

8th International  
Conference on  
**BIG DATA**  
& Data Science for Official Statistics

**BILBAO 2024**

Informing Climate Change and  
Sustainable Development Policies  
with Integrated Data

**BILBAO. SPAIN** | **10-14 JUNE 2024** | **#UNBigData2024**

# IMF Big Data Center: Driving Innovation in Macroeconomic Statistics

Alessandra Sozzi, Data Scientist, IMF





**STATISTICS**

# **IMF Big Data Center: Driving Innovation in Macroeconomic Statistics**

**JUNE 11, 2024**

Alessandra Sozzi

Data Scientist – IMF Statistics Department

**IMF BIG DATA CENTER**

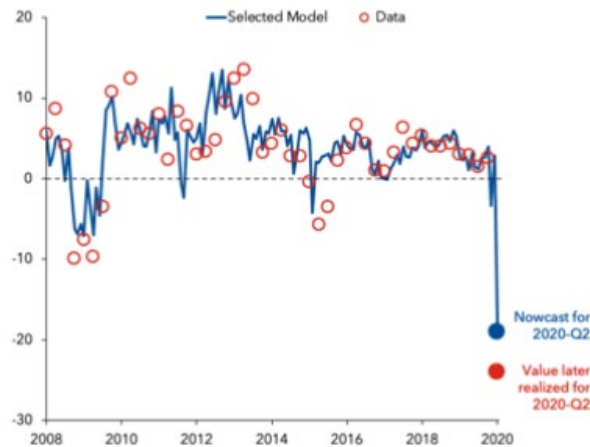
# **Vision for the IMF Big Data Center**

# Advantages of Big Data for macroeconomic and financial analysis

## Timeliness

### Real-time read

Nowcasting for Botswana's GDP helped inform policymaking by predicting trends months before official data were available. (annual change in quarterly GDP, percent)



Sources: Botswana authorities; Haver Analytics; and IMF staff calculations.



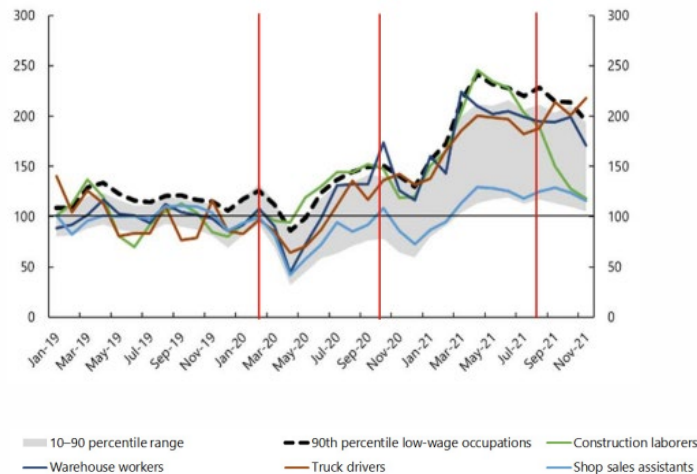
Source: [Choi and Iyer \(2022\)](#).

## Granularity of the analysis

Figure 1.1: Vacancies in Selected Low-Wage Occupations

(Average 2019 = 100)

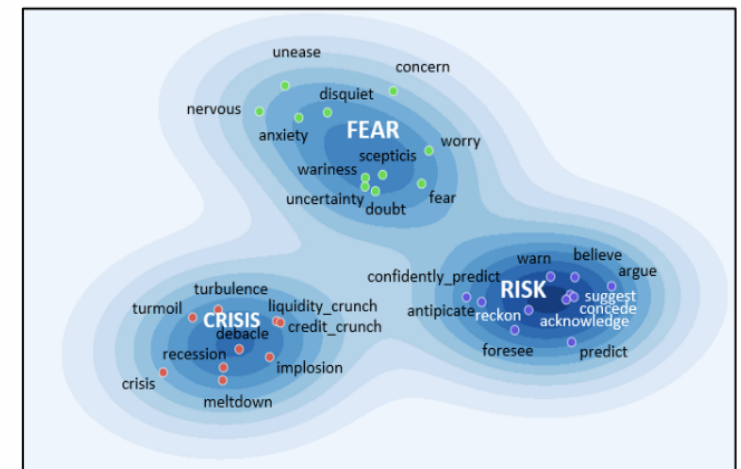
### United States



Source: [Duval and others \(2022\)](#).

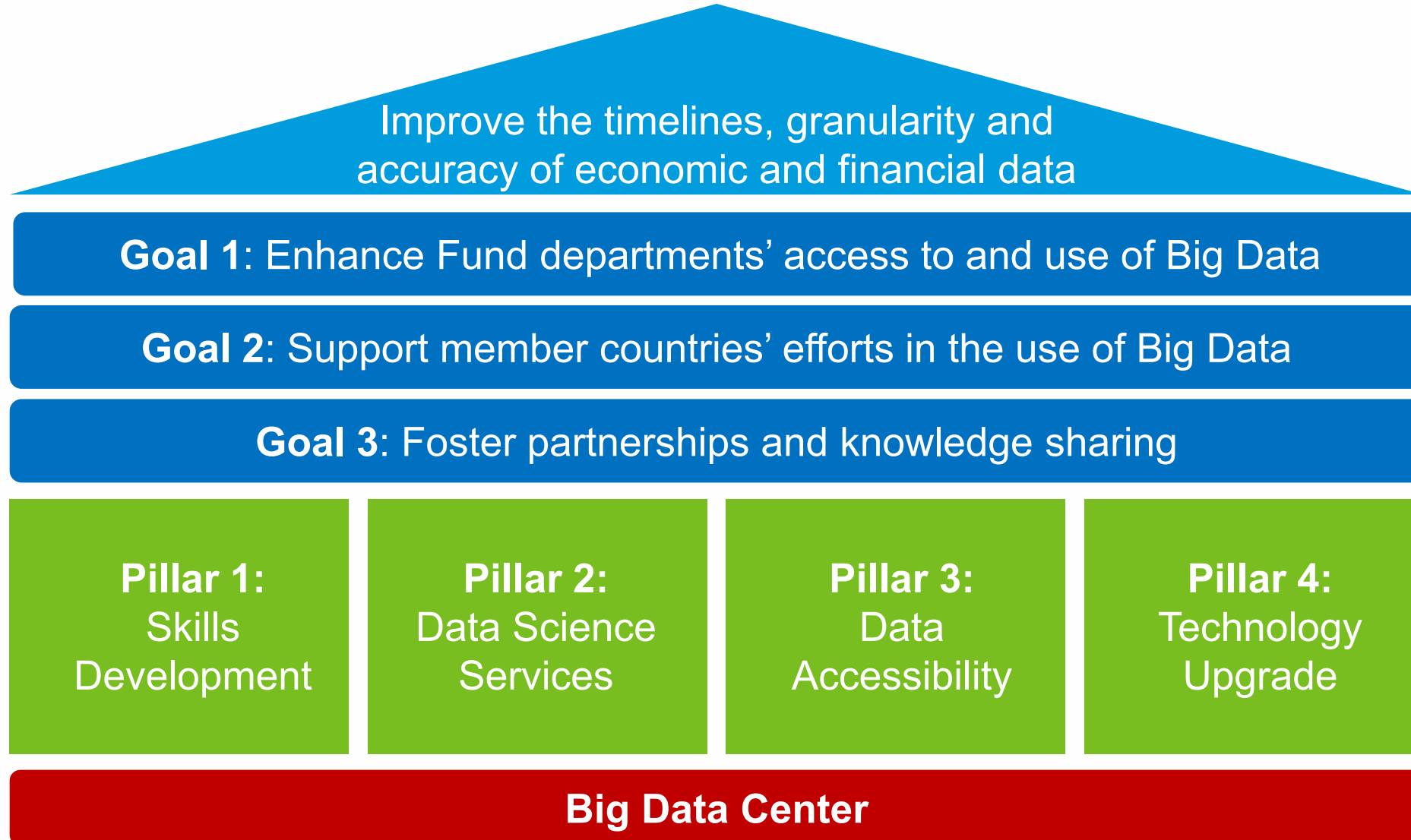
## Filling the areas where official statistics are lacking

Word Vector Representations, Semantic Clustering, and Sentiment



Source: [Huang and others \(2019\)](#).

# IMF Big Data Center: Pillars and Goals



# **What is the IMF Big Data Center doing in support of the IMF mandate?**

# GDP Nowcasting

On-going research addresses delays in GDP data dissemination or lack of capacity to publish quarterly GDP data, enriching policy advice.

- To **nowcast GDP**, alternative predictors complementary to traditional input such as PMI have been studied:
  - **Satellite images** such as nightlights
  - **Google search volumes**
  - **Google Places data**
- Explored with machine learning techniques
  - for an individual country
  - for a panel of countries

## Nighttime Light



(a) China (mainland) 1992

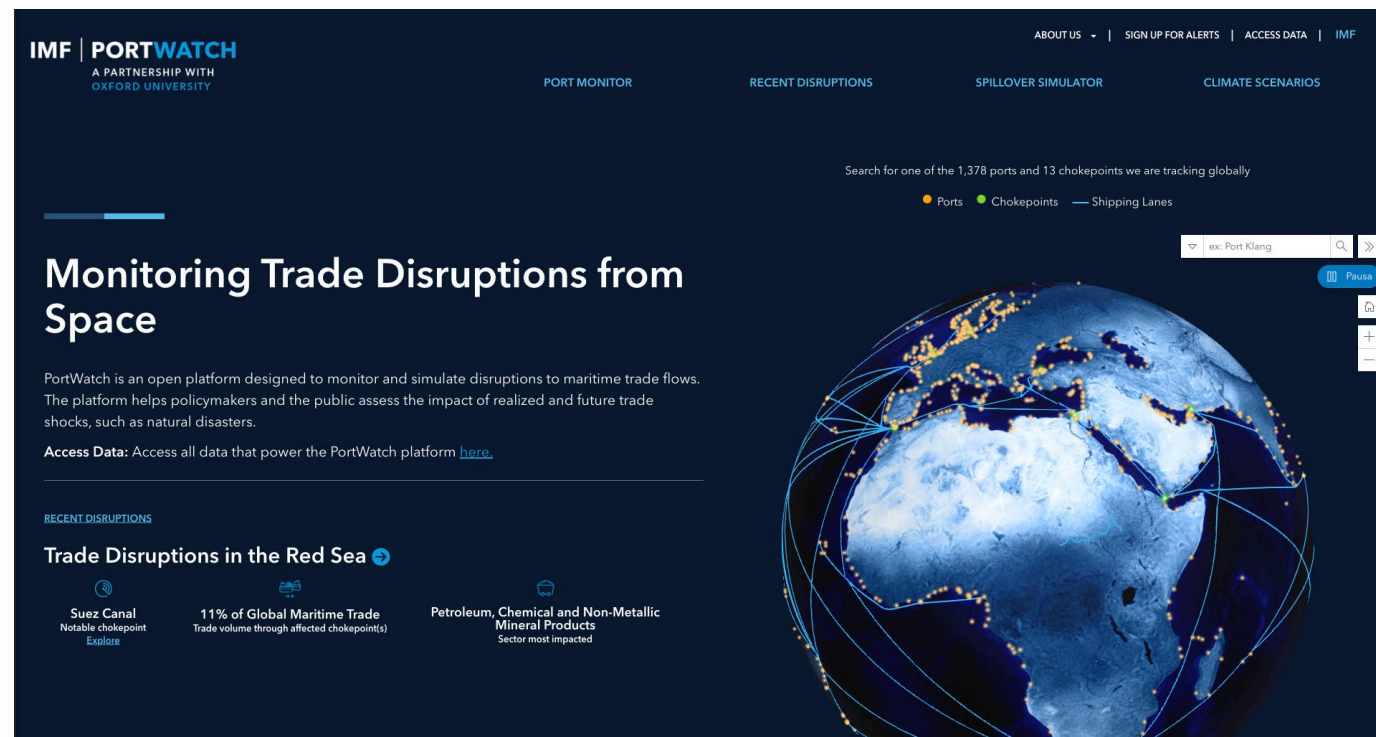


(b) China (mainland) 2013



# AIS Data and Trade Nowcasting

- **Automatic Identification System (AIS) data** (on vessel traffic) have been used in tracking international trade in real time.
- The AIS data are also used in climate-related projects such as [whale conservation](#).
- [PortWatch](#) releases AIS-derived **real-time trade estimates** to monitor disruptions from shocks, assessing their regional and global spillovers through port-to-port linkages.





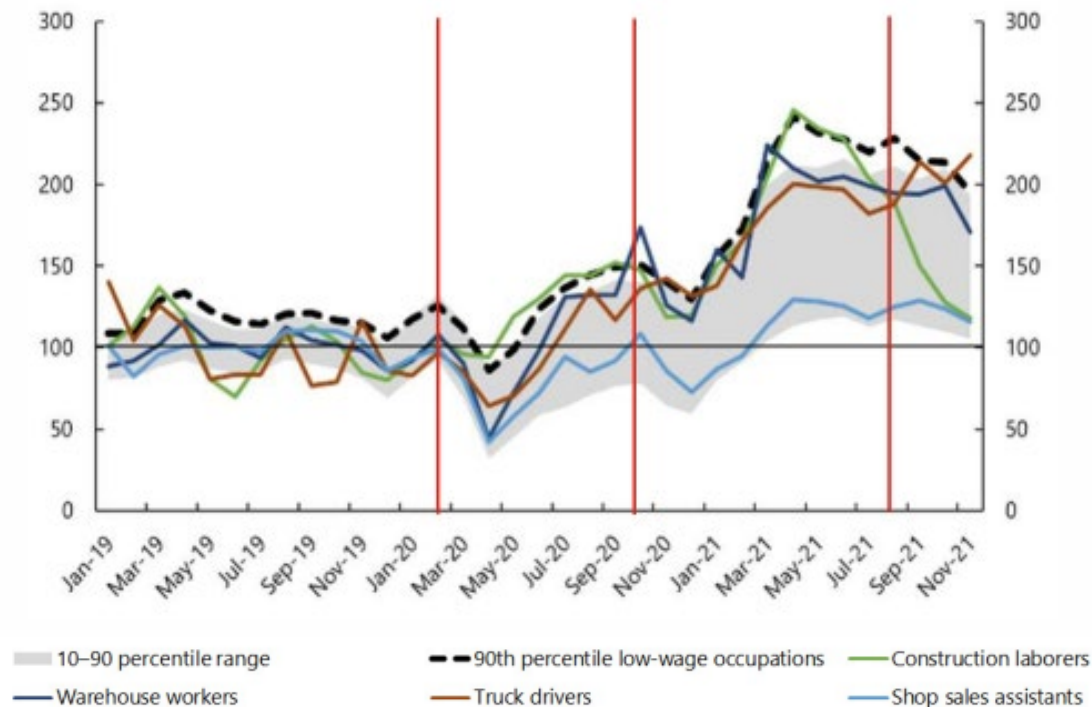
# Job Postings and Text Analytics

- **Job posting data** from Indeed and LinkedIn have been used to:
  - analyze the evolution of the greener employment
  - analyze tightness in the labor market post-pandemic.
- **Texts extracted from newspaper articles** (such as Financial Times) have been used to produce sentiment indicators

Figure 1.1: Vacancies in Selected Low-Wage Occupations

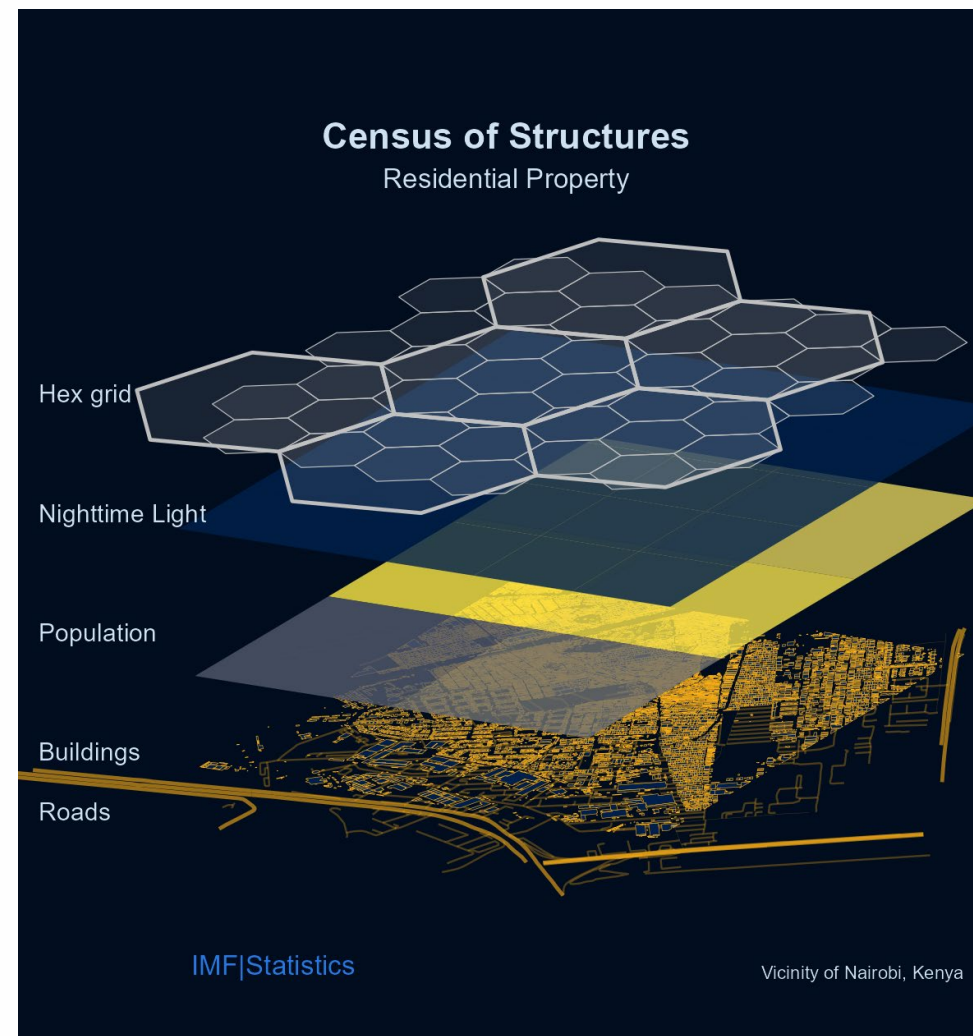
(Average 2019 = 100)

United States



# Climate Change Physical and Transition Risk Indicators

- Leverage open-source data to close **climate data gaps**
- Census of residential properties
  - Building footprint + Gridded population + Nighttime Light
- Application: Accurately assess climate risk, impacts, and damage to properties



## Other Examples



IMF staff produced an index for cross-country comparison of **travel costs** called “the Week at the Beach” index, based on hotel prices and occupancy rates.



**Mobility data** from Vodafone (i.e., people’s movement tracked by smartphones) have been used to assess the economic effects of the pandemic



**Digitalization: Broadband speed data** from Ookla have been used to estimate how digitalization impacts labor market resilience.



**Climate change:** Hundreds of billions of **weather measurements** have been analyzed based on machine learning methods to estimate the impact of weather shocks on GDP and fiscal variables

**Fostering not just the USE of non-traditional data sources but also ACCESS to them**

<https://datapartnership.org/>



## Data Partners

A Data Partner is an organization that provides data and/or services under the Master Data License Agreement.



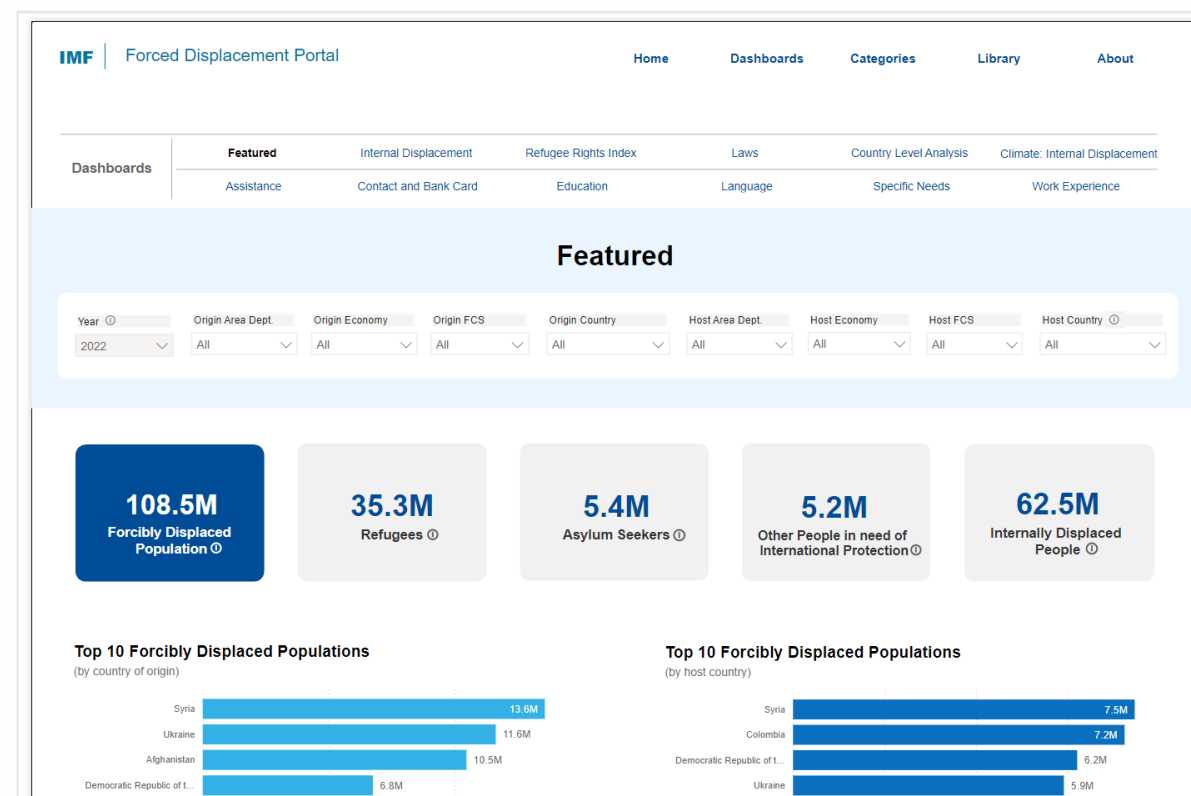
## Development Partners

A Development Partner is an international organization that receives data and/or services under the Master Data License Agreement.



# IMF Forced Displacement Portal

- Internal portal with data on forcibly displaced population
- Provides key open-source datasets to quantify the macroeconomic impact of global forced displacement (refugee spillover and internal displacement).
- Facilitating data accessibility with 23 partners, it identifies 70+ curated indicators with analytics and qualitative data leveraging traditional and non-traditional sources.



# **Looking Forward: Leveraging Big Data for Improving Macroeconomic Statistics in Member Countries**



# A regional approach to match IMF support for Big Data with specific needs of the region

- IMF Big Data Center and AfDB Data Innovation Lab are organizing a joint workshop on “Big Data for Macroeconomic Statistics” in Rwanda, hosted by the UN Regional Hub for Big Data in Africa (September 23-27, 2024)
- Statistics agencies and central banks and from 15 African countries
- Lectures and hands-on workshops covering various Big Data applications pertinent to macroeconomic statistics, offering insights from both producer and user standpoints
  - GDP nowcasting, Satellite data, Natural language processing, PortWatch
- Since March 2022, the IMF Institute for Capacity Development (ICD) has offered a regular **training course on nowcasting**, one of ICD’s most popular and oversubscribed, delivered 6 times between May 2023 and April 2024.



[COURSE DETAILS](#)

