

Generative AI and Large Language Models in Official Statistics: From Exploration to Production

Jakob Engdahl, Statistics Sweden, jakob.engdahl@scb.se

Abstract

Many statistical organisations experiment with machine learning and generative AI, but only 10–20 percent of models reach robust production systems. This paper examines requirements for moving from model-centric experimentation to system-centric, production-grade AI in official statistics. The first part outlines main challenges: organisational maturity obstacles, expanded application space from generative AI, trust and confidentiality concerns, technical integration issues, limited transparency of external models, and reproducibility expectations. A typology of six AI use cases is proposed. The second part presents nine design elements for AI systems, including ML-Ops, evaluation frameworks, human oversight patterns, and agentic architectures. The final part connects these through an illustrative matrix supporting systematic design and governance decisions.

Keywords: Generative AI, Large Language Models, Official Statistics, ML-Ops, Maturity, AI Systems