

FUSE: Forecasting with UNSPSC-Enhanced granularity

Erli Wang^a, Dakshi Kapugama Geeganage^a, Opec Kemp^a

^a *Queensland Government Procurement, Department of Housing and Public Works, Brisbane, Australia*
Email: erli.wang.Dakshi.KapugamaGeeganage.Opec.KEMP@hpw.qld.gov.au

Abstract: Public procurement represents a major share of national economies and a lever for development and policy. Yet strategic use of procurement data in Queensland Public Sector is hindered by broad General Ledger (GL) codes, ill-suited for analytics. This paper introduces FUSE (Forecasting with UNSPSC-Enhanced granularity), a two-stage AI framework for high-resolution demand forecasting. FUSE first maps transactions into the United Nations Standard Products and Services Code (UNSPSC) taxonomy with 92% accuracy, then applies time-series forecasting on these granular streams. Results show forecasts capture the patterns and forecast error (MAPE) for main components is within a reasonable level. By exposing demand patterns hidden by aggregation, FUSE offers a stronger empirical basis for policy, sourcing, and supply chain management.

Keywords: Government expenditure, Artificial intelligence, Multiclass, Forecasting, LLM, UNSPSC

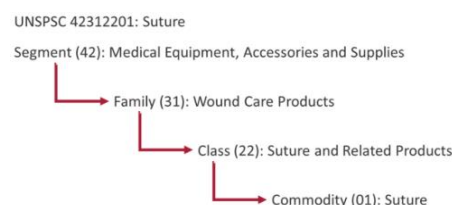
1. INTRODUCTION

Public procurement is a major economic driver, with OECD [1] countries spending around 13% of GDP on it. In Queensland this exceeds \$27 billion annually, underscoring the need for optimization. Governments face rising pressure to deliver more with less, while citizens demand value and transparency. Traditional models are poorly suited to this environment.

A key obstacle is **data granularity**. Procurement transactions are coded to general ledger (GL) accounts, which are built for compliance rather than analysis. Categories like *Office Supplies* or *Medical Supplies* conflate highly diverse items; within “Medical Supplies” one might find implants, lab consumables, and PPE. Such aggregation masks divergent trends, producing forecasts that appear stable but misrepresent real demand.

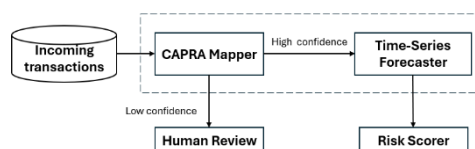
The **UNSPSC taxonomy** [2] offers a four-tier structure (segment, family, class, commodity) for fine-grained categorisation. Forecasting at this level mitigates aggregation bias, exposes heterogeneous trajectories, and yields more policy-relevant insights.

The challenge, however, is monumental: how to automatically and cost-effectively align multi-million transaction records with UNSPSC? This paper presents an AI-driven approach.



2. THE FUSE FRAMEWORK

We present FUSE (Forecasting with UNSPSC-Enhanced granularity), a framework designed to enable robust procurement analytics by moving from high-level GL data to granular, forecast-ready UNSPSC time-series data.



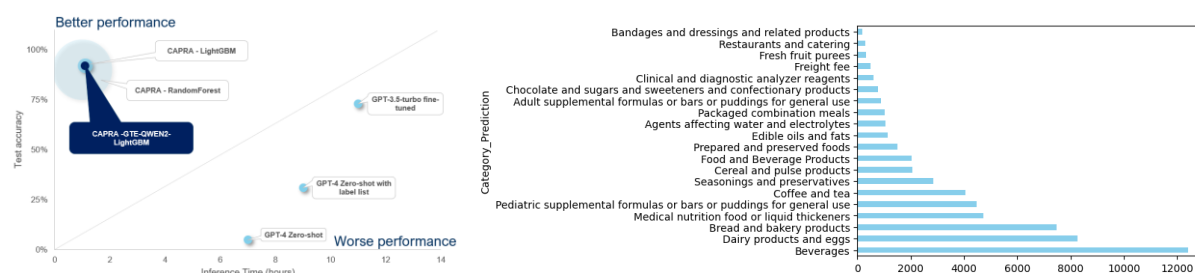
FUSE operates in two primary stages,

- **Stage 1: AI-Powered UNSPSC Categorization.** Raw transaction records, containing unstructured text, GL codes, and other metadata, are fed into our categorization architecture with predictive reasoning and alignment (CAPRA) for UNSPSC pipeline. Then, we introduce and evaluate an AI-feedback-driven validation phase to flag only the lowest-confidence predictions for human review.
- **Stage 2: Granular Time-Series Forecasting.** Once transactions are mapped to UNSPSC commodity codes, they are disaggregated into distinct time series. Standard time-series forecasting models (e.g., ETS) can then be applied to each series, generating forecasting at a level of detail previously unattainable. These forecasts can feed into downstream decision-making modules, such as risk scorers that assess the probability of stockouts or wastage.

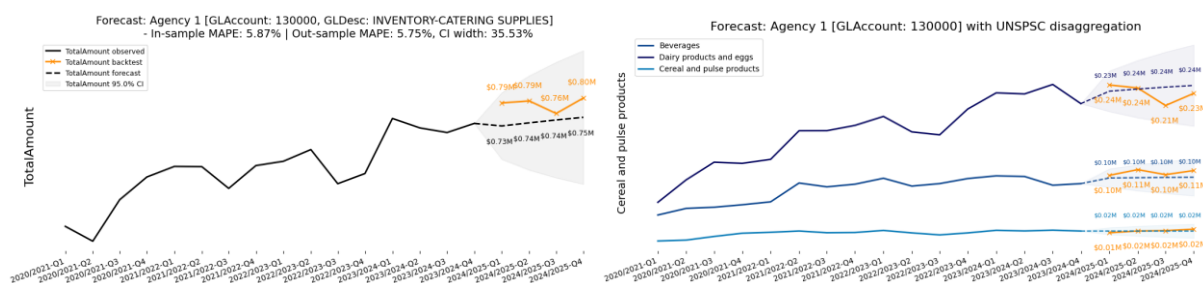
3. EVALUATION AND RESULTS

We test dataset compiled from multiple procurement sources and refined through expert validation and stratified sampling, comprised 258,278 labelled records. We applied an 80:20 stratified split on the final UNSPSC titles to preserve class distribution. This ensured representative coverage across procurement categories. The enabling step of FUSE is the CAPRA classification pipeline. It achieves 92% test accuracy, which is 19 percentage higher than a fine-tuned GPT-3.5-turbo (73%) and over 61 percentage higher than GPT-4 zero-shot reasoning (5–31%). CAPRA delivers results with only 10% of the LLM inference time, enabling seamless production deployment.

For illustrative purposes, consider GLAccount 130000 (Inventory-Catering Supplies) from Agency 1. Once we complete the UNSPSC categorization, we can disaggregate to the operational level within UNSPSC. The right figure shows the frequency of top 20 standard product and service within GLAccount 130000.



The value of UNSPSC-enhanced granularity becomes evident in the forecasting stage. In this example, forecasts generated at the **General Ledger (GL) level** present a seemingly stable view of demand. The aggregate model achieved an in-sample MAPE of 5.87% and an out-of-sample MAPE of 5.75%, suggesting high predictive accuracy. However, the associated confidence interval was wide (35.53%), reflecting considerable uncertainty. This pattern indicates that while GL-level forecasts appear reliable on the surface, they may conceal divergent sub-trends that cancel one another out at higher levels of aggregation. When the same GL account was **disaggregated into UNSPSC categories**, distinct dynamics emerged. *Cereal and pulse products* exhibited a pronounced upward trajectory although in a small amount, whereas *Beverages* remained comparatively flat and *Dairy products and eggs* showed moderate growth. These heterogeneous patterns were obscured in the aggregate forecast but became visible once transactions were classified into the UNSPSC taxonomy.



4. POLICY IMPLICATIONS AND FUTURE WORK

This study introduces FUSE, a framework for forecasting procurement demand at UNSPSC-enhanced granularity. Granular forecasting provides government with a new strategic tool. With forward-looking visibility into demand for specific goods and services, policymakers can:

- **Test market capability** to assess whether suppliers can meet forecasted demand.
- **Drive economic development** by identifying gaps and supporting local industries.
- **Support supplier diversity** by proactively creating opportunities for Small and Medium Enterprise businesses, Indigenous-owned, and Social Enterprises.

By integrating automated UNSPSC classification with forecasting, FUSE reduces aggregation bias and yields policy-relevant insights, advancing procurement analytics from compliance to strategic decision-making.

REFERENCES

1. OECD (2024), “Governing with Artificial Intelligence: Are governments ready?”, OECD Artificial Intelligence Papers, No. 20, OECD Publishing, Paris, <https://doi.org/10.1787/26324bc2-en>.
2. UNDP: United Nations Standard Products and Services Code (UNSPSC), <https://www.undp.org/unspsc>