IN PRESS

CARMA 2026

8th Int. Conf. on Advanced Research Methods and Analytics
1-3 July 2026, Valencia, Spain

Human-in-the-Loop Validation for LLM-Based Policy Audits: Measuring Language-Access Discourse at Scale

Shaheera Sayed
School of Public and International Affairs, Virginia Tech, United States

How to cite: Sayed, S. (2026). Human-in-the-Loop Validation for LLM-Based Policy Audits: Measuring Language-Access Discourse at Scale. In: 8th International Conference on Higher Education Advances (HEAd'26). Valencia, 15-18 June 2026. https://doi.org/10.4995/HEAd26.2026.21870

Abstract:
As large language models (LLMs) are increasingly used in policy research, questions of measurement validity, reliability, and interpretive oversight become central. This paper develops a human-in-the-loop architecture for auditing language-access discourse across large administrative corpora. We implement a four-pipeline LLM framework on 2,847 housing and community development documents (2014–2025) from Washington, D.C., Fairfax County (VA), and Montgomery County (MD), of which 744 advance to multi-pipeline analysis after language-relevance filtering. The framework generates structured outputs on language mentions, conditionality cues, and discursive framing. Validation combines multi-run stability testing, cross-pipeline triangulation, and a manual audit on a 150-document sample (87 percent observed agreement; estimated $\kappa \approx 0.74$). The system produces three scalable indicators of language-access specificity and conditionality. The primary contribution is methodological: rendering LLM-based policy audits transparent, reproducible, and evaluable through structured prompts, validation protocols, and human interpretive control.

Keywords: *large language models; policy text analysis; validation; human-in-the-loop; language access; administrative documents*

Article in Press

This work is licensed under a Creative Commons License CC BY-NC-SA 4.0
Editorial Universitat Politècnica de València