STATISTICS

The IMF's Experience with the GSBPM

APRIL 28, 2023

ModernStats Community of Practice

Marco Marini

IMF Statistics Department

Outline

- Overview of implementation of GSPBM at the IMF Statistics Department

- Current uses of GSBPM to strengthen data governance

- GSBPM extension: Big Data vs. Traditional Data

- Conclusions

GSBPM Adopted as a Framework to Modernize the Data Functions of the IMF Statistics Department (STA)

- In 2018, IMF's [Overarching Strategy on Data and Statistics](#) called for a modernization of its data management functions to increase efficiency of work processes and improve data quality.

- Since then, the IMF's Statistics Department (STA) has made significant progress in modernizing its operations to collect, process, and disseminate official statistics of its member countries – a global public service for the statistical community.

- This presentation highlights how STA implemented and used the **Generic Statistical Business Process Model (GSBPM)** to improve quality and efficiency of the set of business processes and sub-processes managed by the department.

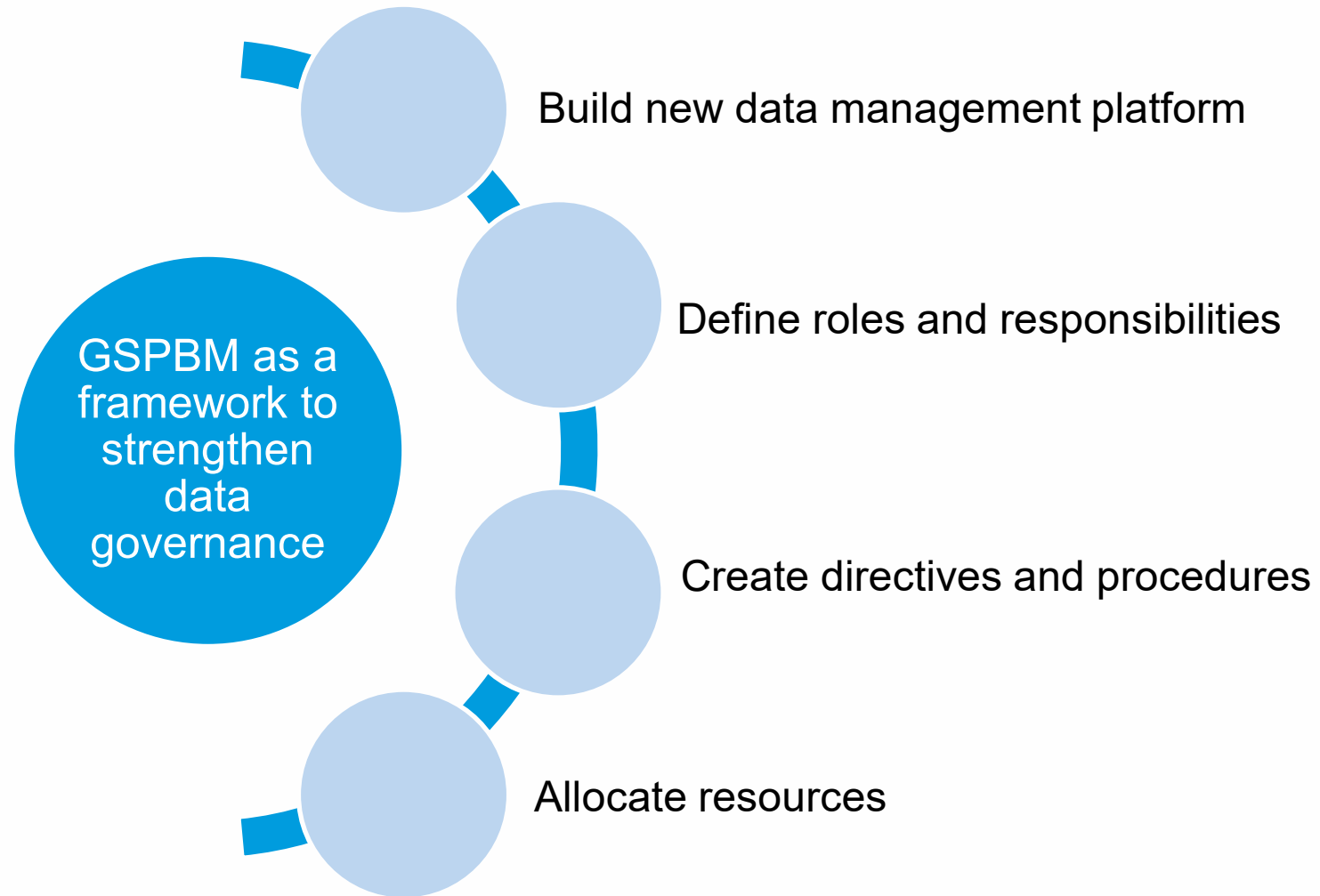# Looking Backward: Key Milestones of STA's Implementation of the GSPBM

❑ Map 'as-is' STA data production processes adopting **Business Process Methodology Notation (BPMN)**, including delineation of roles and responsibilities using the **RACI** approach.

❑ Creation of an **internal data governance body** to oversee the development, implementation and maintenance of departmental statistical business processes model.

❑ Development of **metadata standards** and use of a **harmonized information model** based on **enterprise-wide vocabulary**.

❑ Design a **'to-be' business process** using consistent language and harmonized practices across different data products, aligned with international standards and best practices of statistical production (**GSBPM**).

STA's Implementation of the GSBPM

1 Specify Needs	2 Design	3 Build	4 Collect	5 Process and Compile	6 Analyze	7 Disseminate	8 Evaluate
1.1 Identify needs	2.1 Design product	3.1 Build /enhance collection system	4.1 Maintain registries and calendars	5.1 Classify and code	6.1 Interpret and explain statistics	7.1 Disseminate statistical products	8.1 Gather evaluation inputs
1.2 Consult and confirm needs with internal and external users	2.2 Design collection	3.2 Build / enhance processing system	4.2 Set up collection	5.2 Edit and impute		7.2 Manage release of dissemination products	8.2 Conduct evaluations
1.3 Identify concepts and conceptual frameworks	2.3 Design processing	3.3 Build / enhance dissemination system	4.3 Run collection	5.3 Transform		7.3 Promote dissemination products	8.3 Agree on action plan
1.4 Establish output objectives	2.4 Design dissemination	3.4 Conduct end to end testing	4.4 Close collection	5.4 Derive aggregations and new variables		7.4 Manage users	
1.5 Check data availability	2.5 Design production system and workflow			5.5 Apply disclosure control			
1.6 Prepare and submit business case				5.6 Close Processing			

✓ 8 phases
✓ 34 sub-processes
✓ 137 activities + RACI

How the GSBPM is being used?

GSPBM as a framework to strengthen data governance

- Build new data management platform
- Define roles and responsibilities
- Create directives and procedures
- Allocate resources

Build new data management platform – iData Project

- The IMF is in the process of replacing its legacy data management and dissemination platforms. The new solution—**iData**—is a modern, cloud-based platform built on the Azure stack covering the entire data lifecycle from ingestion to dissemination.

- The objective of the iData project to deliver **modern data management tools and enable seamless access to the IMF's economic and financial data** for anyone anywhere on any device.

- The Statistics Department will leverage the opportunity provided by the new platform to **operationalize the GSPBM**. In particular, it will accelerate the implementation of standardized business processes across different data products and of a common SDMX-based information model using enterprise-wide vocabulary.

Define Roles and Responsibilities (RACI) "Design Collection" Example

2.2 DESIGN COLLECTION

List of Activities / RACI Matrix

Activity	Senior Mgmt.	Business	Design	Build	Strategy	Other IMF
1) Design collection instruments (questionnaires, reporting templates, API access, Data sharing agreements).		A	R	I		
2) Identify and document reporting entities, data providers and the collection calendar.		A/R	I	C		
3) Identify and document the different modes (e.g., IMF collection platform, API, web-scraping) of collection.		A/R	C	C		
4) Identify and document "hard check" collection validations.		A/R	C	I/R		
5) Design collection validation reports		A/R	C	I		
6) Design collection management reports.		A/R	C	I		
7) Design contact management, follow-up and non-response escalation procedures and tools.		A/R	C	I		
8) Develop collection training material for collection staff.		A/R	C	I		
9) Design and document collection sign-off procedures.		A/R	C	I		

Create Directives and Procedures

Potential Directives, Centrally Governed Standards and Methods, and Tools to be established.		
Directives	**Centrally Governed Standards and Methods**	**Tools**
Directive on Disseminating Data	Classification and Reference Lists	Statistical Product Design and Documentation Templates
Directive on Processing Data	Statistical Methods (seasonal adjustment, benchmarking, currency conversion, indexing, weighting)	Business Case Templates
Directive on Collecting Data	Data exchange mechanisms (SDMX)	Uses Feedback Reports
		Data Usage Reports
		Project Planning Template

Example: Directive on Disseminating Statistics

Objective

The purpose of this directive is to ensure that the statistics products disseminated by the IMF Statistics Department are (1) accessible, (2) interpretable, (3) relevant and aligned with the mandate of the IMF, (4) accurately presented, (5) consistent with Fund-wide standards (...)

Directive Statement

1. The Statistics Department will disseminate economic, financial, social, and environmental statistics and their associated metadata, aligned with the relevant statistical standards consistent with the mandate of the IMF (...)

Scope

This directive applies to all statistical products and related content disseminated by STA (...)

Roles and Responsibilities

The Statistics Product Owner is responsible for establishing the release calendar (GSBPM 4.1) and approving the release of the statistics product (GSBPM 7.1) (...)

- Objective

- Directive Statement

- Scope

- Roles and Responsibilities (linked to GSBPM)

Example: Design Templates

Product Design		
Author:	Real Sector Division	
Last Update Date:	6/18/2020	
Dimension / Attribute Description	**SDMX Code**	**Selected Classes / Dimensions**
Frequency	FREQ	A - Annual, Q - Quarterly, M - Monthly
Adjustment	ADJUSTMENT	Y - Calendar and seasonally adjusted data, N - Neither seasonally adjusted nor calendar adjusted data
Reference Area	REF_AREA	
Counter Party Reference Area	COUNTERPART_AREA	
Sector	REF_SECTOR	S1 - Total economy
Counterparty Sector	COUNTERPART_SECTOR	S1 - Total economy
Consolidation	CONSOLIDATION	_Z - Not applicable
Accounting Entry	ACCOUNTING_ENTRY	_Z - Not applicable
Stocks, Transactions and Other Flows	STO	XGCB - Export of Goods - Customs Basis, MGCB - Import of Goods - Customs Basis
Instrument	INSTR_ASSET	_Z - Not applicable

Proposed Suite of Design Templates
- ✓ Product Design
- ✓ Collection Operations
- ✓ Collection Report
- ✓ Collection Validations
- ✓ Processing
- ✓ Processing Validations
- ✓ Processing Report
- ✓ Dissemination Validations

✓ Governed by Data Governance and Services Division.

✓ Embed departmental standards, directives and guidelines.

✓ Serve multiple purposes including:

    ✓ the consistent departmental documentation of statistical programs;

    ✓ the consistent design of statistical programs; and

    ✓ input into the build process.

Allocate Resources for Data Projects
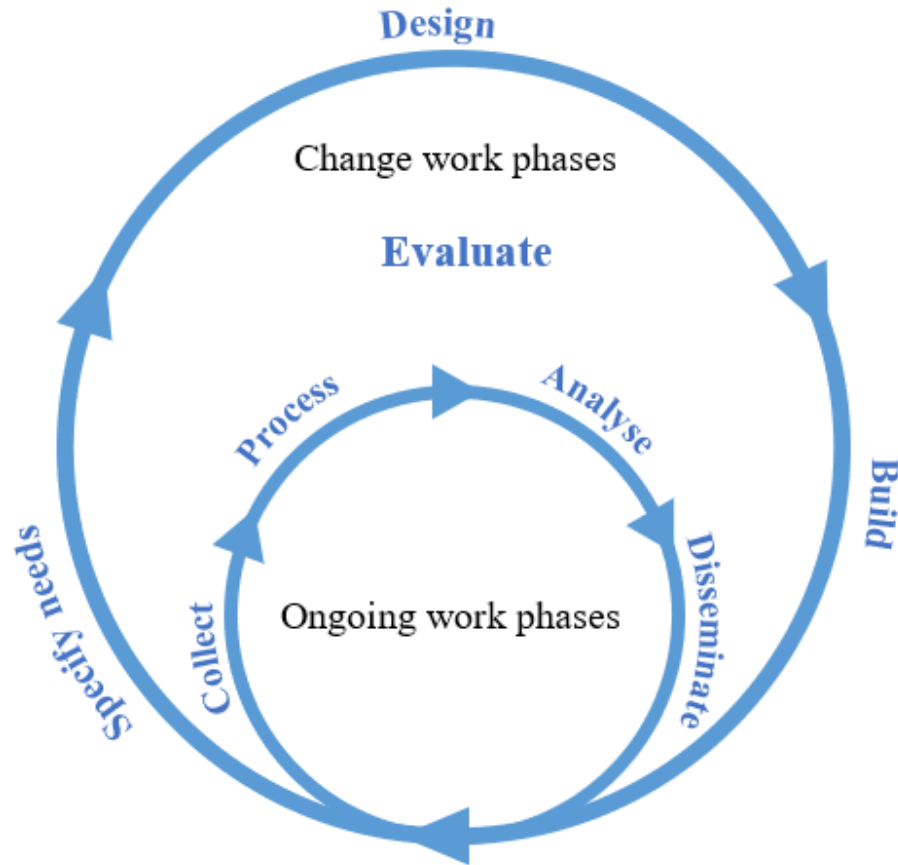
- Internal governance body to conduct **annual assessment of resources needed** for (i) ongoing operations, (ii) enhancements to existing products, and (iii) new initiatives.

- GSBPM is helpful to identify impact of these projects by phase/division, and take **decision on trade offs, reprioritizations, and additional resources needed**.

- **Evaluate phase** needs to be further developed in this process: it is key to consult periodically with internal and external users to ensure that departmental outputs remain fit for purpose.

GSPBM Extension: How "Big Data" is changing the GSBPM

SURVEY — Less timely/Burdensome   **BIG DATA** — More timely/Less Burdensome

Official Statistics

Question → Design → Collect → Process → Analyze

VS

Question → Collect → Design → Process → Analyze

FEATURES OF SURVEYS

There is a need to recognize a fundamental change in the way we produce statistics:

- The statistical process of the 1980s, 1990s and 2000s started with the design.
- Statisticians would meticulously design the process (e.g. determine sample sizes, develop and test questionnaires, set up robust collection systems).
- Data would be collected (often supported by legal frameworks that compelled response) and analyzed. Conclusions would be drawn.
- Data quality indicators could be readily constructed (co-efficients of variation, response rates, confidence intervals)

FEATURES OF BIG DATA

Need to learn how to work in a world that is fundamentally different from it was just 20 years ago...

- The statistical process often starts with the data.
- The statistician needs to design around the data, the data are no longer a result of the design.
- This introduces a number of new constraints into the work of a statistician.
- Techniques such as record linkage, automatic coding take on an increased importance.
- Quality indicators and quality assessment is more difficult.

Conclusions

- Experience of IMF's Statistics Department is that the **GSBPM adapts well to the statistical operations of an international organization.**

  - Excellent framework to map out specific activities that are differentiated from the functions of a national statistical office.

- GSPBM can be a useful reference framework to strengthen data governance in a data-producing organization. Some examples include:

  - build a new **data management platform** to support a process-based data pipeline

  - delineate **accountability/responsibility** across stakeholders

  - establish **directives and procedures** to promote best data management practices

  - **prioritize and plan** resources for data projects.