Irving Fisher Committee on Central Bank Statistics

12th biennial IFC Conference: "Statistics and beyond: new data for decision making in central banks"

22-23 August 2024

Generative AI and official statistics: the project of the UNECE High-Level Group for the Modernisation of Official Statistics[1]

InKyung Choi and Amilina Kipkeeva,
United Nations Economic Commission for Europe

Olivier Sirello,
Bank for International Settlements

Vytas Vaiciulis,
Central Statistics Office Ireland

[1] This contribution was prepared for the conference. The views expressed in this publication are those of the authors and do not necessarily represent the official views of the Committee, its members, or the BIS.

Generative AI and official statistics: the project of the UNECE High-Level Group for the Modernisation of Official Statistics

InKyung Choi, Amilina Kipkeeva, Olivier Sirello, Vytas Vaiciulis[1]

Abstract

There is little doubt that generative artificial intelligence (AI) has the potential to play an important role in official statistics, as it can boost productivity across the statistical production chain. However, the rapid pace of change presents challenges for its adoption by statistical offices, including central banks. As producers of official statistics, central banks ensure that the use of these technologies complies with the Fundamental Principles of Official Statistics, particularly upholding trust with data providers and users. Pooling experiences and knowledge from statistical organisations is invaluable to facilitate the adoption of the new technology.

The High-Level Group for the Modernisation of Official Statistics (HLG-MOS) of the United Nations Economic Commission for Europe has been actively steering the modernisation of official statistics. In this paper, we present the HLG-MOS project on generative AI. Building on earlier work, the project aims to explore prominent use cases for statistical offices while establishing best practices and approaches for governance, project management and architecture.

[1] InKyung Choi and Amilina Kipkeeva are Statisticians at the United Nations Economic Commission for Europe. Olivier Sirello is a Senior Statistical Analyst in the Monetary and Economic Department at the Bank for International Settlements. Vytas Vaiciulis is a Lead Data Scientist at the Central Statistics Office of Ireland.

The views expressed here are those of the authors and do not necessarily reflect those of the United Nations Economic Commission for Europe, the Bank for International Settlements or the Central Statistics Office of Ireland.

Contents

1. Introduction 3

2. Generative AI in official statistics: a view from the top on the key opportunities and limitations 4

3. The UNECE High-Level Group for the Modernisation of Official Statistics: the project on generative AI for official statistics 5

4. Taking stock of use cases across statistical organisations 6

4.1 Data processing and analysis 7

Box A: Survey on the use of generative AI in statistical organisations 8

4.2 Code and IT development 9

4.3 Communication of official statistics 10

4.4 Data dissemination and information search 11

4.5 Other applications 12

5. Building a unified framework for sharing experiences, best practices and knowledge on generative AI for official statistics 12

6. A roadmap for preserving trust in official statistics in the age of AI 14

Annex 15

References 17

1. Introduction

Generative artificial intelligence (AI) holds the potential to revolutionise the field of official statistics. First, for producers of data, it can automate a wide range of manual tasks throughout the statistical production chain and streamline key phases, such as data collection, editing, processing and validation. This automation enhances efficiency, reduces the burden on human resources and improves the accuracy of statistical outputs. Second, generative AI offers innovative solutions for users of data as well, by simplifying queries and improving data accessibility, discoverability and findability. This dynamic interaction ensures that user needs are better understood and met. Finally, generative AI enables official statistics to engage more actively in areas such as communication that previously received less attention than the core business of statistical organisations. By leveraging AI, statistical offices can enhance their outreach and better disseminate their findings to the public and, more importantly, bridge the gap between producers and users.

This revolution has already found concrete applications among statistical organisations and central banks. The project on generative AI for official statistics, under the aegis of the United Nations Economic Commission for Europe (UNECE) High-Level Group for the Modernisation of Official Statistics (HLG-MOS), primarily serves as a forum for organisations and practitioners to share experiences, map the footprint of generative AI across the field and explore its potential and limitations. Through this collaborative effort, statistical offices have demonstrated that they have already implemented AI-driven solutions in several key processes.

The imperative to pool knowledge and experiences has never been more compelling, as generative AI advances rapidly and is not free of risks. The limitations associated with its implementation are significant. Operationally, the use of AI introduces black box algorithms that lack transparency, making it difficult to trace outputs back to inputs and vice versa. Scarce traceability ultimately poses a threat to the core principles of official statistics, such as accuracy, impartiality and, more fundamentally, trust. Ensuring that AI applications adhere to these principles is crucial to maintaining the integrity of official statistics.

This paper serves three main purposes. First, it introduces the Generative AI for Official Statistics project under the High-Level Group for the Modernisation of Official Statistics. It also provides an overview of both the key opportunities and limitations associated with this frontier technology. Second, it takes stock of and summarises the diverse collection of use cases gathered by the project, spanning several areas, including code assistants, data editing, dissemination and communication of official statistics. Finally, the paper underscores the critical importance of developing comprehensive frameworks to share knowledge and establish international best practices, both in terms of governance and implementation. These frameworks would serve to effectively and safely advance the adoption of generative AI in official statistics.

2. Generative AI in official statistics: a view from the top on the key opportunities and limitations

Generative AI offers significant potential benefits for official statistics, promising improvements across the entire statistical production chain (UNECE (2019)). First, by leveraging this frontier technology, statistical organisations can enhance productivity and streamline manual operations, enabling more efficient resource allocation and improved service delivery. Large language models (LLMs), a subcomponent of generative AI, are particularly proficient in generating human-like text, summarising large volumes of information and producing reports (HLG-MOS (2023)). They have the potential to streamline a wide range of routine tasks in statistical organisations, thereby freeing up human resources for more strategic and complex activities. Second, generative AI shows great promise in the areas of data processing, standards and analytics. Typical use cases range from validating and classifying data and updating statistical standards to generating synthetic data sets for safe exchange. Perhaps more importantly, this technology also allows the production of charts and graphs, while helping users to refine their data queries and code. This aspect is critical to bridge users and producers of data, particularly when using granular data, as data retrieval might be particularly complex for non-expert users. Finally, dissemination and communication of official statistics can also benefit greatly from this innovation wave. Solutions such as AI-powered chatbots, interactive queries and customised information delivery can help users better navigate statistical technical jargon and access the data, thereby improving statistical literacy (IFC (2024b)). Thanks to generative AI's ability to generate audio and video, some statistical organisations already employ this technology to improve their communication content, foster engagement with users and, ultimately, expand the outreach of official statistics.

Despite these promising opportunities, the integration of generative AI in official statistics also comes with several challenges. These include high uncertainty and rapid technological changes, limited transparency and, more fundamentally, critical considerations in terms of ethics and misuse of data. First, the rapid pace of technological advancements in AI presents a significant challenge for statistical organisations. It can indeed create an environment marked by high risks and volatility where fast decision-making becomes a pressing need to cope with the pace of technological development. As a result, strategic foresight has never been so critical within and across statistical organisations, not only to plan in advance for the emerging needs but, more critically, also to prevent or minimise failures. A second strand of limitations has to do with the poor transparency and traceability of the technology. This implies opacity in the algorithms – often referred to as black boxes – and processes used to generate outputs from inputs or vice versa (Garrett (2024)). It also entails a lack of visibility in terms of technological developments, inflating operational risks in terms of vendor lock-in, production disruptions and, ultimately, business continuity (HLG-MOS (2024a)). Finally, generative AI leads statistical organisations to face new types of ethical limitations. This technology can inadvertently reproduce biases and errors present in the training data, leading to biased or incorrect outputs. Moreover, language divides are even more apparent with AI since the majority of the data used for its training is in English, posing a number of cultural challenges (OECD (2024)). For official statistics, these limitations of generative AI are anything but trivial. They have the potential to seriously undermine key aspects rooted in the Fundamental Principles of Official Statistics (UN (2014)), notably accuracy, authenticity, impartiality and, more fundamentally, trust. Pooling experiences and best practices across the official statistics community is therefore critical.

3. The UNECE High-Level Group for the Modernisation of Official Statistics: the project on generative AI for official statistics

With the pace of technological change accelerating, collaboration and sharing the experimental burden have become increasingly important. At the early stage of innovative technology adoption, individual organisations often have limited resources available to navigate the full potential of the technology on their own. Also, as public agencies, statistical organisations face common challenges and concerns such as alignment with quality criteria and integration of the technology within legacy information technology (IT) systems. This makes pooling experiences and knowledge among different organisations invaluable to facilitate the adoption of innovative solutions.

To cope with these challenges, **the UNECE High-Level Group for the Modernisation of Official Statistics (HLG-MOS) serves as an international platform** to exchange experiences and lessons learned. Established by the Conference of European Statisticians (CES) in 2010 to advance the modernisation of official statistics, HLG-MOS is led by chief statisticians of 13 national and international statistical organisations[2] who set the vision, mission and priority topics and approve the work programme.[3] Driven by bottom-up ideation and top-down steering, it aims to ensure voluntary collaboration, flexibility and agility while aligning with a broad vision of modernisation for official statistics. The community of experts engaged in the HLG-MOS work (ModernStats community) spans several national statistical offices, central banks and other international organisations, including the Irving Fisher Committee on Central Bank Statistics (IFC (2024a)). It engaged over 500 members in its various meetings in 2023.

Undoubtedly, frontier technologies such as machine learning and generative AI have been high on the agenda of HLG-MOS. Discussions on machine learning emerged within the ModernStats community in the mid-2010s and gained significant momentum following the position paper in 2018 (HLG-MOS (2018)) and subsequently with the creation of the HLG-MOS Machine Learning Project (2019–20) and the Machine Learning Group led by the UK Office for National Statistics and UNECE (2021–22).[4] With the fast expansion of generative AI, HLG-MOS began fostering international collaboration efforts around the LLMs, delivering the first HLG-

[2] Statistical offices of Australia, Canada, Ireland, Italy, Mexico, the Netherlands (co-chair), New Zealand, Poland (co-chair) and the Republic of Korea, as well as Eurostat, the Organisation for Economic Co-operation and Development and UNECE.

[3] It is the custodian of some of foundational standards for modernising official statistics production, such as the Generic Statistical Business Process Model (GSBPM) or the Generic Statistical Information Model (GSIM), and it spearheaded collaboration on topics of growing strategic importance such as data governance for interoperability (see UNECE (2024a)), strategic communication (see UNECE (2021)) and cloud computing (see UNECE (2024b)).

[4] For more information, see the UNECE Machine Learning for Official Statistics wiki (statswiki.unece.org/display/ML).

MOS white paper on the implications and opportunities for official statistics (HLG-MOS (2023)).

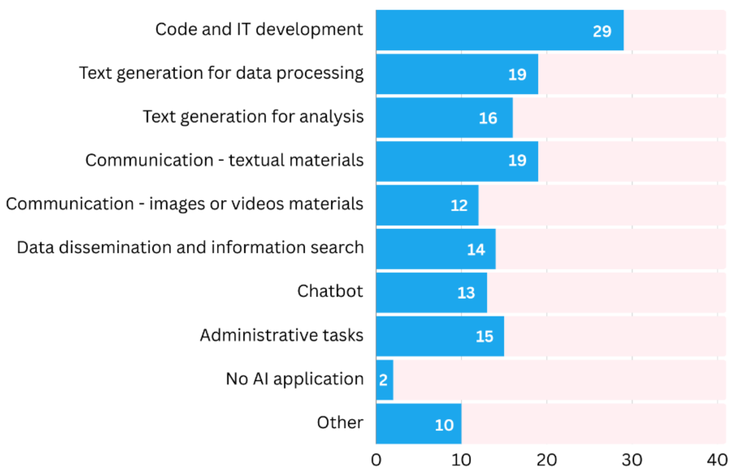
Drawing on these efforts, HLG-MOS launched the project on generative AI to shed light on governance and implementation aspects, taking stock of experiences among statistical organisations. Starting from use cases shared by statistical offices,[5] the first goal is to investigate the strategic considerations that arise when statistical organisations seek to use generative AI effectively and responsibly (eg governance, open-source models). Here, critical questions revolve around governance and ethics (analysis of security, legal and ethical risks and mitigation) as well as project management and the development journey (organisational aspects). The second purpose is to explore opportunities to co-develop concrete solutions and shed light on strategies related to prompt engineering (best practices and quality assurance) and the architecture and IT stack (infrastructure and tools).

4. Taking stock of use cases across statistical organisations

Although generative AI is a relatively recent development, statistical offices and central banks have already begun experimenting with it, providing valuable use cases for exploration. In fact, results from a CES survey clearly show that virtually all statistical organisations are using or experimenting with AI in their business processes, with those not using it constituting only a small share (5%) (Box A). Generative AI applications can be categorised into several key areas: data processing and analysis, code translation and generation, data dissemination and information search, as well as communication (Graph 1). Each category offers unique benefits and opportunities for innovation.

Application areas for generative AI by statistical offices

Graph 1



Note: the graph shows the number of application areas within organisations using generative AI (multiple selection). Use cases include both production-grade and pilot projects.

Source: UNECE.

[5] The generative AI project created a use case repository (unece.github.io/genAI/).

4.1 Data processing and analysis

Data processing and analysis are the most popular areas in the use of generative AI by statistical organisations. These include a wide range of tasks, such as editing and classification, but also analysis of the outputs. The CES survey indicates that 46% of the respondents are using or experimenting with AI-based solutions for data processing, such as editing, synthetic data generation and classification of codes, and 39% for analysis, such as producing reports or text summaries from news articles.

An area of ongoing exploration is data and metadata editing. Editing typically involves the correction of missing or outdated values through imputation methods. This process is no doubt extremely important for statistical organisations, not least because it ensures the highest quality of the statistical outputs. Here, some statistical offices indicate that they use or explore solutions leveraging generative AI to perform data imputation. Generative AI is also particularly cost-effective to edit metadata. These typically are "documentation about the data", for example, describing the underlying compilation, quality and methodology. Although metadata play a critical role in statistics, not least to ensure correct interpretation of the data, their processing is particularly challenging for compilers. In fact, editing tasks are often time-consuming, labour-intensive and prone to errors for compilers. Also, as metadata often come as unstructured data or text, automation and standardisation are difficult to achieve, lowering processing efficiency. Using generative AI can be a game changer in this regard because of the advanced ability of LLMs to generate human-like sentences due to the underlying technology of transformers. A concrete solution tested by the **Bank for International Settlements** (BIS) includes extensive metadata as part of its statistical offering in the BIS Data Portal.[6] The BIS Data Portal team has developed an AI-powered assistant, called the BIS AI Metadata Editor, to automate the metadata editing process (Sirello (2024)). This assistant applies English capitalisation rules, ensures the correct spelling of institution names, performs comprehensive formatting cleanups and checks for grammar and syntax errors. This AI-driven approach significantly reduces the time and effort required for metadata editing, enhancing consistency and accuracy.

Another strand of applications is the use of AI to assist with generating, updating and maintaining statistical classifications. Statistical organisations employ various statistical classifications, such as the Classification of Individual Consumption According to Purpose (COICOP) and the Statistical Classification of Economic Activities (NACE) to ensure comparability across time and domains as well as between countries. Regular updates to these classifications are crucial to reflect changes in the corresponding domain and maintain their relevance. LLMs can be used to draft initial versions of category descriptions, which are then reviewed by human experts. For example, the **Australian Bureau of Statistics** used LLMs to update occupation codes, significantly reducing the manual workload and potentially saving over 1,600 working hours. A draft version of the updated ANZSCO classification was released for stakeholder feedback, including full disclosure of the use of generative AI and details of the human-centred review process. Quality and legal concerns were addressed through thorough reviews and consultations. Notably, no stakeholders

[6] The BIS Data Portal is available at data.bis.org.

Box A

Survey on the use of generative AI in statistical organisations

As part of the 72nd Conference of European Statisticians in 2024, the UNECE and HLG-MOS project on AI organised an electronic consultation with national statistical offices and international organisations. The survey aimed to collect information on use cases of generative AI, related organisational policies and challenges. Thirty-seven national and four international organisations that are involved in official statistics took part in this exercise (HLG-MOS (2024b)).

This box presents an overview of the survey results with **six main highlights:**[7]

1. **There is wide agreement that generative AI has or is expected to have a high impact on statistical organisations, particularly on efficiency and productivity**. Responses show that, over the next two or three years, IT-related areas, including coding, will be the most impacted, closely followed by dissemination and communication. Areas such as collection, processing, analysis and administrative functions are also expected to be moderately impacted. Further, statistical organisations anticipate that efficiency and productivity will be the most impacted, followed by enhancement of service delivery to users and fostering creativity.
2. **Virtually all organisations have used generative AI, although with significant variation across application areas**. Specifically, code/IT development as well as data processing and analysis are by far the most popular applications. Most of the identified solutions revolve around code translators or generators, as well as synthetic data generation, maintenance of statistical classifications and report generation. Statistical organisations have also started to adopt generative AI for data dissemination, search and communication, including for improving search engines, data accessibility through chatbots and generating textual (eg press releases, social media posts) and non-textual (e.g. video, audio and images) content.
3. **The majority of statistical organisations do not yet have a policy in place for governing generative AI**. In fact, results show that few statistical organisations have established guidelines, while the bulk of them are still developing and/or have not formalised their approach. Common themes include data security, transparency and ethical use, focusing on minimising risks, ensuring transparency and promoting accountability. Some countries' frameworks and ethical charters also guide AI use.
4. **Statistical organisations nevertheless express significant concerns regarding risks, particularly in terms of data security, privacy and accuracy**. Furthermore, ethical considerations, copyright and legal issues are of substantial importance to the majority of respondents. Additionally, the potential for misuse and the lack of reproducibility are major issues.
5. **To mitigate those risks, statistical offices are taking various approaches**. These include development of internal policies, staff training and self-hosted solutions, particularly to minimise the sharing of confidential information. Investments in infrastructure, including hardware and data centres, remain a key limitation for certain countries. Some organisations also prefer using open source software, especially to control and reduce dependency on external providers. More fundamentally, there is a clear need for internal capacity building, particularly to raise awareness about the risks of using generative AI with sensitive information. Most of the respondents consider the availability of staff with appropriate skills and capabilities to use generative AI as a limiting factor.
6. **Ultimately, stronger international cooperation is required**. Results underline strong demand from offices for more international cooperation along several dimensions, including sharing knowledge, best practices and training resources, as well as practical examples, experiences and use cases. Perhaps more importantly, there is demand for international cooperation to develop common standards and guidelines.

raised concerns about the use of generative AI, demonstrating that the transparency and rigorous review process ensured strong governance and positive public

7 A detailed analysis of the survey is available at unece.org; see HLG-MOS (2024b) and UNECE (2024c).

perception (HLG-MOS (2023)). The ability of LLMs to generate data that fulfil certain characteristics indicates their potential to be used for synthetic data – a topic of growing importance in statistical organisations, not just for internal operational purposes (eg improving an imbalanced sample for training ML models) but also for the provision of micro-level data for the users while ensuring data confidentiality (UNECE (2022)). The CES survey indeed showed that several statistical organisations are working on generating synthetic data using LLMs. LLMs can also be used to classify free-text responses into statistical classification codes, improving the efficiency of this highly resource-intensive task in the production process. For example, the **UK Office for National Statistics** has developed an experimental text classification pipeline – "ClassifAI" – to code textual responses from labour market surveys into the Standard Industrial Classification (ONS (2024)).

Finally, generative AI is also transforming data analysis in statistical organisations and central banks, for instance by generating reports and presentations through metadata and data-driven pipelines. The **Deutsche Bundesbank** (Germany) developed the Text-based Intelligent Assistant (TIA) platform, using various versions of ChatGPT on Azure Microsoft Cloud Service. TIA supports both English and German, allowing for generative queries and document processing with flexible access control (Düker (2024)). In the same vein, **Statistics Canada** is experimenting with LLMs to generate draft reports from Census data that are reviewed for accuracy (HLG-MOS (2023)). The process includes generating summary statistics and drafting articles, with improvements being needed in style and conciseness. The team is also exploring bilingual report translations and secure AI practices.

4.2 Code and IT development

Code generation and translation are areas where generative AI has already been proven to be particularly critical for organisations compiling official statistics. Generative AI can assist users in a number of coding tasks, typically through coding assistants or co-pilots. Some models, such as Codestral, are specialised in code tasks. They can automatically generate code snippets based on user input, with significant gains in terms of precision, quality and robustness of the code. They can also help with debugging, generalising scripts and documenting pieces of code to enhance understanding. These features are critical for statistical offices and central banks, not least because many of them have embarked on open source technology transition projects, including migrating from proprietary statistical software, such as SAS and SPSS, to open source programming languages, such as R and Python (CSO (2023), IFC (2023), Statistics Netherlands (2023)).

A large share of statistical organisations indicate that they use generative AI for coding. According to the CES survey, 71% of the respondents are experimenting with or using AI solutions for coding, such as co-pilots, code generators and translators. The latter refers to assistants that help translate scripts from one language to another. Noteworthy projects, including an experimental phase, are at the **National Institute of Statistics and Economic Studies of France** (Eusebio and Givernaud (2023)) and the **Central Statistics Office (CSO) of Ireland**. The CSO has developed the "SAS to R Transcompiler" using OpenAI's application programming interface (API) to automate the translation of SAS code to R (HLG-MOS (2023)). The application is currently in production and has been open sourced, available on GitHub.[8] The tool converts SAS code into R by leveraging GPT-4, tailored to the specific packages used by the CSO. The primary advantage is saving time, with significant reductions in the processing time for the translation. Also, the AI-powered translators used by the CSO have proven effective in providing high-quality outputs requiring minimal manual correction. User-friendly applications, built with R's Shiny framework, allow users to input SAS code and instantly receive R code. There are also some challenges, starting with addressing privacy concerns. RegEx is used to detect and block sensitive information during translation. Future developments include adding features like code explanation, script writing and potentially running the model locally with open source LLMs.

4.3 Communication of official statistics

Undoubtedly, AI helps foster clearer communication, better engagement with the public and, ultimately, trust in official statistics. For example, visual and audio content summarising key statistics can reach broader audiences, ensuring public confidence in official statistics and supporting literacy efforts (IFC (2024b)). On a more practical level, a key goal is to simplify technical jargon to plain language that is easily accessible to analysts. Chatbots can provide a more intuitive and user-friendly interface for answering questions about the data or metadata. Many respondents to the CES survey express interest in exploring or using AI to improve their communication of official statistics. More specifically, 46% indicate using AI for producing textual materials, such as press releases or social media posts, while 29% use it for non-textual content, such as images and videos.

The use of LLMs for generating social media content is also diverse.[9] For example, the **Statistics Division at UNECE** already employs LLMs to help create social media posts, such as event announcements and project introductions. By training LLMs with manually written posts, organisations can maintain coherence and brand recognition in their social media content. In the same vein, other organisations, such as **Statistics Indonesia**, indicate ongoing work in using generative AI for social media content creation (Faris (2023)).

Additionally, AI can be leveraged to generate non-textual content, such as audio, images and videos. By manipulating prompts, statistical organisations can create images that maintain a coherent style and use brand colours, ensuring visual consistency across platforms. This allows for the efficient production of high-quality visual content and stories, supporting data dissemination and communication efforts. For example, **Statistics Canada** is testing approaches to generate content with narration and visuals, expanding data accessibility both internally and externally.[10]

[8] Codes are available at github.com.

[9] In addition to the generative AI project, HLG-MOS has a task team under its Capability and Communication Group dedicated to communication aspects. The task team is exploring how to enhance the productivity of communication experts with the use of AI and how to communicate the use of AI to the public while maintaining trust in official statistics. The use cases presented here were featured in the task team’s work.

[10] A demo is available at vimeo.com.

4.4 Data dissemination and information search

An additional area spans across data dissemination, particularly to improve data discoverability, findability and searchability. Users of official statistics often face challenges navigating multiple discovery tools, dealing with domain-specific tables and, more broadly, searching for data. Around one third of statistical organisations involved in the CES survey indicate interest in leveraging this frontier technology to improve information searches, queries and dissemination, including for internal purposes.

Generative AI is set to break new ground in improving the discoverability, findability and searchability of data. It fundamentally enables semantic search, rather than searches based on keywords. This means that LLMs can interpret the meaning of the text, for example by leveraging the relationships between metadata attached to time series or data sets, rather than matching keywords. The **World Bank** has been actively exploring solutions in this area, for example by developing search solutions based on World Development Indicators (Solatorio and Dupriez (2024)).[11] The **International Monetary Fund** (IMF) has also been making strides with "StatGPT", an AI tool that simplifies accessing and combining statistical data using plain language, enhancing search capabilities on the IMF's data platform. Like the World Bank's, the IMF's approach leverages statistical standards, such as the Statistical Data and Metadata eXchange (SDMX), to better refine the accuracy of the results. Other similar initiatives integrate the use of LLMs with search platforms, such as Apache Solr (SEASE (2023)). The UK **Office for National Statistics** (ONS (2023)) has developed StatsChat, an AI-powered search tool utilising LLMs to enhance search accuracy on its website, significantly improving search efficiency and data accessibility. Finally, other organisations indicated in the CES survey that they are exploring AI solutions to design next-generation statistical information systems, for instance, to help users interpret statistics and create visualisations based on personalised content and usage patterns.

Relatedly, **using natural language processing (NLP) techniques to enable easier data queries** is also an active area of work among national and international statistical organisations. A key goal is to simplify the ability to query data by transforming natural language into programmatic statements, for example by using SQL. This is particularly important for analysis of record-level data, which are typically complex, as they require an advanced understanding of performance considerations often scarcely mastered by business users. By assisting with the query generation process, AI allows analysts to better focus on the correct interpretation of the data, rather than constructing the data themselves. A successful example of this is the project StatBot.Swiss by the **Federal Statistical Office of Switzerland**, a bilingual text-to-SQL system (Nooralahzadeh et al (2024)). This system simplifies access to Swiss open government data by enabling users to interact through natural language queries in English and German. The workshop "SDMX+AI: Unlocking the potential of NLP to enhance data access", co-organised by the **Organisation for Economic Co-operation and Development** (OECD) and the **BIS** in March 2024 (SIS-CC (2024)), revealed a burgeoning activity to improve SQL data accessibility.

[11] Codes are available on GitHub (github.com/worldbank/llm4data).

An additional aspect is leveraging generative AI to facilitate access to statistical content. It is worth mentioning here the efforts by **United Nations Conference on Trade and Development (UNCTAD)** to develop an open source retrieval augmented generation (RAG) LLM application (UNCTAD (2024)). This tailored, in-house solution makes LLMs more relevant and useful for the organisation's specific needs, offering a cost-effective and flexible alternative. **Statistics Norway** features another key project**,** which has been developing an AI assistant using open source tools like Python, LangChain and Hugging Face models to support employees in creating and communicating statistics (Statistics Norway (2024)). This ensures flexibility, cost-effectiveness and security, with prototypes including chatbots for data retrieval and a "methodology bot".

4.5 Other applications

The use of generative AI by official statisticians goes beyond data processing, analysis, dissemination, communication and code development. It also includes a number of other areas, such as administrative tasks, which are crucial for improving organisational efficiency and productivity. According to the CES survey, more than a third of the respondents indicate leveraging generative AI for administrative tasks, such as drafting emails, preparing presentation slides or drafting minutes. Finally, and perhaps more interestingly, some organisations also expressed interest in using AI for streamlining surveys.

5. Building a unified framework for sharing experiences, best practices and knowledge on generative AI for official statistics

Given the fast pace at which generative AI is revolutionising official statistics, **developing common, structured and consistent frameworks is critical** in at least three directions.

First, **understanding the opportunities and limitations of generative AI requires a unified and international approach to exchange experiences and best practices** among statistical organisations. In this regard, one of the primary goals of the HLG-MOS project on generative AI for official statistics has been to establish a framework to facilitate knowledge-sharing at the international level. This framework aims to gather use cases from various application areas by providing a level playing field to enhance comparability (Annex 1). It includes three main sections. The first section is on the business requirements and how AI helps to address them. Specifically, it describes the issues that the AI solution aims to resolve as well as its impact on business. Here, the template requires qualitative and quantitative indicators, such as cost savings, revenue enhancement or productivity gain. This section aims to shed light on the specificities of generative AI compared with other data science projects, particularly in terms of scalability, costs and dependencies. The second section turns to the results as well as the underlying requirements, including in terms of human and IT capacity. Results are quantified by key performance indicators as well as the stage of the use case in the process (eg proof of concept, production). Turning to the requirements, the template covers questions on the availability of any training programs and the IT stack, including system architecture, infrastructure and tools. The last section sheds light on strategic considerations at the level of the organisation. It addresses critical aspects such as project management and organisational barriers, strategies for sharing techniques and prompts, and, more fundamentally, any AI-related policy to mitigate risks and establish governance of generative AI at the level of the organisation. Key points under examination include ethical considerations, control mechanisms and standards.

Beyond framing a common approach to share knowledge, **what is needed is to develop guidelines on governing generative AI tailored to official statistics from the concrete use cases** shared by statistical organisations. Substantial progress has undoubtedly been made over the past few years in designing and proposing governance frameworks for generative AI.[12] Yet there is no framework tailored specifically to organisations involved in official statistics. These organisations have several specificities, not least that they are bound to data sovereignty and, more fundamentally, responsible for creating trustworthy, accurate and timely data which ultimately serve to train AI.[13] In this sense, producing official statistics in the age of AI requires a high degree of awareness and foresight in statistical organisations on the new way statistics will impact society through machines (Bogdanova et al (2024)). At stake are issues such as preserving and fostering the relevance, traceability and trust in official statistics. To cope with these fundamental issues, the HLG-MOS project on AI for official statistics aims to build a framework to address two critical aspects. First, the goal is to identify and mitigate AI-related risks for official statistics. Drawing on concrete use cases, the key components include developing an interoperable framework for managing AI risks, with a focus on accuracy, privacy, copyright, ethics and operational risks, including reputational ones, as well as dependencies. The second aspect is to shed light on governance-related issues among statistical organisations. One goal is to gather evidence on existing structures across organisations to manage AI-related projects and strategies, their benefits and limitations. These may include top-down approaches, project boards and AI advisory boards or, conversely, bottom-up solutions with decentralised cross-organisational committees. Another goal is to look at the design, execution and management of AI-related projects within the organisation, with particular emphasis on the roles (eg data scientists, statisticians, data engineers) and cross-unit partnerships (eg statistics, communications).

Finally, on a technical level, **the generative AI project includes a focus on prompt engineering and the architectural stack**. The purpose is to investigate various prompt engineering strategies to enhance AI implementations, emphasising the role of "guardrails" to ensure AI systems operate within desired parameters. The project will establish best practices, including prompt versioning and knowledge-sharing, utilising tools such as LangSmith for versioning features. A potential deliverable can be a centralised prompt library for official statistics hosted on platforms such as GitHub. Additionally, quality assurance guidelines for prompt engineering are being developed, and the best prompts are being collected. Another technical focus is exploring a design of robust architecture for AI systems, considering infrastructure, tools and IT stack selection tailored to the needs of official statistics. Criteria for selecting appropriate generative AI models will be defined, along with methods for evaluating models and generative AI's output accuracy through

[12] A noteworthy example is the work conducted by the OECD AI Policy Observatory (oecd.ai).

[13] For example, see UN CEB (2023) making the case for international data governance.

standardised procedures. The project is exploring options of cloud versus non-cloud solutions and issues related to open source versus proprietary tools. Key considerations include scalability, performance and cost-effectiveness in the chosen architecture.

6. A roadmap for preserving trust in official statistics in the age of AI

Generative AI holds the potential to revolutionise official statistics in the forthcoming years. Undoubtedly, it will bring several benefits across the wide statistical spectrum, not least in terms of productivity and efficiency gains for statistical organisations. Yet a long term vision for generative AI in official statistics involves a fundamental shift in how statistical organisations operate and, more importantly, in the way that data, statistics and information are retrieved and communicated. Making the most of the opportunities offered by AI in official statistics may require building a proactive and adaptive approach to ensure that official statistics remain relevant and trusted along four dimensions.

First, there is a pressing need to ensure the accurate and clear origin of statistics. The paradigm of information retrieval is quickly changing as we move into the era of AI, with provision and traceability of data key to ensuring trustworthy information. Official statisticians need to weigh in even more on the quality of their statistics, producing more, finer and better metadata, particularly to explain the source, the methodology and any model used in the production. This requires also preparing and building institutional capabilities in order to prepare AI-ready data.

Second, it is critical to foster better communication of official statistics. The information revolution has made data ubiquitous. Yet the public is often insufficiently statistically literate and swayed by sensationalist media. To address misuse and misinformation – which generative AI has the potential to exacerbate – it is essential for statistical organisations to provide clear, authoritative information – through analysis, text, visuals and other content – to complement the raw data.

Third, sound governance and stewardship by statistical offices will be key. The use of generative AI, particularly LLMs, raises concerns about data privacy, security and the potential misuse of statistical information. Organisations must navigate these complexities while advocating for open data principles and ensuring responsible data stewardship. Further, the use of generative AI in official statistics necessitates adherence to ethical guidelines and evolving regulations. Statistical organisations must actively participate in policy discussions, contribute to developing ethical frameworks for AI use and ensure compliance with relevant regulations.

Ultimately, international collaboration and knowledge-sharing will be critical to preserve and promote trust. The importance of collaboration among statistical organisations is crucial in navigating the evolving landscape of generative AI. This includes establishing platforms for sharing best practices in prompt engineering, collaborating on generative AI-related research projects and engaging in open dialogue about the ethical and societal implications of AI in official statistics.

Annex

Use case template

Table 1

	Title of the use case and the name of the organisation (Please provide a title of the use case with enough information so that readers can understand the context.)
Executive summary	Overview of the business problem, solution and key outcomes.
Business need and solution	**Description:** specific business need or problem addressed by the generative AI solution. **Impact:** effect on business, including metrics and qualitative factors. **Overview of the solution:** description of the solution. **Value added by the solution**
Specificities of generative AI compared to other data science projects	**Benefits:** tangible and intangible benefits (eg including cost savings, revenue enhancement, productivity gain and other value metrics. **Challenges**, including scalability, costs, dependencies.
Preliminary results	**Metrics:** key performance indicators and results. **Feedback and analysis:** user and stakeholder feedback and the significance of the preliminary results.
Requirements and training	**Human resources:** skills and roles for implementation and maintenance. **Training:** required training programs and materials. **Design and components:** system architecture, infrastructure, tools and IT stack selection. **Cloud or non-Cloud:** what was used. **Model selection:** what model was used and why. **Strategies and costs:** strategies for scalability and performance, and cost considerations.
Validation and testing	**Methodology and testing:** methods and types of testing conducted for validation. **Results:** outcomes of the validation process.

Stakeholders/collaboration	**Internal:** teams and individuals within the organisation. **External:** partners, vendors and external entities.
Organisational barriers and solutions	**Challenges:** key challenges faced during implementation. **Solutions:** strategies to overcome barriers. **Risks:** potential risks associated with the solution, including any potential vendor dependency. **Mitigation and monitoring:** steps to mitigate risks and ongoing risk monitoring and management.
Prompt engineering	**Strategies and examples:** strategies, techniques and successful prompts. **Guardrails and best practices:** ethical considerations, control mechanisms, industry standards, guidelines, documentation and versioning, if any.
Implementation status	**Progress:** implementation status. **Challenges and next steps:** ongoing challenges and solutions, and immediate next steps.
Future opportunities and directions	**Opportunities and developments:** potential future enhancements and strategic long-term plans. **Project journey and integration:** timeline, key milestones and organisational workflow integration. **Co-development, co-investment and collaboration:** directions about collaboration plans, within and outside the organisation. **IT development**: directions about upgrade of IT architecture and plans.

References

Bogdanova, B, M Erdem, B Ligani and O Sirello (2024): "Enhancing metadata with generative AI: the case of BIS statistics", presentation at the 12th Biennial Conference of the Irving Fisher Committee on Central Bank Statistics, August.

Central Statistics Office (CSO) (2023): "Journey to using R – experience of the Central Statistics Office of Ireland", ECE/CES/2023/19, June.

Düker, M (2024): "Textbased intelligent assistant: TIA", presentation at the HLG-MOS Communication and AI task team internal meeting, 6 June.

Eusebio, P and F Givernaud (2023): "SAS to R conversion: the large language models", presentation at the UNECE HLG-MOS Modernisation Workshop 2023, November.

Faris, M (2023): "Use of AI in communication and dissemination statistics", UNECE Expert Meeting on Dissemination and Communication of Statistics, Lisbon, 11–13 October.

Garrett, A (2024): "The devil, the detail, and the data", *Journal of the Royal Statistical Society*, *Series A: Statistics in Society*, vol 187, no 4, pp 857–78.

High-Level Group for the Modernisation of Official Statistics (HLG-MOS) (2018): "The use of machine learning in official statistics", Blue-Skies Thinking Network Position Paper, November.

——— (2023): *Large language models for official statistics*, December.

——— (2024a): *Organisational aspects of implementing ML based data editing in statistical production*, February.

——— (2024b): "Use of generative AI in statistical organizations: CES survey results", July.

Irving Fisher Committee on Central Bank Statistics (IFC) (2023): "Central banks' use of time series products", *IFC Guidance Note*, no 4, December.

——— (2024a): *IFC Annual Report*, January.

——— (2024b): "Communication on central bank statistics", *IFC Bulletin*, no 60, April.

Nooralahzadeh F, Y Zhang, E Smith, S Maennel, C Matthey-Doret, R de Fondville, K Stockinger (2024): "StatBot.Swiss: bilingual open data exploration in natural language", arXiv:2406.03170, June.

Office for National Statistics (ONS) (2023): "Using large language models (LLMs) to improve website search experience with StatsChat", *Data Science Campus*, 23 August.

——— (2024): “ClassifAI – exploring the use of large language models (LLMs) to assign free text to commonly used classifications”, *Data Science Campus*, 9 July.

Organisation for Economic Co-operation and Development (OECD) (2024): *OECD Digital Economy Outlook 2024 (Volume 1) - Embracing the Technology Frontier*, May.

SEASE (2023): *PoC final report: support for natural language search through large language models*, September.

Solatorio, A and O Dupriez (2024): “Beyond keywords: AI-driven approaches to improve data discoverability”, *World Bank Blogs,* 22 May.

Statistical Information System Collaboration Community (SIS-CC) (2024): “SDMX+AI: a path to explore generative AI for better data accessibility”, *SIS-CC News,* 25 April.

Sirello, O (2024): “Editing metadata with generative AI”, *UNECE Expert Meeting on Statistical Editing*, October.

Statistics Netherlands (2023): “Free and open-source software at Statistics Netherlands”, ECE/CES/2023/20, June.

Statistics Norway (2024): “Adopting artificial intelligence in the production and dissemination of official statistics“, ECE/CES/2024/21, May.

United Nations (UN) (2014): “Fundamental principles of official statistics”, A/RES/68/261, March.

United Nations Chief Executives Board for Coordination (UN CEB) (2023): “International data governance: pathways to progress”, CEB/2023/1/Add.2, May.

United Nations Economic Commission for Europe (UNECE) (2019): *Generic Statistical Business Process Model (GSBPM)*, version 5.1, January.

——— (2021): *Strategic Communication Framework for Statistical Institutions*, July.

——— (2022): *Synthetic data for official statistics: a starter guide*

——— (2024a): “Data Governance Framework for Statistical Interoperability (DAFI)”, High-Level Group for the Modernisation of Official Statistics, March.

——— (2024b): *Cloud for Official Statistics*, March.

——— (2024c): “Generative AI already having significant impact in statistical organizations”, Press release, September.

United Nations Trade and Development (UNCTAD) (2024): “Free to play: United Nations Trade and Development’s experience with developing its own open-source retrieval augmented generation large language model application”, ECE/CES/2024/23, May.

Generative AI and official statistics:

the project of the UNECE High-Level Group for the Modernisation of Official Statistics

Disclamer: The views and opinions expressed in this presentation are those of the author and do not necessarily reflect the official policy or position of UNECE

Authors:

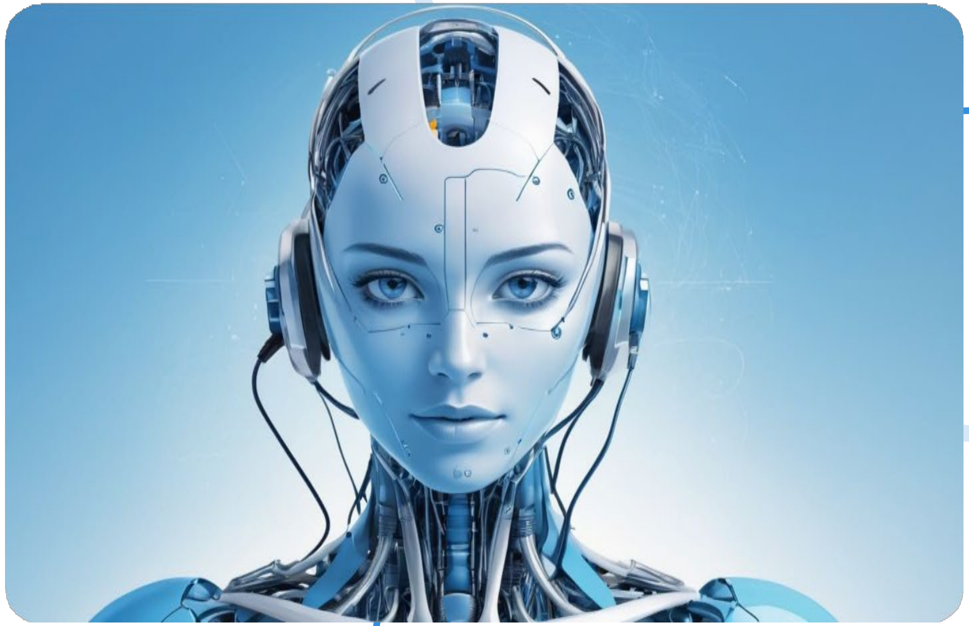
Amilina Kipkeeva, InKyung Choi, Olivier Sirello, Vytas Vaiciulis

The UNECE HLG-MOS Project on Generative AI

The HLG-MOS, established by the CES, leads the modernization of official statistics through collaboration among national and international organizations.

The project on generative AI focuses on governance, project management, technical aspects and explores prominent use cases.

Key objectives include sharing experiences, mapping AI's impact on statistics, and developing best practices for responsible AI implementation.



Collaborative Effort

The project serves as a platform for national and international organizations to share experiences, map the footprint of AI, and discuss its opportunities and limitations.

Strategic Considerations

The project investigates governance and ethical considerations, as well as practical aspects like prompt engineering, infrastructure, and tools.

CES Survey on Generative AI in Statistical Organizations

Overview of Respondents

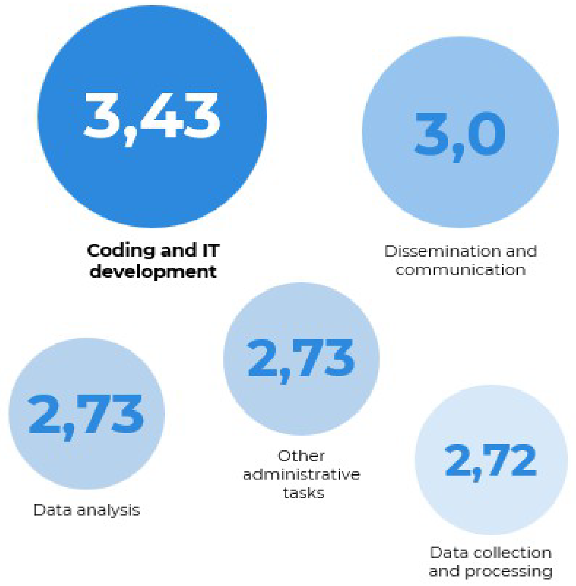
41
Number of responses

36
National organizations

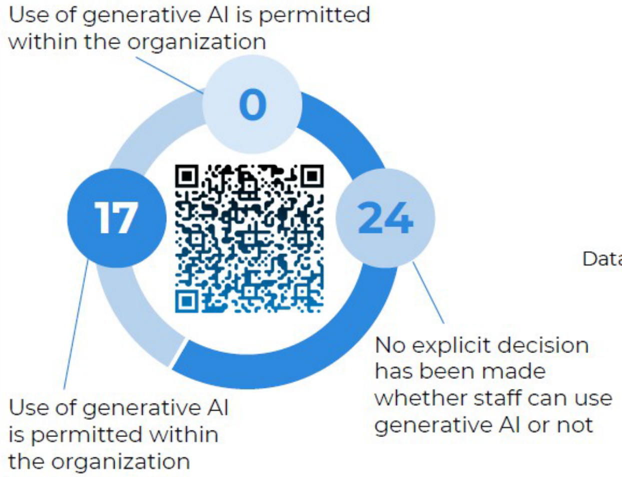
+

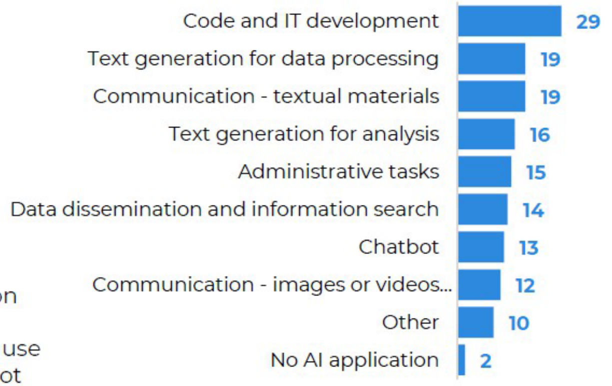
5
International organizations

Impact of Generative AI*

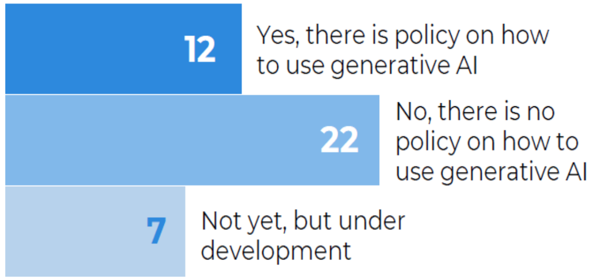


Usage and Permission*





Organizational Policies and Guidelines



Data security

Transparency

Ethical use

Challenges and Concerns*

* Weighted average score

Use Cases in Data Processing and Analysis

Metadata Editing

The Bank for International Settlements (BIS) developed AI-powered assistants for metadata editing, significantly reducing the time and effort required and enhancing consistency and accuracy.

Synthetic Data Generation

LLMs are used to generate synthetic datasets for safe data exchange and internal operations, improving data confidentiality and balancing training datasets.

Statistical Classifications

The Australian Bureau of Statistics used LLMs to update occupation codes, achieving 70% quality compared to human-created lists and saving over 1600 working hours.

Report Generation

Statistics Canada experiments with LLMs to generate draft reports from Census data, creating detailed articles that are reviewed for accuracy. The process includes generating summary statistics and bilingual report translation.

Use Cases in Code and IT Development

Code Translation and Generation

The Central Statistics Office (CSO) Ireland developed the "SAS to R Transcompiler" using GPT-4, automating the translation of SAS code to R. This tool saves significant time and reduces manual correction needs.

Coding Assistants

Generative AI assists in various coding tasks, such as code generation, debugging, and documentation, enhancing code quality and robustness.

Open-Source Transition

Many statistical offices are transitioning to open-source programming languages like R and Python, facilitated by AI-powered code translators and generators.

Use Cases in Communication

Social Media Content

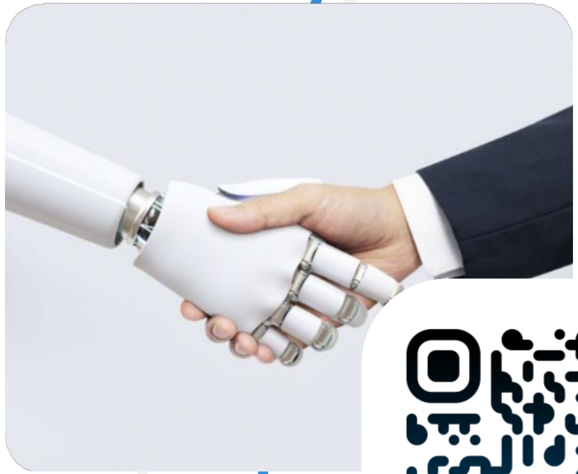
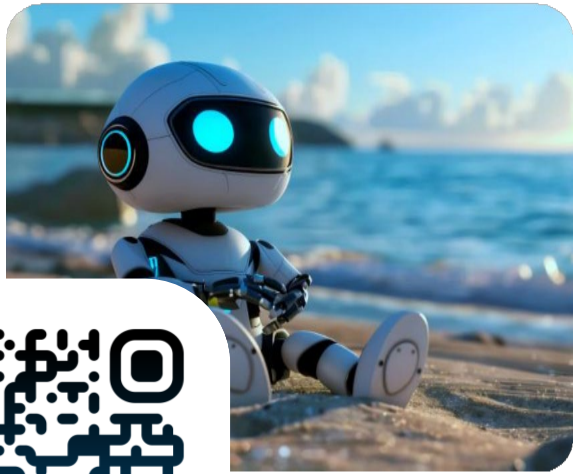
The UNECE Statistics Division uses LLMs to create coherent social media posts, ensuring consistency and brand recognition.

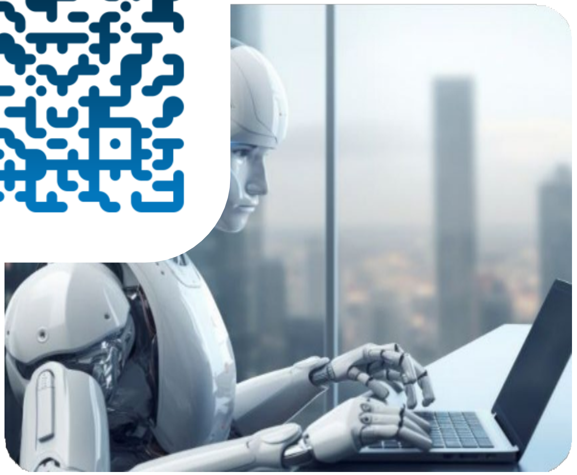
Non-Textual Content

Statistics Canada explores generating narrated content and visuals to enhance data accessibility and engagement with the public as well as experimenting with using AI for video production.

AI-Powered Chatbots

AI chatbots improve user interaction with statistical data by providing intuitive and user-friendly interfaces for answering data-related questions.



Frameworks for Generative AI in Official Statistics

Unified Framework

Establishing a unified framework for sharing experiences and best practices is crucial for the responsible adoption of generative AI in official statistics.

Governance and Ethics

Developing guidelines tailored to the needs of statistical organizations, focusing on accuracy, privacy, transparency, and trust.

Technical Considerations

Emphasizing prompt engineering strategies, quality assurance, and the selection of appropriate AI models. Criteria include scalability, performance, and cost-effectiveness.

International Collaboration

Highlighting the need for international cooperation to share knowledge, best practices, and develop common standards and guidelines for AI use in official statistics.

Thank you for attention

Amilina Kipkeeva

amilina.kipkeeva@un.org