



Artificial Intelligence & Machine Learning for Official Statistics

AIML4OS ESSnet Project

Data Science Leaders Network – January 2025

Brendan O'Dowd PhD Senior Statistician, Data Science

Outline

- Overview of project
- Governance initiatives
- Progress from two use cases





Project Overview

- 14 participating countries
- €4 million in total funding
- 4 years (Apr 2024 Mar 2028)
- 13 workpackages CSO in charge of WP1, Project Management and Coordination
- 55 separate tasks, 39 deliverables





Project Overview

Provide a comprehensive suite of resources, guidelines, methodologies and use cases for the implementation of AI/ML in Official Statistics.

Build communities around AI/ML and innovation	Support transition from prototype to production	Provide working use cases
Provide sandbox for development	Promote shared standards and guidelines	Deliver economies of scale



Project Structure

#	Supporting Workpackages	Lead
1	Coordination	
2	Communication	
3	Lab	
4	State-of-play	-
5	Standards, methodology, implementation	
6	Knowledge repository and training	

#	Use Case Workpackages	Lead
7	Earth Observation Data	
8	Editing	
9	Imputation	
10	Text to code	
11	Supply Chain Networks	
12	Large Language Models	
13	Synthetic Data	

As many as 12 countries on each workpackage

Governance Initiatives

- WP Lead meetings every six weeks, with progress reports for each WP
- Quarterly expenditure tracking
- Review panel
- Shiny App for tracking project tasks/deliverables



WP7 – Earth Observation Data Model One – Crop Mapping

- Start with open-source three crop model, choose regions for each country within the subgroup.
- Create inventory of information and software required (try Copernicus Data Space Ecosystem)
- Q1 2025: Create detailed process map
- Q2 2025: Begin model on one region.



WP7 – Earth Observation Data Model Two – Land Cover Model

- Start with CoSIA model from IGN. Arrange workshop to gain detailed understanding of the model.
- Choose regions to test. Examine data availability and harmonization.
- Q4 2025: Apply CosIA to selected regions.
- Report on representativity and comparability.



WP10 – Text Classification

Clusters have been formed around five common challenges:

- 1. Insufficient training data / Data gaps
- 2. Addressing the complexity and nuances of natural language
- 3. Traditional models neglect hierarchical structures
- 4. Challenges in deploying/maintaining classification models in production environments
- 5. Adapting to classification system revisions



Thank you!

- www.linkedin.com/company/essnet-aiml4os/
- cros.ec.europa.eu/dashboard/aiml4os



Funded by the European Union



