

10TH EXPERT GROUP MEETING ON

# Statistical Data and Metadata eXchange

JANUARY 25-28, 2021

## **SDMX-based Automatic Data Collection Platform**

January 27, 2021

Khaled AlAjmi

Section Chief – IMF

The views expressed in this presentation are those of the authors

# Outline

- IMF Data Collection Platform
- ADCP Architecture
- Performance improvements
- New user Interface
- Next steps

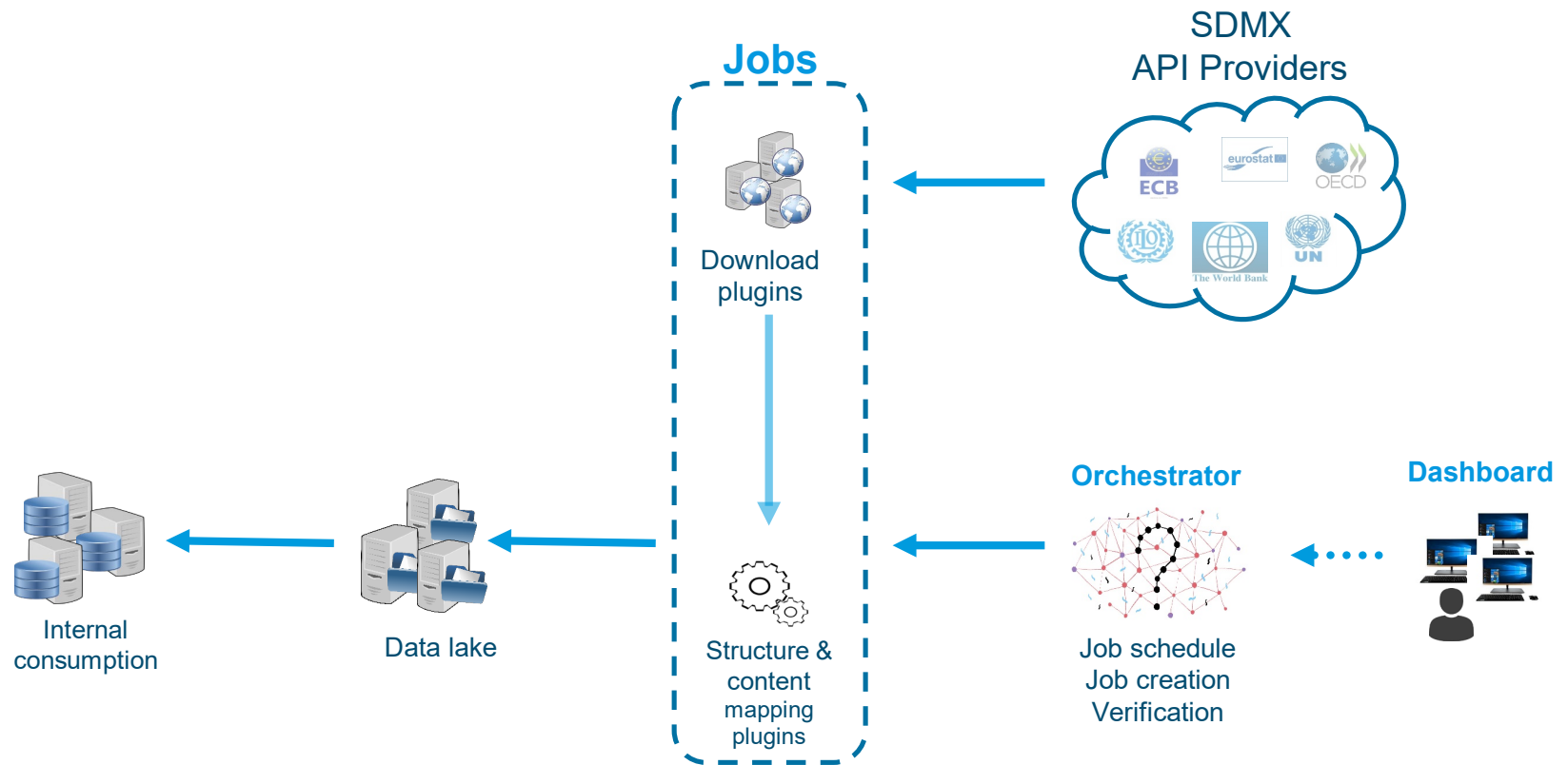
# IMF Data Collection

- Over 70 external datasets
- From 9 data providers  
EuroStat, ECB, OECD, UN, ILO, WB, BIS, US Census Bureau, and Bloomberg
- NSDP underlying data links – 689 URL instances;

# Challenges

- Diverging formats  
SDMX 2.0 or 2.1 standards, GESMES, JSON and CSV format
- Heterogeneous protocols  
RESTful, RESTful secure, SOAP, SOAP Secure web service
- Complex codebase
- Unsatisfactory performance
- Business overreliance on development team

# Architecture



# Improvement

- Better resource utilization
- Improved performance
- Empowerment of end user

# Performance

		Pre	Post
<b>Eurostat</b> DS-645593	Download Processing	3 h 30 minutes 1 h 40 minutes	6 minutes 4 minutes
<b>Eurostat</b> DS-575274	Download Processing	4 h 26 minutes 5 h 50 minutes	19 minutes 14 minutes
<b>Eurostat</b> PRC_HICP_MIDX	Download Processing	17 minutes 2 h 24 minutes	12 minutes 12 minutes
<b>World Bank</b> Quarterly Public Sector Debt (PSD), Quarterly External Debt Statistics SDDS (QDS), Quarterly External Debt Statistics GDDS (QDG)	Download Processing	1 h 45 minutes 3 h 25 minutes	15 minutes 9 minutes

# User Interface - Mapping

### Incoming Structure Components

Include additional field if any

▼ Dimension

COLLECTION	+
REF_AREA	+
FREQ	+
SURVEY	+
REPRESENTED_VARIABLE	+
SEX	+
AGE	+

▼ TimeDimension

TIME_PERIOD	+
-------------	---

▼ Attribute

OBS_STATUS	+
UNIT_MEASURE_TYPE	+
UNIT_MEASURE	+
UNIT_MULT	+
FREE_TEXT_NOTE	+
DECIMALS	+
SOURCE	+
INDICATOR	+

▼ PrimaryMeasure

Apply Scale

OBS_VALUE	+
OBS_VALUE(UNIT_MULT)	+

### Output Columns

Choose a delimiter if an output column contains multiple incoming components: |

COLLECTION	(column name)
COLLECTION	- +
REF_AREA	(column name)
REF_AREA	- +
FREQ	(column name)
FREQ	- +
SURVEY	(column name)
SURVEY	- +
REPRESENTED_VARIABLE	(column name)
REPRESENTED_VARIABLE	- +
SEX	(column name)
SEX	- +
AGE	(column name)
AGE	- +
TIME_PERIOD	(column name)
TIME_PERIOD	- +
OBS_STATUS	(column name)
OBS_STATUS	- +
UNIT_MEASURE_TYPE	(column name)

### Output Columns Summary

(Drag and drop to re-order the column)

COLLECTION
REF_AREA
FREQ
SURVEY
REPRESENTED_VARIABLE
SEX
AGE
TIME_PERIOD
OBS_STATUS
UNIT_MEASURE_TYPE
UNIT_MEASURE
UNIT_MULT
FREE_TEXT_NOTE
DECIMALS
SOURCE
INDICATOR
OBS_VALUE



# User Interface - Scheduling

Service Provider: OECD  
SDMX Dataflow ID: SNA\_TABLE1  
SDMX Dataflow Name: 1. Gross domestic product (GDP)  
 Is Active

Publication Schedule: [Clear](#)

Schedule	Quarter of Year	Month of Quarter	Week of Month	Day of Week	Time
<input checked="" type="radio"/> Yearly	<input type="text" value=""/>	<input type="text" value=""/>	<input type="text" value=""/>	<input type="text" value=""/>	12 : 0 AM
<input type="radio"/> Quarterly					
<input type="radio"/> Monthly					
<input type="radio"/> Weekly					
<input type="radio"/> Daily					

# User Interface - Monitoring

All monitoring Dataset(s): Show my dataset(s) only Reset filter Refresh

Completed In progress

Provider	Dataset Name	Dataset ID	Owner	Start Date	Finish Date	Execution Status	Content Status	Up-to-date Status	Time Download	Time Process or Transform	Action
ILO	[1.1.1] Working poverty rate (percentage of employed living below US\$1.90 PPP)	DF_SDG_ALL_SDG_0111_SEX_AGE_RT	Voonna, Srinivas Venkata	02-07-2019 15:59	02-07-2019 16:04	●	●	●	00:00:20	00:00:07	
ILO	[1.1.1] Working poverty rate (percentage of employed living below US\$1.90 PPP)	DF_SDG_ALL_SDG_0111_SEX_AGE_RT	Yuxin, Jiang	03-05-2019 00:03	03-05-2019 00:07	●	●	●	00:01:40	00:00:10	📅 🔔
ILO	[1.1.1] Working poverty rate (percentage of employed living below US\$1.90 PPP)	DF_SDG_ALL_SDG_0111_SEX_AGE_RT	Gu-Fang, Gracia	03-05-2019 11:03	03-05-2019 11:06	●	●	●	00:01:46	00:00:04	
ILO	[1.1.1] Working poverty rate (percentage of employed living below US\$1.90 PPP)	DF_SDG_ALL_SDG_0111_SEX_AGE_RT	Yuxin, Jiang	03-05-2019 00:04	03-05-2019 00:07	●	●	●	00:01:19	00:00:04	📅 🔔
ILO	[1.1.1] Working poverty rate (percentage of employed living below US\$1.90 PPP)	DF_SDG_ALL_SDG_0111_SEX_AGE_RT	Eang, Arunrith	03-01-2019 14:33	03-01-2019 14:37	●	●	●	00:00:25	00:00:06	
ILO	[1.3.1] Proportion of population covered by social protection floors/systems	DF_SDG_ALL_SDG_0131_SEX_SOC_RT	Voonna, Srinivas Venkata	02-07-2019 16:00	02-07-2019 16:04	●	●	●	00:02:00	00:00:02	
ILO	[1.3.1] Proportion of population covered by social protection floors/systems	DF_SDG_ALL_SDG_0131_SEX_SOC_RT	Yuxin, Jiang	03-05-2019 00:04	03-05-2019 00:05	●	●	●	00:01:39		👤 📅 🔔
ILO	[10.4.1] Labour income share as a percent of GDP	DF_SDG_ALL_SDG_1041_NOC_RT	Voonna, Srinivas Venkata	02-07-2019 16:25	02-07-2019 16:29	●	●	●	00:00:35	00:00:04	
ILO	[10.4.1] Labour income share as a percent of GDP	DF_SDG_ALL_SDG_1041_NOC_RT	Yuxin, Jiang	03-05-2019 00:04	03-05-2019 00:05	●	●	●	00:01:37		👤 📅 🔔
ILO	[5.5.2] Female share of employment in managerial positions	DF_SDG_ALL_SDG_0552_OCU_RT	Voonna, Srinivas Venkata	02-07-2019 16:25	02-07-2019 16:29	●	●	●	00:02:04	00:00:02	

1 2 3 4 5 10

Inactive by service provider

# Moving Forward

- Confidential data sources
- Cloud infrastructure
- Streams and unstructured data

# Thank you

Khaled AlAjmi

[KAIAjmi@imf.org](mailto:KAIAjmi@imf.org)