is already used by statistical authorities for real-time/high-frequency input for policymaking. *Statistics Netherlands* and *Eurostat* are among the statistical agencies most advanced in using Big Data, with the former launching the *Center for Big Data Statistics* and developing "beta products" using Big Data. One example is using the intensity of traffic flow on national roads to track GDP growth. The *Netherlands* also uses Facebook and Twitter data to estimate consumer sentiment. Along with Google and Data provider, it published a Big Data study to measure the size of the Dutch digital economy. *Eurostat* has partnered with European statistical offices to pilot Big Data projects and uses Big Data for macroeconomic nowcasting. *Eurostat*, in collaboration with the University of Bergamo, used the number of Wikipedia page views to predict tourism flows to three European cities.

Central banks and other agencies also use Big Data for compilation and research. The *European Central Bank* receives data from Google searches to follow weekly volume changes in different categories of goods by geographical area. The *Deutsche Bundesbank* has used Google searches as a proxy for investors' worry to signal the risk of a run on bank deposits. The *U.S. Federal Reserve* uses Big Data for nowcasting GDP with its GDPNow Index. The *Central Bank of Estonia* uses mobile positioning to collect data on travel services for balance of payments statistics. The *Bank of England* used transaction-level data on over-the-counter derivatives markets in January 2015 to investigate the impact of the de-pegging in the euro-Swiss franc market by the *Swiss National Bank*. The finance ministry in *Colombia* uses Google trends to monitor economic activity. *China and Italy* approximate job vacancy rates through web-scraping. *Canada* is considering the use of scanner data collected for the consumer price index to publish weekly indicators of retail trade activity. *Bank Indonesia* produces a job vacancy indicator based on advertisements posted in on-line job market portals, and a property price index in the secondary market using sales information. Other examples (reported in the SDN/17/06): include:

- mobile phone data for tourism, transportation, and urban statistics (e.g., *Belgium, Brazil, Eurostat, Indonesia, Israel, Italy, Poland*);
- web-scraping for price indices, labor market indicators, and enterprise profiling (*China, Ecuador, Eurostat, Finland, Germany, Hungary, Japan*);
- smart meters for energy and environmental statistics (*Belgium, Canada, Eurostat*); and
- credit card, cash register, and scanner data for price and other economic statistics (*Denmark*).

IFIs are collaborating with academia, the private sector, and others on Big Data projects. The UN is cooperating with country authorities, *Eurostat*, the *OECD*, the *World Bank Group*, *NASA*, *Google*, *IBM*, Harvard University, and others to develop Big Data ideas. And the *UN* and some countries are working with telecom companies to use mobile data to monitor progress on the Sustainable Development Goals, including through geospatial data. The *World Bank Group*, in collaboration with Deloitte, undertook some pilot studies to help formulate trade and competitiveness policy. One project used regulatory data to classify and assess the impact of non-tariff measures on economies and competitiveness. To overcome the challenges facing private-sector organizations wanting to use their Big Data for social good, the Data-Pop Alliance, MIT, Orange Group, the World Economic Forum, and Telefónica have come together in the Open Algorithms Project to extract key development indicators.

Annex I. Taskforce and Other Contributing Staff

Members of the Interdepartmental Taskforce on Data and Statistics Strategy

Louis Marc Ducharme (Chair)—STA	
Patrizia Tumbarello—STA	Hua Zhang—ITD
Michael Atingi-Ego (Co-Chair EDSC)—AFR	Ivanna Hollar Vladkova—KMU
Dominique Desruelle—KMAC and AFR	Katharine Christopherson—LEG
Elena Loukoianova—APD	Karen Ongley—MCD
Ricardo Davico—APD	Jonathan Manning—MCD
George Tsibouris—EDT	Miguel Savastano—MCM
Ben Kelmanson—EUR	Karl Driessen—MCM
Beata Kudela—EUR	Brian Christensen—OBP
Manal Fouad—FAD	Andrea Richter-Hume—OBP
Thomas Krueger—FIN	Gisela Ulmschneider—OBP
Heikki Hatanpaa—FIN	Oya Celasun—RES
Robert York—HRD	Toh Kuan—RES
Charleen Gust—ICD	Johan Mathisen—SPR
Charlie Kramer—ICD and iLab	Roberto Cardarelli—WHD
El Bachir Boukherouaa—ITD	Bennett Sutton—WHD
Margherita Bruni—ITD	

Working Groups (Taskforce members and other contributing staff)

Staff Incentives for Data Management: Karen Ongley (MCD—Lead); Robert York (HRD); Margherita Bruni (ITD); Gisela Ulmschneider (OBP); Ezgi Ozturk (FIN); Serkan Arslanalp (STA); and volunteers: Tom Dowling (WHD; Juan Carlos Flores (AFR); Leni Hunter (APD); Jehann Jack (AFR); Etibar Jafarov (WHD); Ahn Van Le (APD); Esther Perez Ruiz (WHD); Amir Sadeghi (MCD); Dustin Smith (EUR); and Farid Talishli (MCD).

Data Management: George Tsibouris (EDT—Lead); Margherita Bruni (ITD); Oya Celasun (RES); Toh Kuan (RES); Ezgi Ozturk (FIN); Hua Zhang (ITD); Alexandre Chailloux (STA); Jonathan Manning (MCD); Ben Kelmanson (EUR); Beata Kudela (EUR); Gustavo Ramirez (AFR); Edgardo Ruggiero (STA); Laurent Meister (STA).

Data Needs for the Fund and the Membership: Andrea Richter-Hume (OBP—co-Lead); Brian Christensen (OBP); Miguel Savastano (MCM); Katharine Christopherson (LEG); Jemma Dridi (AFR); Subir Lall (MCD); Lorraine Ocampos (MCD); Johan Mathisen (SPR); Serkan Arslanalp (STA); Patrizia Tumbarello (STA–co-Lead).

Statistical Innovation and Big Data: Charlie Kramer (ICD/iLab—co-Lead); El Craig Beaumont (EUR); Bachir Boukherouaa (ITD); Karl Driessen (MCM); Stephanie Medina-Cas (STA); Elad Meshulam (CSF/Creation Lab); Papa N'Diaye (SPR); Mwanza Nkusu (FIN); Patrizia Tumbarello (STA—co-Lead).

Cross-Comparable Data: Miguel Savastano (MCM—co-Lead); Roberto Cardarelli (WHD—co-Lead); Serkan Arslanalp (STA); Elena Loukoianova (APD); Ricardo Davico (APD); Margherita Bruni (ITD); Ben Sutton (WHD).

Capacity Development and Data: Charleen Gust (ICD—Lead); Ivanna Vladkova Hollar (KMU); Manal Fouad (FAD); Dominique Guillaume (FAD); Carlos Sanchez-Munoz (STA).

Data Security: Mohammad Alhammouri (ITD).