



StatGPT

Exploring Official Statistics using Generative AI and SDMX

CES MEETING

Bert Kroese, IMF

Users face several challenges when accessing official statistics



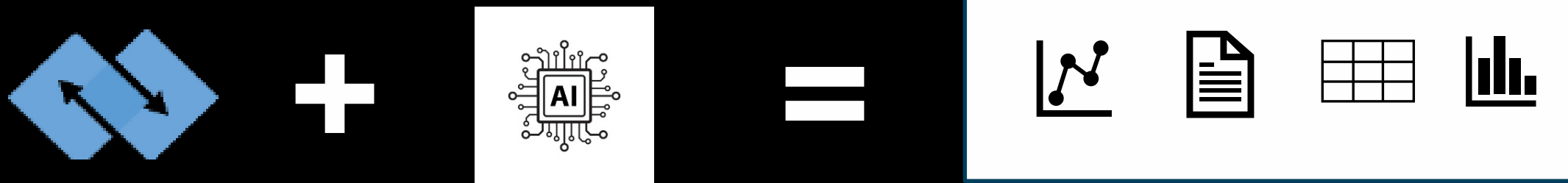
Official statistics are published across multiple platforms, in multiple formats, using multiple discovery tools.

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..

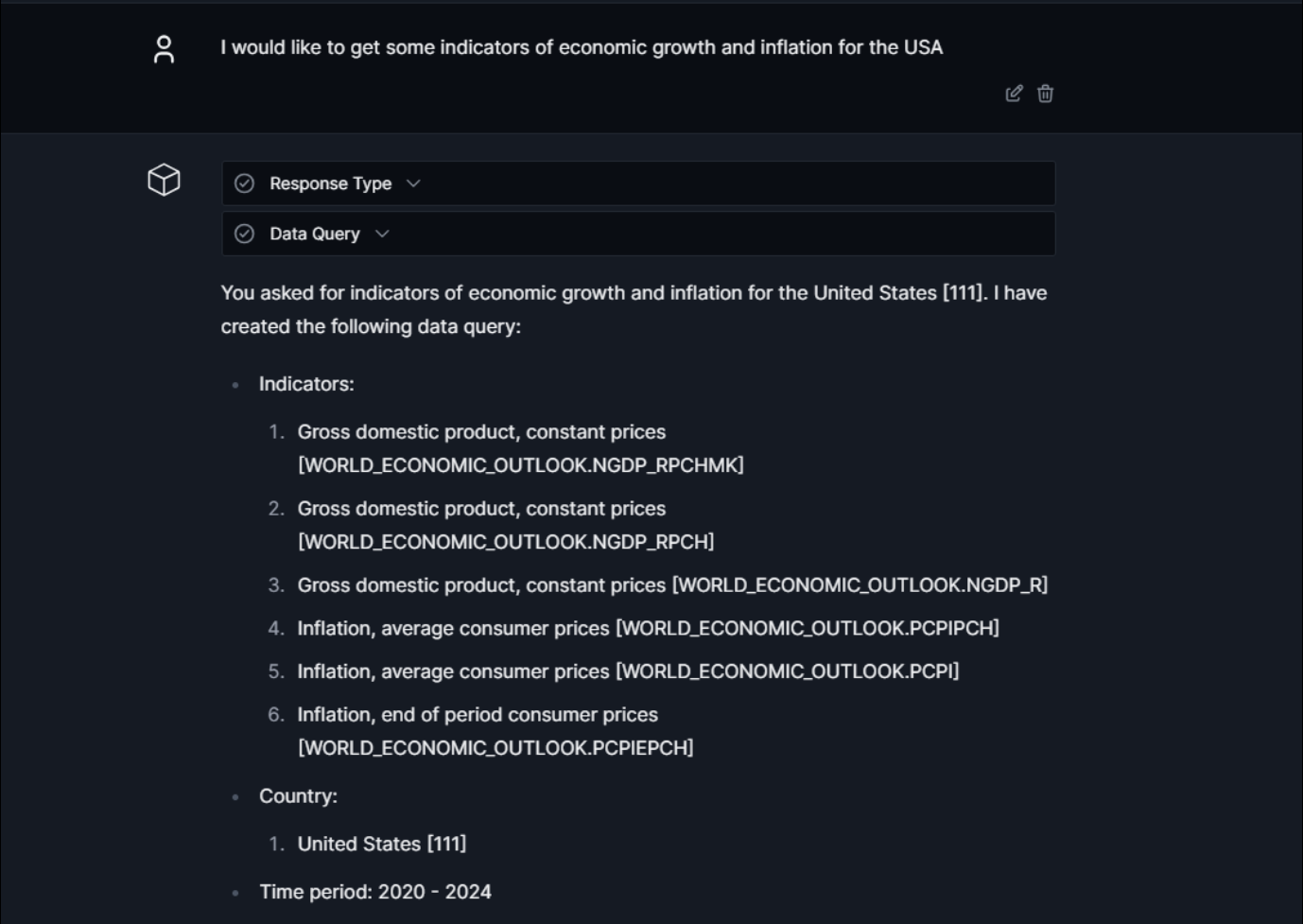
SDMX and AI can help address these challenges



- 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

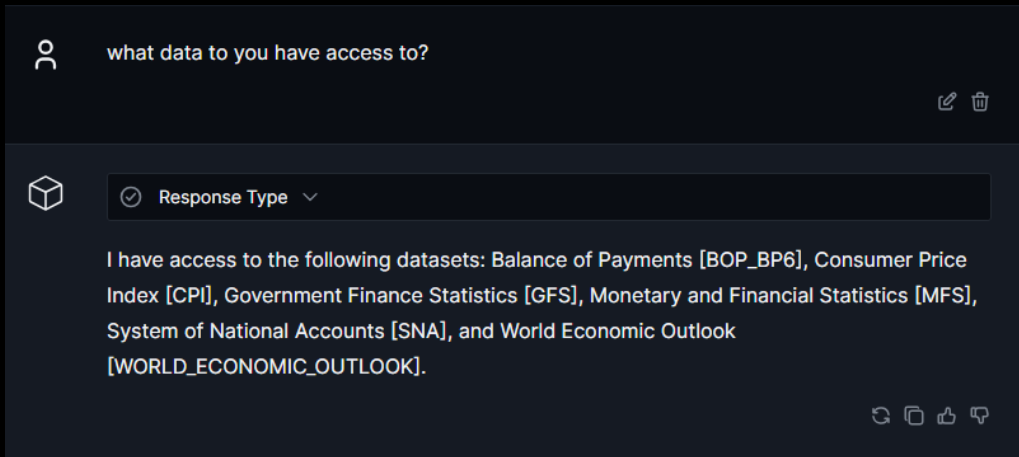
- Natural language interface to search for statistical data and answer questions
- Breaks down silos between datasets
- Faster than faceted search
- Output is an SDMX query for data and metadata retrieval with certainty
- User is an integral part of the query-building process



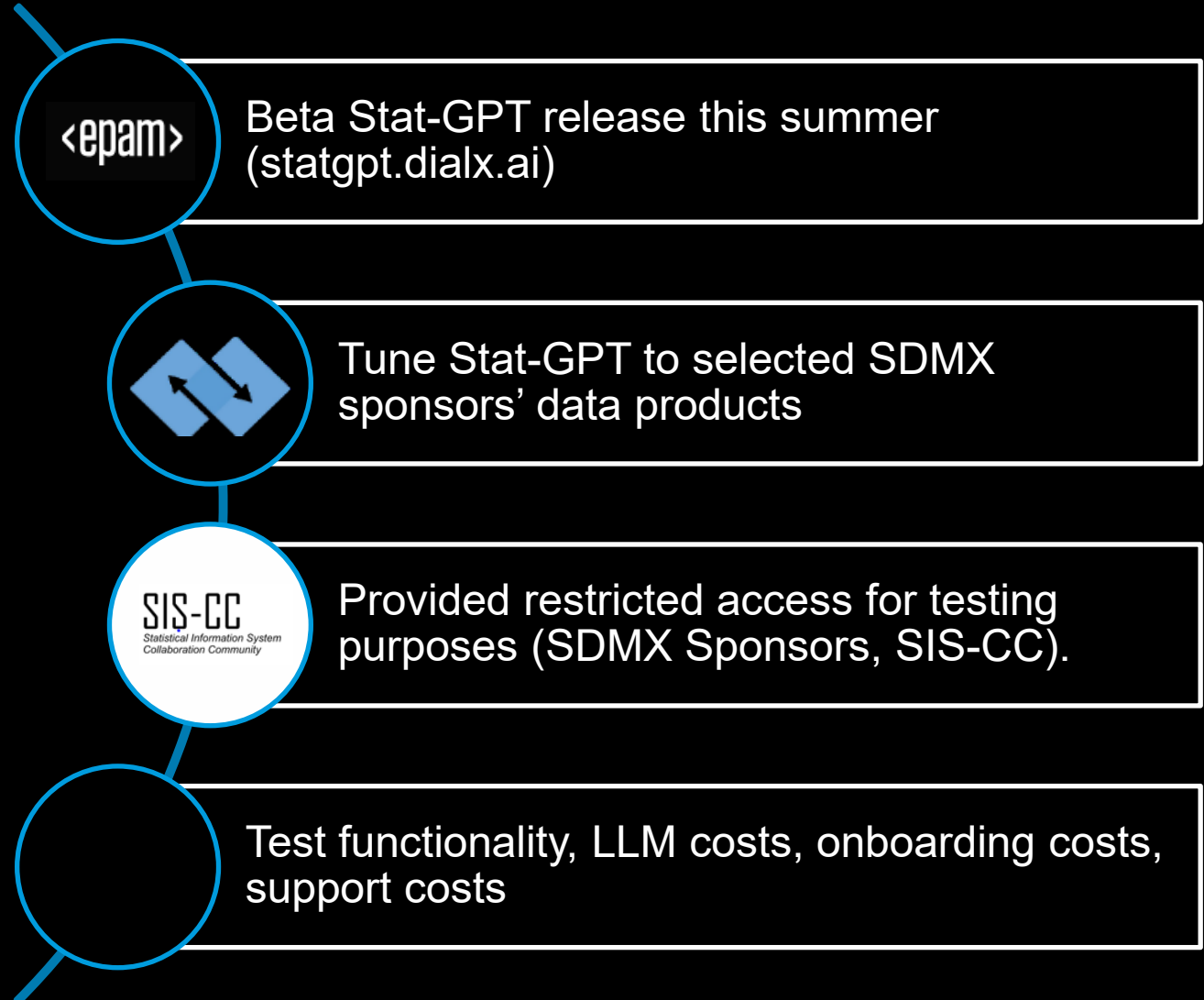
The screenshot displays the StatGPT application interface. At the top, a user profile icon is followed by the query: "I would like to get some indicators of economic growth and inflation for the USA". To the right of the query are icons for editing and deleting. Below the query, there are two dropdown menus: "Response Type" and "Data Query". The main content area shows the system's response: "You asked for indicators of economic growth and inflation for the United States [111]. I have created the following data query:". This response is structured into three categories: "Indicators:", "Country:", and "Time period:".

- 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.PCPIPCH]
 5. Inflation, average consumer prices [WORLD_ECONOMIC_OUTLOOK.PCPI]
 6. Inflation, end of period consumer prices [WORLD_ECONOMIC_OUTLOOK.PCPIEPCH]
- Country:
 1. United States [111]
- Time period: 2020 - 2024

Next Steps



We can help you **reimagine** your business through a digital lens.

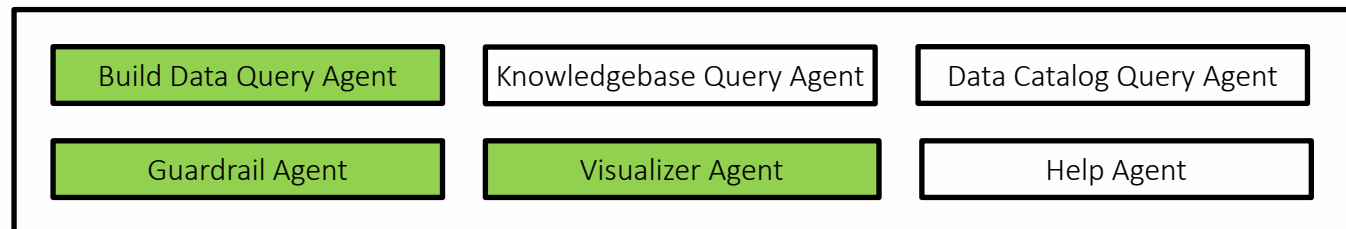


QUESTIONS

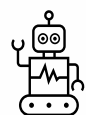
ANNEX

Stat-GPT: How it works

Stat-GPT Agents:



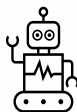
Given a user can ask anything, we need several agents.



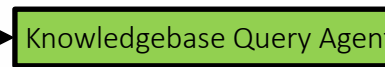
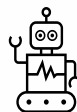
Orchestrator Agent builds execution sequence



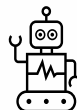
Please give me data for GDP per capita for Baltic countries for the last five years.



What is the difference between constant price GDP and current price GDP?



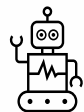
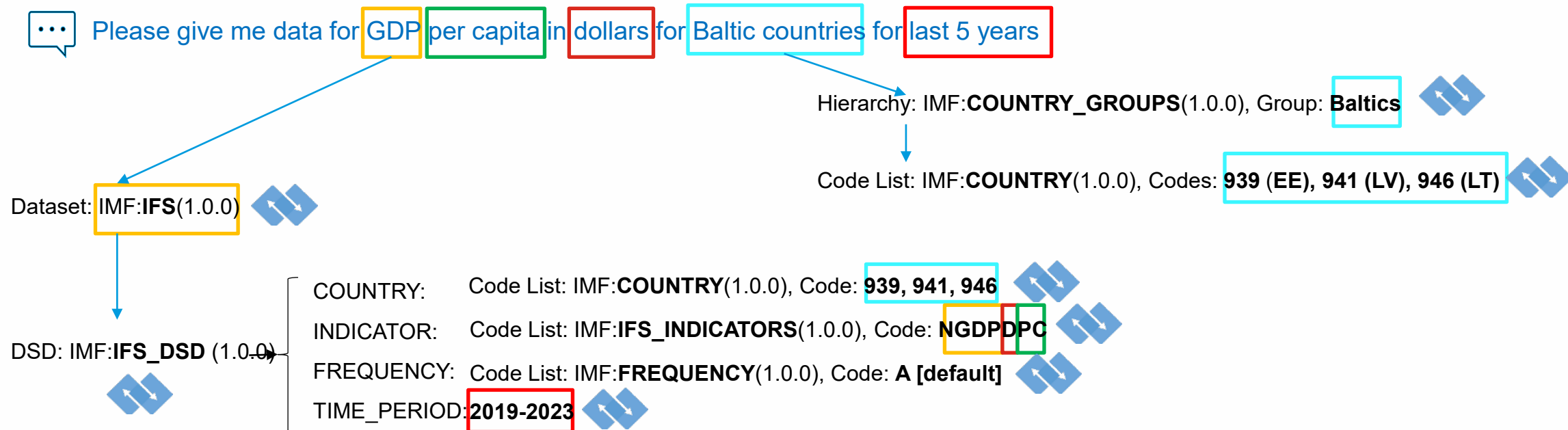
Who is Taylor Swift?



Stat-GPT: How it works



Please give me data for GDP per capita in dollars for Baltic countries for last 5 years



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

Stat-GPT: How it works

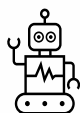


What is the difference between constant price GDP and current price GDP?

Orchestrator Agent

Guardrail Agent

Knowledgebase Query Agent



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.



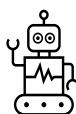
Who is Taylor Swift?

Orchestrator Agent

Guardrail Agent

not executed

Knowledgebase Query Agent



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.