Users face several challenges when accessing official statistics

Official statistics are published across multiple platforms, in multiple formats, using multiple discovery tools.

Official statistics are published using statistical language (which is not necessarily ‘natural language’).

Official statistics are presented in thematic tables (prices, health, education, economy), when most questions cross themes.

Official statistics are redistributed and then redistributed some more raising issues about vintages and questions about ownership.
SDMX and AI can help address these challenges

• SDMX data model contains descriptive metadata, i.e. names and description attributes in the code lists

• This allows effective interactions between LLMs and SDMX datasets without developing semantic layer from scratch
StatGPT application allows users to talk to data

- Natural language interface to search for statistical data and answer questions
- Breaks down silos between datasets
- Faster than faceted search
- Output is an SDMX query for data and metadata retrieval with certainty
- User is an integral part of the query-building process

You asked for indicators of economic growth and inflation for the United States [111]. I have created the following data query:

- Indicators:
  1. Gross domestic product, constant prices [WORLD_ECONOMIC_OUTLOOK.NGDP_RPCHMK]
  2. Gross domestic product, constant prices [WORLD_ECONOMIC_OUTLOOK.NGDP_RPCH]
  3. Gross domestic product, constant prices [WORLD_ECONOMIC_OUTLOOK.NGDP_R]
  4. Inflation, average consumer prices [WORLD_ECONOMIC_OUTLOOK.PCPI] [WORLD_ECONOMIC_OUTLOOK.PCPIEPCH]
  5. Inflation, average consumer prices [WORLD_ECONOMIC_OUTLOOK.PCPI]
  6. Inflation, end of period consumer prices [WORLD_ECONOMIC_OUTLOOK.PCPI] [WORLD_ECONOMIC_OUTLOOK.PCPIEPCH]

- Country:
  1. United States [111]

- Time period: 2020 - 2024
Next Steps

Beta Stat-GPT release this summer (statgpt.dialx.ai)

Tune Stat-GPT to selected SDMX sponsors’ data products

Provided restricted access for testing purposes (SDMX Sponsors, SIS-CC).

Test functionality, LLM costs, onboarding costs, support costs
QUESTIONS
ANNEX
Stat-GPT: How it works

Stat-GPT Agents:

- Build Data Query Agent
- Knowledgebase Query Agent
- Data Catalog Query Agent
- Guardrail Agent
- Visualizer Agent
- Help Agent

Orchestrator Agent builds execution sequence

Given a user can ask anything, we need several agents.

- Please give me data for GDP per capita for Baltic countries for the last five years.
  - Guardrail Agent → Build Data Query Agent

- What is the difference between constant price GDP and current price GDP?
  - Guardrail Agent → Knowledgebase Query Agent

- Who is Taylor Swift?
  - Guardrail Agent
Please give me data for GDP per capita in dollars for Baltic countries for last 5 years.

Dataset: IMF:IFS (1.0.0)
DSD: IMF:IFS_DSD (1.0.0)

Hierarchy: IMF:COUNTRY_GROUPS (1.0.0), Group: Baltics
Code List: IMF:COUNTRY (1.0.0), Codes: 939 (EE), 941 (LV), 946 (LT)

Indicators:
Code List: IMF:IFS_INDICATORS (1.0.0), Code: NGDPDPC

Countries:
Code List: IMF:COUNTRY (1.0.0), Codes: 939, 941, 946

Frequency:
Code List: IMF:FREQUENCY (1.0.0), Code: A [default]

Time Period: 2019-2023

I have modified existing data query in a following way:
• Added indicators: Gross domestic product, current prices, U.S. dollars, per capita [NGDPDPC]
• Added following countries: Estonia [939], Latvia [941], Lithuania [946]
• Set interval 2019-2023

SDMX Query: `<host>/sdmx21/data/IFS/939+941+946.NGDPDPC.A?startPeriod=2019&endPeriod=2023`
What is the difference between constant price GDP and current price GDP?

The main difference between the two is that constant price GDP provides a more accurate representation of economic growth by isolating the effects of price changes, whereas current price GDP can be influenced by inflation or deflation, making it less reliable for comparing economic performance across different time periods.

Who is Taylor Swift?

I'm sorry, but I'm designed to provide information related to Economics, Statistics, and SDMX standard. I can help you with queries related to these topics. Unfortunately, I can't provide information about individuals like Taylor Swift.