

Week 12 Tutorial: LLM Studies

POP77032 Quantitative Text Analysis for Social Scientists

Tom Paskhalis

Exercise 1: Reviewing LLM Studies

- In the first part we will look at an example of state-of-the-art industry-led social science research on AI.
- First, make sure you read through Huang et al., (2026), paying particular attention to the methodology and substantive conclusions of the carried-out study.
- In groups discuss the following questions:
 1. What was the research question of the study?
 2. What were the dependent and independent variables?
 3. What was the sampling strategy?
 4. What were the main findings of the study?
 5. What are the key strengths and innovations of the study's design and methodology?
 6. What are the weaknesses or limitations of the study's design and methodology?
 7. How does the study fare in terms of its internal and external validity?
 8. Would you expect the results to hold had it been conducted by a different company working in this space?

Exercise 2: Designing LLM Studies

- Imagine that you work for an LLM-development company and you are tasked with designing a study to answer the following research question:

How do people perceive generative AI?

- Assuming as a researcher your utility function is maximising the chances of openly publishing the results while keeping your current position, design a range of studies to answer this question with the following constraints in mind:

- Standard **ethical** considerations, including users' privacy, informed consent, etc.
- **Practical** constraints, such as finite, even if large, user base, non-response, etc.
- **Time** frames (excluding corporate approval and related processes) of:
 - a. *1 year.*
 - b. *6 months.*
 - c. *1 month.*
- **Financial** with 4 different scenarios. Assuming US rates for compensation with higher costs for participant recruitment and lower labour costs for engineering/analysis in other countries.
 - a. *Unlimited budget.* Effectively, an idealised study that you would have conducted under no financial constraints but with other constraints still in place.
 - b. *Budget of \$1,000,000.*
 - c. *Budget of \$100,000.*
 - d. *Only internal company resources.*
- **Corporate** with 3 different scenarios:
 - a. *No corporate constraints.* Effectively, *carte blanche* on part of the senior management to design and conduct the study as you see fit with results being allowed to be published irrespective of the findings.
 - b. *Management review* - the study results reviewed by senior management with the possibility of withholding publication of the results if they are deemed to be damaging to the company's interests.
 - c. *Internal consumption only* - the study results will be used internally by the company for strategic purposes and will not be published.
- Which aspects of the study design would you change/keep across different scenarios?