Generative AI
in official statistics

High-Level Group for the Modernization of Official Statistics

Sprint on Artificial Intelligence and Data Science for Economic Statistics

United Nations Network of Economic Statisticians

Online, 12 December 2024

Vytas Vaiciulis (CSO Ireland), Olivier Sirello (BIS)

UNECE

modernstats
BY HLG - MOS

The project at a glance

- **13** statistical offices and central banks, 1 academia

- **3** international organizations, plus UNECE

- **37** members

🇳🇿	Pubudu Senanayake	🇨🇦	Joanne Yoon, Jason Solomon, Li Wang
🇬🇧	Samantha Trace, Fahad Anwar	🇮🇪	Vytas Vaiciulis (co-chair)
🇦🇺	Claire Clarke, Adam Leinweber	🇮🇹	Pierpaolo Massoli, Sara Letardi, Marco Silipo
🇪🇺	Cristiano Tessitore	🇸🇪	Jakob Engdahl
🌐	Jens Dossé, Jeanne Jadeau	BIS	Olivier Sirello (co-chair)
		UN	Amilina Kipkeeva, InKyung Choi

UNECE

modernstats
by HLG – MOS

Goals

- **Applications**: use cases and current implementations

- **Expectations** and priorities

- **Challenges**, limitations and risks

- Ultimately, **gaining a comprehensive view** on these three aspects with a report on "**Generative AI in official statistics**"

Stock-taking

1 Defining and setting up a template for sharing and classifying use cases

2 Creation of a UNECE repository

3 International survey for the Conference of European Statisticians

UNECE

modernstats
by HLG - MOS

Focusing on two distinct yet complementary aspects

Governance ⟷ **Technical**

Chapters

Building organizational capability

Using and implementing

Governing and managing

Mitigating and monitoring

Sharing

Preparing for the future

UNECE

modernstats
by HLG - MOS

Chapter 1 - Building organizational capability

Six building blocks:

1. **Data**: making them AI-ready

2. **Technology**: preparing for the wealth (and overload) of unstructured data

3. **Governance**: finding the right balance between innovation and compliance

4. **Skills**: the road towards AI literacy and AI proficiency

5. **Cultural**: breaking down silos with data at the center as strategic asset

6. **Legal**: adapting to quick changes in a polarized and uncertain environment

UNECE

modernstats
by HLG - MOS

Chapter 2 - Using and implementing

- **Where to use Gen AI:** identify target processes, domains, and workflows (eg across GSBPM)

- **How to implement Gen AI:** internal development, external partnerships, or managed services

- **Criteria for using Generative AI:** business value, requirements, technical feasibility

- **Prompt engineering:** apply targeted strategies for effective prompts, emphasizing clarity, context, and specificity, and address working with prompts in multiple languages

UNECE

modernstats
by HLG - MOS

Chapter 2 - Using and implementing

- **Model selection:** tailor solutions to domain specificities or opt for general-purpose models, and assess available options (including open-source)

- **Architecture and infrastructure:** choose frameworks, tools, and stacks that deliver the right balance of performance, scalability, and cost

- **Deployment options:** weigh the flexibility of cloud services against the control and compliance benefits of on-premise or hybrid setups

- **Scalability and performance:** plan for sustainable growth, maintaining responsiveness while managing costs

UNECE

modern stats
BY HLG - MOS

Chapter 3 - Governing and managing

- **Adapting structures, frameworks and processes**: breaking down silos and achieving coordination within and across the ecosystem with data at its center

- **Finding the right balance between top-down and bottom-up** models: centralized, decentralized, federated, hybrid and spoke-hub

- **Coping with multiple (conflicting) objectives**: offensive (innovation, growth) vs defensive (risks and uncertainties, compliance)

- **From governance to stewardship**: the central place of ethics

UNECE

modernstats
BY HLG - MOS

Chapter 4 - Mitigating and monitoring

A **broad palette of risks**:

- Reliability, **traceability** and transparency
- **Robustness**, security and safety
- **Accountability**, ethics and privacy
- **Sustainability**
- Ultimately, a risk for **information** and **trust** with huge **social costs**

Building mitigation strategies is an imperative:

- Technical
- Policies
- Governance

A **call for monitoring and evaluating the use of generative AI** in official statistics

UNECE

modernstats
ᵇʸ HLG - MOS

Chapter 5 - Sharing

Sharing is a critical requirement to fully harvest the benefits of generative AI while also mitigating its risks:

1. Sharing **data**: balancing (meta)data openness for better traceability with data sovereignty and safety

2. Sharing **tools**: open-source software as fundamental backbone of open AI

3. Sharing **knowledge**: guidelines, frameworks and best practice to mature governance at the institutional level and achieve data literacy/proficiency at the individual level

UNECE

modernstats
BY HLG - MOS

Chapter 6 - Preparing for the future

In the fast-paced technological landscape, it is critical for statistical offices to act (on top of reacting):

1. Preparing for **new ways of working**: (generative) AI will not only affect official statistics, but statisticians and their interactions with non-human users

2. Preparing for **new ways of accessing, retrieving and generating information**: the primary focus on information, on top of data

3. Preparing for **new threats**: countering mis- and disinformation in an increasingly uncertain and polarized landscape, cybersecurity, super artificial intelligence

UNECE

modernstats
by HLG - MOS

Thank you for your attention

vytas.vaiciulis@cso.ie

olivier.sirello@bis.org