INAPP
ISTITUTO NAZIONALE PER L'ANALISI DELLE POLITICHE PUBBLICHE

ALESSIA ROMITO & BORIS SOFRONIC

ADDRESSING THE SKILL MISMATCH USING GENERATIVE AI

Building a GenAI Model for Education, Training and Labour Market in Italy

Work 2025 - Work in the era of unruly AI

Turku, 20-22 August 2025

CONTENT

SKILLS MISMATCH OVERVIEW

Skills Mismatch: A Global Challenge with Local Impact

Global Impact:
2 out of 5 workers affected

1.3 billion worldwide

European Context: **around 30% of European employees are mismatched**

Italian Specificity: approximately 1.6 million job vacancies were difficult to fill, due to a shortage of suitable candidates or adequate skills

RESEARCH DESIGN

Can Generative AI be trusted for rigorous scientific research in workforce development?

1 **Phase 1 (2023):**
ChatGPT-3.5 testing and limitation identification

2 **Phase 2 (2025):**
Frontier Models evaluation with advanced methodologies

Methodology:

✓ **Rigorous fact-checking**
Multiple validation layers

✓ **Accuracy assessment**
Systematic evaluation

✓ **Institutional data validation**
ISTAT, INAPP, ESCO

Longitudinal Study Design
Same research questions across different technological generations

PHASE 1 (2023)
THE REALITY CHECK

- Domain Knowledge Gaps
- Knowledge Base Obsolescence
- Artificial Hallucinations

PHASE 1: THE REALITY CHECK

Domain Knowledge Gaps:

ChatGPT misidentified Italy's CP2011 as European economic classification.
Actually it's the Italian professional classification system.
Could invalidate any skills analysis.

Knowledge Base Obsolescence:

September 2021 training cutoff created dangerous temporal gaps.
Recommended closed training centers in Rome with detailed addresses.
Provided course descriptions for non-existent services.

Artificial Hallucinations:

Generated fabricated academic references with plausible titles and authors.
Created fake DOI numbers that led nowhere.

Severe risks for evidence-based policy making.

Critical Failure Modes Identified
Scientific rigor demanded systematic evaluation of GenAI limitations

PHASE 2 (2025) BREAKTHROUGH RESULTS

- The Improvement Journey
- Model Selection & Training
- Practical Applications & Validation
- Mismatch Analysis

PHASE 2: BREAKTHROUGH RESULTS

The Improvement Journey

Remarkable progress from 2023 to 2025 through targeted technological advances.

Chain-of-Thought Reasoning

Step-by-step logical processing for complex classifications.
Structured reasoning paths reduce domain knowledge gaps.
Enhanced navigation of occupational systems.

RLHF

Systematic feedback from INAPP domain experts.
Continuous learning loops for Italian labor market.
Understanding nuances of terminology.

RAG

Real-time data from CEDEFOP, INAPP, ISTAT, MUR.
Current information instead of static training data.
Solved knowledge obsolescence problem.

Benchmark Improvements

15-30% reduction in hallucinations (TruthfulQA).
Substantial progress toward reliable AI intelligence.
Quantitative validation of our progress.

80%

Progress toward fully reliable AI system for skills mismatch analysis

Systematic approach demonstrates targeted improvements deliver measurable results in specialized domains

PHASE 2: BREAKTHROUGH RESULTS

Model Selection & Training

Model Evaluation:
Benchmark scores, training recency
advanced features, cost-effectiveness

Selection:
Claude 3.7 Sonnet Pro for
superior domain performance

Custom Project Architecture:
Specialized instructions, institutional data integration

Training Sources:
CEDEFOP, INAPP datasets
ISTAT classifications, MUR databases

Methodology:
Expert supervision with
hybrid human-AI validation

Comprehensive approach: evaluation, selection, architecture, training, and validation

PHASE 2: BREAKTHROUGH RESULTS

Practical Applications & Validation

Fact-checking Protocol:
Cross-validation with:
- INDIRE (ITS Academy monitoring)
- Unioncamere-Excelsior (employer demand surveys)

Accuracy Assessment:
Systematic comparison of model outputs with official statistics

Bias Identification:
Detection of model preferences for non-representative data sources

Quality Assurance:
Multi-layer validation with domain experts

Comprehensive Validation Framework

Ensuring reliability and accuracy through systematic multi-source verification

PHASE 2: BREAKTHROUGH RESULTS

Mismatch Analysis

Three-case study comparison table with GPT match scores

Gap Types Identified:

1

Instrumental Gap

Data Analyst case

GPT match: 0.5

Solid theoretical foundation, missing specific tools (SQL, Power BI)

Solution:

Targeted 40h SQL + 60h Power BI

2

Technological Gap

Software Developer case

GPT match: 0.6

Strong algorithmic base, outdated technology stack

Solution:

Bootcamp + industry certifications + practical experience

3

Paradigmatic Gap

Digital Marketing case

GPT match: 0.3

Fundamental approach transformation required

Solution:

Comprehensive retraining with industry-standard certifications

Gap Analysis Results

Lower GPT scores indicate greater skills mismatch requiring more intensive interventions

CONCLUSIONS AND FINAL CONSIDERATIONS

- Future Perspectives
- Challenges and Ethical Considerations

FUTURE PERSPECTIVES

For VET Systems:
Differentiated intervention design based on gap typology

For Policy Makers:
Data-driven decision support with real-time labor market intelligence

For Individuals:
Personalized career guidance with precise competency mapping

For Society:
Reduced skills mismatch and improved labor market efficiency

CHALLENGES AND ETHICAL CONSIDERATIONS

Data Bias and Representativeness:
Risk of perpetuating existing labor market inequalities

Algorithmic Transparency:
Need for explainable AI in public policy applications

Privacy Protection:
GDPR compliance in processing CV and career data

User Training Requirements:
Essential competencies for effective human-AI collaboration

Responsible AI Implementation: Balancing Innovation with Ethical Safeguards

THANK YOU FOR YOUR ATTENTION

INAPP
ISTITUTO NAZIONALE PER L'ANALISI
DELLE POLITICHE PUBBLICHE

www.inapp.gov.it