Artificial intelligence – a buzz, must or bust for official statistics

Mariana Kotzeva,
Director-General, Eurostat

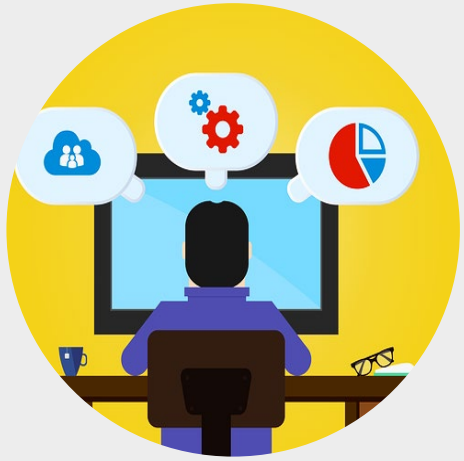
15-17 May 2024, Mexico City

IOS-ISI 2024 MEXICO CONFERENCE

Fast-changing technological landscape

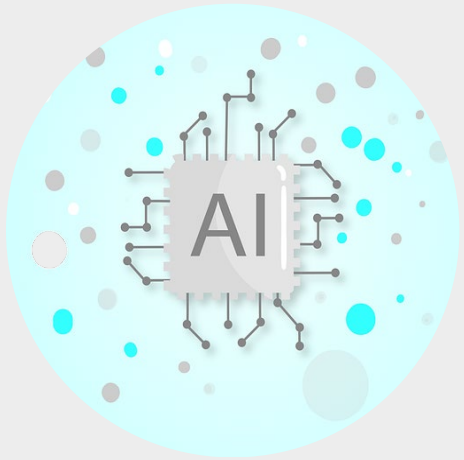
- Generative AI
- Cloud and edge computing
- Cybersecurity & Privacy enhancing technologies
- Internet of things (IoT)
- Augmented reality and digital twins

The challenge for official statistics



- To understand technology trends and follow them appropriately to stay relevant in a rapidly evolving digital landscape
- To have in mind that different technology trends have different impacts on official statistics
- To focus on those that are expected to have the greatest impact on official statistics

Artificial Intelligence System: Definition



EU Artificial Intelligence Act (2024)

AI system is a machine-based system that:

- is designed to operate with varying levels of autonomy and;
- may exhibit adaptiveness after deployment, and;
- for explicit or implicit objectives, infers from the input it receives, how to generate outputs such as predictions, content, recommendations, or decisions that can influence physical or virtual environments.

IOS-ISI 2024
MEXICO CONFERENCE

AI landscape

Generative AI

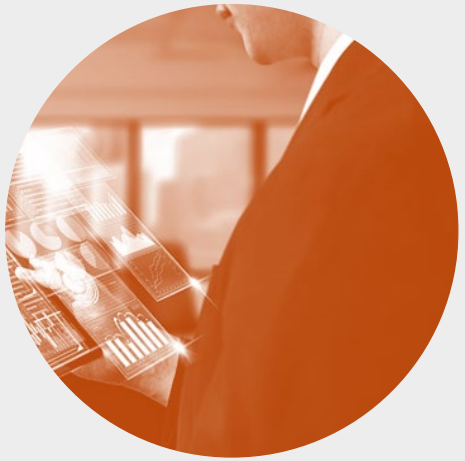
Large Language Models (LLM) create new content

Machine Learning (ML)

Algorithms and Deep learning (DL) based on neural networks to mimic brains of humans trained on available data to predict, classify and cluster them

IOS-ISI 2024 MEXICO CONFERENCE

Is AI a must for official statistics?



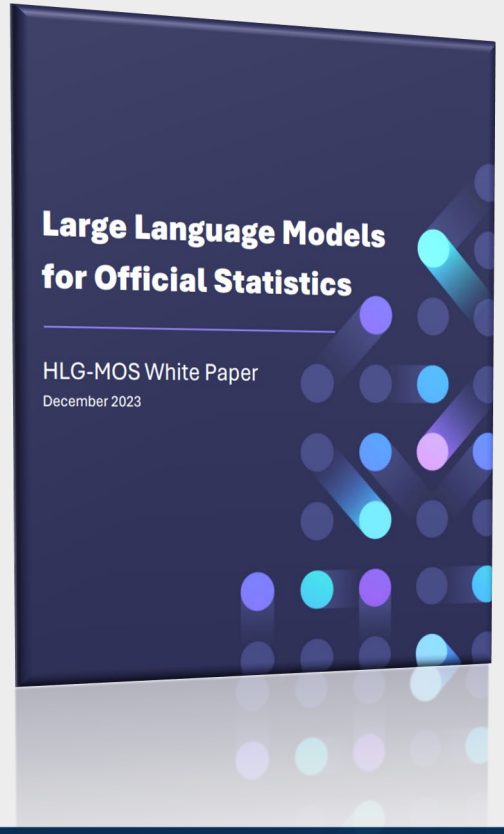
- Yes, because of the fear of missing out on the next big thing
- Yes, BUT based on deep and hard thinking about which use cases might be appropriate for implementing AI and about which type of AI might bring benefits to official statistics

AI in official statistics: possible use cases

- AI for data collection
- AI for data processing and classifications
- AI for data editing (validation and imputation)
- Generative AI for improving dissemination and interaction with users

IOS-ISI 2024 MEXICO CONFERENCE

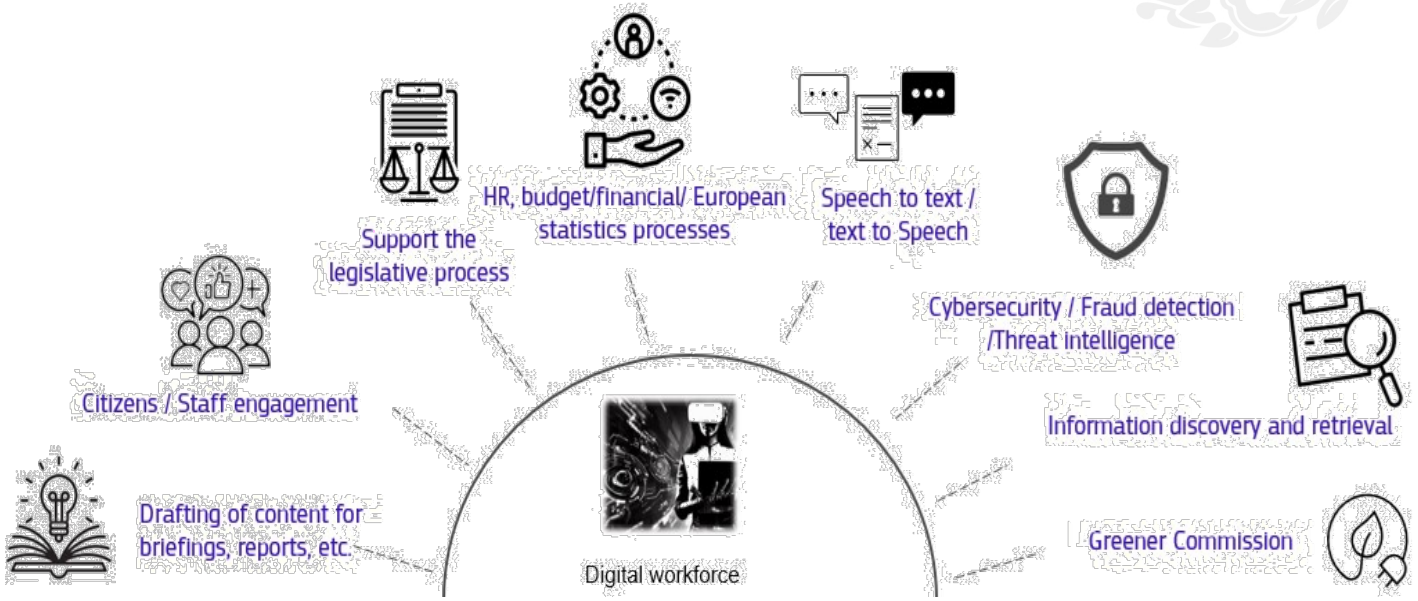
Examples of AI applications in official statistics



- Web intelligence systems for deriving top skills
- EU LUCAS survey images processing
- ML for nowcasting economic indicators
- Chatbot assistants and advisors based on text and problem description
- Web scraping to predict NACE level of enterprises

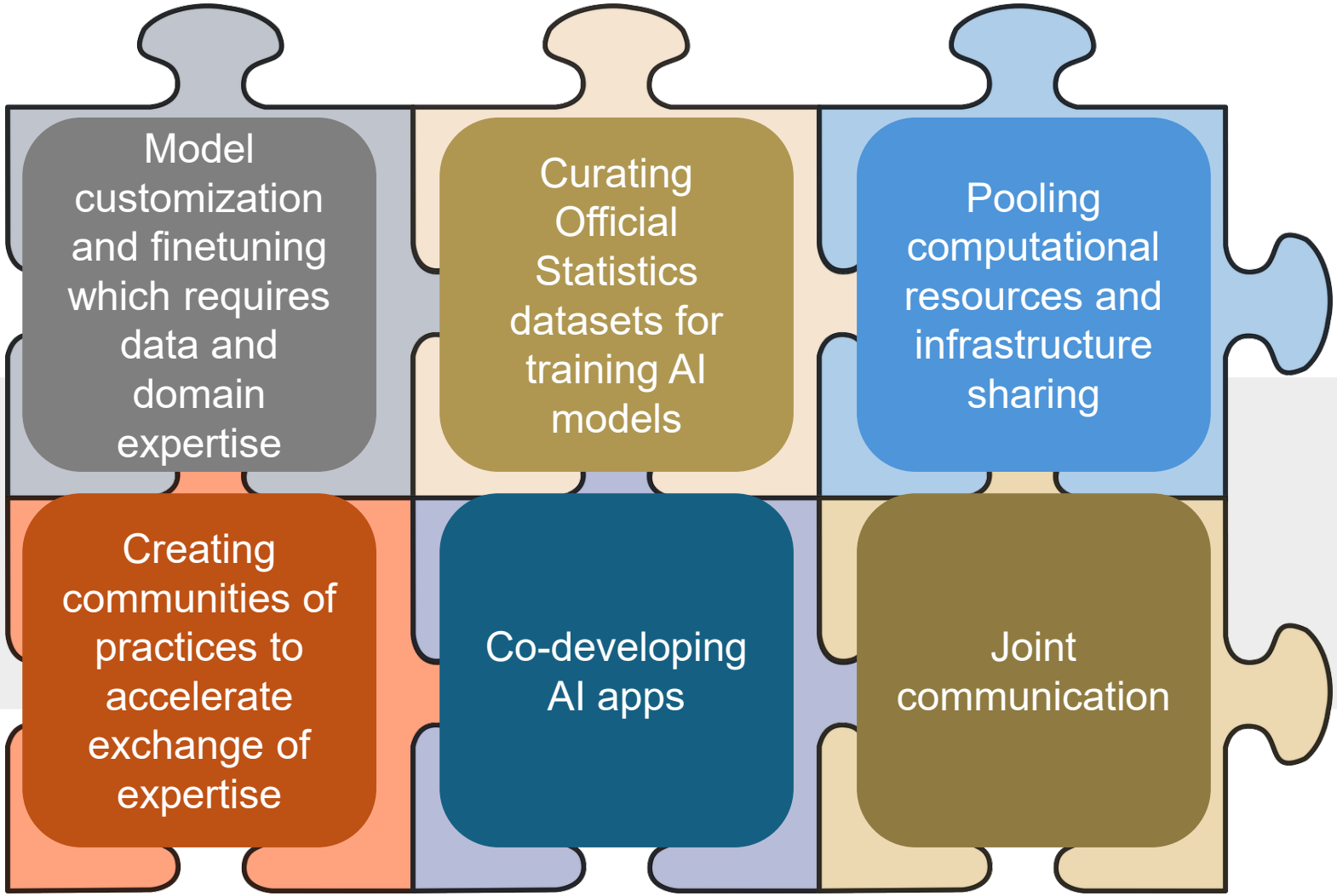
IOS-ISI 2024
MEXICO CONFERENCE

AI@EC: AI powered digital workplace



IOS-ISI 2024 MEXICO CONFERENCE

Working together for AI in official statistics



IOS-ISI 2024 MEXICO CONFERENCE

Is AI a bust for official statistics?

- Discussions should not focus on technical issues but rather on opportunities that AI might bring to official statistics

- The efforts should also focus on:
  - How to use AI in a responsible and ethical manner?
  - How to comply with the legal AI and data protection frameworks?
  - How to mitigate the risks associated with the AI use?

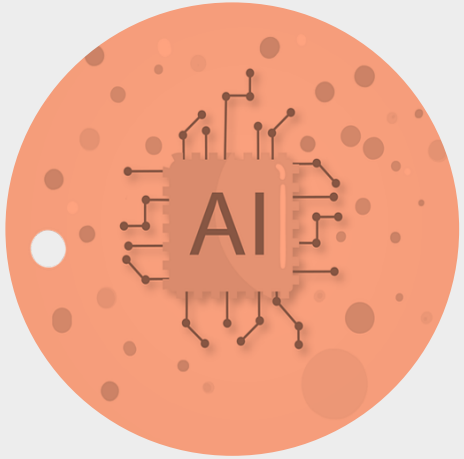
Ethical challenges in using AI in official statistics

Do we need UN principles of responsible AI for official statistics?

- For the public good while balancing the public good with the issue of privacy intrusion
- Safe and secure protection privacy and confidentiality
- Fair and unbiased AI ensuring accuracy and output quality
- Transparent and accountable regarding data and models used

IOS-ISI 2024 MEXICO CONFERENCE

Ethical AI guidelines inside statistical organisations



In addition to general principles of responsible AI for official statistics, more specific guidelines are needed

- Guidance and tools at product level (e.g. on which metrics to use to assess whether an AI model delivers an accurate or explainable product)
- Break down big ethical concepts like privacy and explainability into processes, infrastructure and actions that realise these values
- Allocate responsibilities and ownership for each AI system and function (i.e. pairing humans with AI)

AI might bring benefits but might also scale up risks

- Reputational risks related to ethical challenges
- Regulatory and legal risks related to non-compliance with AI and data protection legal acts
- Risks of wasting resources, inefficiency of AI products deployment and inability to use data to train AI models appropriately

There is a need for an AI risk management strategy outlining the responsible use of data and the development of AI products while minimizing potential risks

ISI 2024 MEXICO CONFERENCE

Compliance with legal AI and data protection frameworks

EU Artificial Intelligence Act (2024)

- ✓ the first comprehensive horizontal legal framework for AI in the world
- ✓ It provides EU-wide rules on data quality, transparency, human oversight and accountability regarding AI systems offered on the market or shared free of charge
- ✓ AI systems should follow certain rules based on risk assessment of their impact for society

IOS-ISI 2024
MEXICO CONFERENCE

AI impact on the environment under which official statistics function

- Generative AI changes the way the users access information - from searching for data and information to asking specific questions
- The growing public concern about the potential unreliability of AI-generated content will create a risk of increasing the overall mistrust to any information provided on-line
- Generative AI systems will provide new types of information products and will open the door for new players in an already crowded information market that will compete with official statistics for getting users attention

IOS-ISI 2024 MEXICO CONFERENCE

Conclusions

- AI and in particular generative AI are evolving at a rapid pace, receiving a lot of buzz recently
- The ever-expanding traditional and generative AI promise a bright future to those who can master this technology according to their business needs and goals
- Success requires investment in understanding both the benefits that AI might bring and risks it might entail
- Joining efforts in implementing AI responsibly and communicating it appropriately is paramount for the statistical community

IOS-ISI 2024
MEXICO CONFERENCE

Thank you for your attention

INEGI
International Statistical Institute
IAOS
IMPROVING OFFICIAL STATISTICS