



STATISTICS

How AI Is Transforming the Dissemination of Official Statistics

WTIS 2026 | 8 SEPTEMBER 2026

Marco Marini, IMF Statistics Department
Data Governance & Services Division (STADS)

Why This Topic Matters Now

- ❑ Users now expect conversational, instant, and contextualized access to data
- ❑ AI tools are rapidly changing how statistics are discovered and used
- ❑ Risk: official statistics lose visibility if they are not AI-accessible
- ❑ Opportunity: reposition official statistics as the trusted backbone of AI

The Core Challenge

- ❑ AI can democratize access to data — but it introduces real risks:
 - Hallucinated or incorrect numbers
 - No provenance or traceability
 - Authoritative and non-authoritative sources mixed together
- ❑ Key question: how do we make official statistics AI-accessible without losing trust?

What Happens Today with Standard LLMs

- ❑ General-purpose assistants (ChatGPT, Claude, Gemini) answer data questions, but often:
 - No clear source attribution
 - Outdated or inconsistent values
 - Limited dataset coverage
- ❑ Result: official statistics become just another input
- ❑ Takeaway: we need structured, machine-readable, authoritative data pipelines

From Dissemination to Intelligent Access

1

Traditional model

- Websites, static tables, and fixed APIs
- Users search for the data they need
- Data delivered as documents for people to read

2

Emerging model

- Chatbots, copilots, AI dashboards, automated reports
- Users ask questions in natural language
- Data delivered as systems that machines can act on

The shift: from searching for data → to asking questions of data

What Makes Data "AI-Ready"

- ❑ Machine-readable formats
- ❑ Rich metadata: concepts, definitions, units, time, and geography
- ❑ Semantic interoperability across datasets
- ❑ APIs for structured, programmatic access
- ❑ Provenance and versioning
- ❑ SDMX is the key enabler

SDMX: The Backbone of AI-Ready Statistics

- ❑ A global standard for structuring data and metadata, with shared vocabularies and API-based dissemination
- ❑ Why SDMX is critical for AI:
 - Provides structure that AI systems can understand
 - Enables cross-country comparability
 - Supports traceability and validation

SDMX in Practice: Enabling AI

- ❑ Structured datasets feed directly into AI pipelines
- ❑ Metadata enables:
 - Disambiguation (e.g., real vs nominal GDP)
 - Unit conversions and contextual explanations
- ❑ APIs allow real-time retrieval for AI tools
- ❑ Outcome: from documents for humans → data systems for AI

IMF Example: StatGPT and the GTDC

- ❑ The IMF approach combines SDMX + APIs + AI
- ❑ StatGPT: a natural-language interface to official IMF datasets
- ❑ Objective: responses that are accurate, traceable, and linked to authoritative data

Global Trusted Data Commons & StatGPT Go Live

Global Trusted Data Commons
Talk to Official Statistics

+ New Chat

All Chats

Last 7 days

☐ prepare table with gdp for united st...

Earlier

☐ compare the growth of china and th...

☐ tell me us gdp growth in 2025

☐ tell me what us gdp growth in 2025

What data are you looking for?

Search 9288 official indicators from the IMF, BIS and 130+ national statistical agencies

Ask about GDP, inflation, trade or any other economic indicator

What data is available?

What is the most recent global rate of inflation?

What is the latest Global growth forecast?

Compare the growth of China and the USA over the last 10 years

What is CPI?

What can you do?

Search across 9,288 official indicators — IMF, BIS and 130+ national statistical agencies

• Live now · portal.statgpt.org



Launched in June 2026

The IMF released the StatGPT conversational interface.



Talk to official statistics

Natural-language search across 9,288 official indicators.



Tier 0 datasets first

Initial coverage centres on IMF and BIS, with 130+ national summary pages indexed.

ONBOARDING STRATEGY

Tier 0 datasets first (SDMX Sponsors + ready NSOs) — then progressively onboard trusted official statistics, tier by tier.

Extending Access with MCP and APIs

- ❑ MCP (Model Context Protocol): a standardized way to connect AI models to data services
- ❑ Benefits:
 - Real-time access to official data
 - Reduced hallucination risk
- ❑ Example: India's MCP and others are exploring MCP-style integrations
- ❑ Building plug-and-play ecosystems for AI + official statistics

Other AI Tools in the Ecosystem

- ❑ Conversational chatbots for data access
- ❑ AI-generated dashboards and visualizations
- ❑ Automated press releases and report drafting
- ❑ Document and statistical-manual summarization tools
- ❑ Cross-dataset search assistants
- ❑ Key takeaway: a rapid move from prototypes → operational tools

In Practice: Examples from IOs and NSOs

1

SDMX AI Assistant (BIS)

- Chatbot for fit-for-purpose access to the full SDMX library — docs, guidelines and tools
- Built on BIS AI infrastructure; public since May 2026

↗ sdmx.io

2

OECD .Stat & AI (SIS-CC)

- Natural-language OECD Data Explorer pilot with an MCP service
- AI-readiness anchors the SIS-CC 2026–30 strategy; co-leads HLG-MOS “AI-Ready Dissemination”

↗ siscc.org

3

Talk2 Manuals (IMF)

- AI navigation of statistical manuals — RAG-based, with references for every answer
- Available to early adopters via the IMF Compilers Hub

↗ compilershub.imf.org

AI Readiness Across the Statistical Community

- ❑ **PARIS21** (OECD) — paris21.org
- ❑ **HLG-MOS** (UNECE) — unece.org
- ❑ **Trusted Data Observatory** (Swiss FSO) — trusteddataobservatory.org
- ❑ **Kigali Group** (UN Statistical Commission) — unstats.un.org
- ❑ **SDMX AI Task Team** (SDMX Sponsors) — sdmx.org

AI is not just tools — it requires ecosystem readiness

Working with Technology Companies

Reaching users in the AI era means partnering with the companies that build AI tools.



Why it matters

- AI assistants are becoming a primary gateway to data
- Official statistics must stay visible, citable and correctly represented
- Engagement shapes how trusted data appears in AI answers



Experience so far

- Google: UNSD [Data Commons](#); plus IMF and OECD collaboration
- Productive, but concentrated on a single provider
- Limited engagement so far with other LLM providers (OpenAI, Anthropic, Microsoft, Meta)

Priority: broaden partnerships across LLM providers for balanced, trusted representation of official statistics

Key Takeaways

- ❑ AI is transforming dissemination — from access to interaction
- ❑ Trust is the differentiator of official statistics
- ❑ SDMX is foundational for AI readiness
- ❑ Tools like StatGPT show what is already possible
- ❑ Collaboration across the statistical community is critical

Closing: A Call to Action

- ❑ Move from experimentation → implementation
- ❑ Invest in standards, metadata, and interoperability
- ❑ Work together so that official statistics remain the trusted source in the AI era



Q&A