

International Journal of Population Data Science

Journal Website: www.ijpds.org



Simulating Policy Discussions with Digital Footprints and Large Language Models

Weiyao Meng¹, John Harvey¹, Christopher Carter¹, Georgiana Nica-Avram¹, James Goulding¹, Paul Frobisher², and Mina Forrest²

¹University of Nottingham, United Kingdom

²Strategic Innovation Ltd, United Kingdom

This research was supported by the Innovate UK Knowledge Transfer Partnership grant (project number KTP13737).

Introduction & Background

Parliaments play a significant role in democratic decision-making, drawing on evidence from a wide range of stakeholders. This evidence data, published on official government websites, is a valuable archive of digital traces reflecting the policymaking process. Among the key data sources is Hansard, the UK Parliament's official verbatim record which captures the full transcripts of parliamentary debates and discussions, revealing diverse perspectives, stances and interactions shaped by participants' political and social contexts. While datasets like Hansard offer rich insights into the policymaking process, their scale and varied dialogic structure present challenges for traditional analysis methods.

Recent studies indicate large language models (LLMs) can exhibit intelligent and collaborative decision-making behaviours in social simulations. This makes them a useful tool for simulating complex discussions and analysing group dynamics.

Objectives & Approach

This study explores the potential of LLMs to simulate UK parliamentary debates, with a focus on speaker roles and stance-taking. The immediate objective is to assess whether LLMs can mimic the structure and dynamics of real debates. Longer-term, this work aims to lay the foundation for a more comprehensive deliberation sandbox which provides an experimental environment that supports multi-stakeholder negotiation and collaborative problem-solving in a low-risk setting. Such a tool could enhance transparency and support more inclusive, evidence-informed policymaking.

We evaluate model performance by comparing LLM-generated debates with real parliamentary discussions. The study tests general-purpose LLMs (including Gemini 2.5 and ChatGPT-o3) and GovernmentGPT (an open-source LLM fine-tuned on Hansard). All models were prompted on the same topic (i.e., ultra-processed foods) to enable comparison. Each generated output includes a speaker's party affiliation and speech content.

Relevance to Digital Footprints

The Hansard data in this study qualify as digital footprints, capturing the digital traces of stakeholder engagement in shaping public policy. By analysing and simulating these footprints, this study shows how such data can be used to model complex social interactions and support public understanding of policy discussion.

Results

Two key patterns emerge from analysis. First, the simulated debates approximate the party composition of real parliamentary debates, i.e., the proportion of utterances by party broadly aligns with actual distribution. Second, many statements exhibit neutral or unclear stances, highlighting challenges in generating clear argumentative positions.

Conclusions & Implications

These initial results suggest that LLMs can reproduce key elements of real parliamentary debates, offering a promising step toward simulating more complex, multi-actor policy dialogues. Future work will seek to connect these outputs to public understanding more directly by introducing new stakeholder voices, simulating responses to citizen concerns, and identifying opportunities for consensus or clarification in contested policy areas such as sustainable food systems, public health, and wellbeing.

Keywords

Large Language Models, Parliamentary Debates, Social Simulation