LLM Inflation Forecasting Pilot — the Czech National Bank's Public Test of AI Against Its Core Model

AI in the Public Sector · Good practice · Czechia

Quality score: 60/100

Evidence strength: Quasi-experimental

Summary

In a public blog experiment, the Czech National Bank pitted general-purpose LLMs (OpenAI o1, Grok 2) against its own core inflation model and market analysts: o1 scored the lowest overall forecast error (RMSE 5.28 vs the bank's 5.48), but the bank's model won decisively in 2023–2

Key dimensions

Dimension	Score
Evidence of impact / measured public value	2/3
Transparency, fairness & accountability	3/3
Transferability / demonstrated replication	2/3
Scalability beyond pilot	1/3
Governance, capability & sustainability	1/3

Evaluated by C-NAPSE

Data sources

C-NAPSE web research — accessed 2026-07-19

Czech National Bank (■eská národní banka, ■NB) — accessed 2026-07-19

■NB blog: First use of AI in inflation forecasting at the ■NB — accessed 2026-07-19

■NB Working Paper 09/2026 (Quantile Regression Forest) — accessed 2026-07-19

Central Banking: CNB study highlights AI's limits as inflation forecasting tool — accessed 2026-07-19

Cite this — Evidoria. *LLM Inflation Forecasting Pilot — the Czech National Bank's Public Test of AI Against Its Core Model*. Persistent ID: 2cadf280-164c-4cab-b89f-3221c15d2431.