Mapping Regional Labour Market Skill Demands Through Automated LLM-Based Analysis: Implications for EU Cohesion Policy

Ioana Ioanas, Lucian Vilcea, Andra Sandu, Adriana Cosac, Liviu-Adrian Cotfas

Bucharest University of Economic Studies

Cohesion policy, as a cornerstone of the European Union's strategy for reducing regional disparities, increasingly relies on evidence-based diagnostics of labour market competencies to guide investments in human capital development and smart specialisation. Yet, generating granular, comparable, and timely data on skill demands across regions remains a persistent methodological challenge, particularly when traditional instruments such as surveys and official statistical classifications struggle to capture the rapid evolution of occupational requirements.

This paper addresses this gap by presenting an automated, scalable workflow for extracting and classifying labour market intelligence from large volumes of job posting data, using locally deployed open-weight large language models. Drawing on a corpus of over 56,000 job advertisements collected from a major Romanian recruitment platform over a period of nearly nine months, the pipeline combines web data extraction with a multi-stage semantic enrichment process. This includes occupational classification aligned with the ESCO taxonomy — the EU's standardised framework for Skills, Competences, Qualifications and Occupations — as well as automated skill extraction and mapping.

The resulting dataset covers 2,152 validated IT job postings, classified across 75 ESCO occupations and enriched with 2,797 unique skills mapped to 881 ESCO skill concepts. By anchoring the output to ESCO, the methodology directly enables cross-regional and cross-national comparisons, a key requirement for cohesion policy monitoring and evaluation. The approach runs entirely on consumer-grade hardware without reliance on commercial cloud APIs, making it replicable and cost-effective for regional development agencies, managing authorities, and research institutions in less developed regions operating under tighter resource constraints.

From a cohesion policy perspective, the methodology offers a practical instrument for tracking skill mismatches, supporting the design of European Social Fund Plus interventions, and aligning regional educational and vocational training strategies with actual employer demand. The paper discusses how continuous application of this pipeline can serve as a monitoring tool for smart specialisation strategies and contribute

to a more dynamic, data-driven approach to reducing territorial inequalities in the digital economy.