



One-stop-shop for AI-ML for Official Statistics

Francesca Kay, CIO, Central Statistics Office, Ireland

November 2024

Background....

- Proposal to form One-stop-shop (1SS) for AI-ML for Official Statistics
- Funded by EU action grant
- Project kick off in April



Key Objectives

Develop, maintain and evolve coherent set of capabilities

Set up sandpit platform/hub

Provide support & guidance for integration & maintenance of AI/ML solutions through training and active support

Build communities around OSS solutions developed by NSI's Share ideas, experiences, lessons and success/failure stories to stimulate AI/ML innovation

Enable transition from experimental solutions to industrialised and production solutions

PROJECT OUTCOMES



Central Statistics Office

Build a framework for developing AI/ML solutions to be used in the context of official and European statistics

ESS staff and partners have access to established and proven AI/ML solutions/resources to be leveraged in the context of official statistics production

ESS organisations are encouraged to innovate and eventually realise benefits from the use of AI/ML approaches for statistical products

Economies of scales and resources are achieved through cooperation inside and outside the ESS and transition from ideas to production is accelerated



Who's Involved?















Statistisk sentralbyrå Statistics Norway



INSTITUTO NACIONAL DE ESTATÍSTICA Statistics Portugal





REPUBLIC OF SLOVENIA STATISTICAL OFFICE RS





Project



Overarching Work Packages

WP1 Coordination – CSO

WP2 Comms & Engagement – ISTAT

WP3 Platform – INSEE

WP4 AI/ML Ecosystem - DeStatis

WP5 Standards – DeStatis & CBS

WP6 Knowledge Sharing & Training – Statistics Poland



Use Cases

AI/ML on earth observation data, satellite imagery

Editing focus - Statistically valid and efficient editing in official statistics by AI/ML

Imputation focus - Statistically valid and efficient imputation in official statistics by AI/ML

From text to code - Experiences and potential of the use of AI/ML for classifying and coding

pplying ML for estimating firm-level supply chain networks

Use of generative large language models in statistics

Generation of synthetic data in official statistics: techniques and applications



Methodology in ML/AI

- Overcoming "adhocracy"
- Reproducibility
- Transparency and explainability
- Ethical considerations



Methodology in AI/ML

Methodological and implementation guidelines for applying AI/ML in official statistics

Generalised knowledge, identified norms and best practices for developing and operating AI/ML-based solutions in official statistics

Standards in applying AI/ML in official statistics to facilitate the transition from development and experimentation of AI/ML-based solutions to industrialisation and production



Methodology in AI/ML

Standards

- Clarity
- Interpretability
- Transparency
- Data Quality
- Auditability
- Ethics

Methodology

- Unbiased
- Reproducible
- Valid statistical outcomes

Implementation

- MLOps
- AI/ML Labs
- Experimentation to Production

%

Current Status

- Each WP has commenced work
- Project dashboard launched on CROS
- Newsletter launched





Early Work.....

- Research started across the work packages, for example:
 - Comparison of Earth Observation models across countries to identify which ones can be "generalised"
 - Application of Belgian software solution & Oxford University model on estimating firm-level supply chain networks to Portuguese & Dutch datasets
 - Options on privacy risk measurement for synthetic data
 - Text to code considering use cases e.g. NACE revision, hierarchical methods for text classifications, synthetic training datasets in multiple languages
 - Using AI/ML for error detection & editing in large datasets



Early Work.....

- Most work packages are still agreeing the specific use case & associated methodology to work on, but
- Testing environment (sandbox) is now available



Key Events

NTTS 2025 Quality Conference 2026 (tbc) NTTS 2027





If you want to learn more and/or receive our newsletter please contact:

1ssai-mlcsoprojectteam@cso.ie