Quantitative Skill-Building Workshop

Simulating Human Responses with Large Language Models?

Instructor: Do Won Kim
Pronouns: She/Her
Email: dowonkim@umd.edu
Office Hours: Tuesdays/Thursdays, 3:15–4:00pm, HCIL, Hornbake Building, South Wing, 2nd Floor.
- Office hours can also be scheduled flexibly via email.
Course Dates: September 23-October 9, 2025
Course Times: Tuesdays/Thursdays, 2:00–3:15pm
Classroom: Hornbake 2105

Workshop Description

A persistent challenge in both human-computer interaction (HCI) and the social sciences is how to define and measure subjective constructs, such as human values, preferences, and attitudes. Traditional methods such as surveys, interviews, and user studies offer valuable insights, but they are time- and resource-intensive.

Recent advances in large language models (LLMs) have opened new avenues for augmenting or simulating human judgment in such contexts. This workshop examines the potential of LLMs to serve as proxies or collaborators in social science and HCI research by asking a core question: *To what extent can LLMs reliably and validly capture human values and other subjective social constructs?*

Students will learn how to translate abstract social constructs into measurable variables, prompt LLMs to simulate human responses, and evaluate LLM outputs using metrics from survey methodology. They will also explore human-LLM collaborative designs, in which LLMs scaffold or enhance human decision-making rather than replace it.

No prior experience with LLMs or survey design is required. Students will engage hands-on with LLMs, examine the ethical and methodological implications of using them in research, and critically assess their potential roles in simulating, supporting, or scaling human judgment across a range of social and computational contexts.

This workshop also serves as a gateway to a mentored undergraduate research project in Spring 2026. Building on the skills and concepts introduced in this course, students will critically evaluate the capabilities and limitations of LLMs in modeling subjective human preferences, particularly in contexts that require value-sensitive judgment and human-centered design.

Learning Outcomes

By the end of this workshop series, students will be able to:

- Understand how subjective constructs are defined and measured in social science or HCI research.
- Design and test prompt-based simulations of human responses using LLMs.
- Evaluate LLM outputs using established metrics from survey methodology.
- Critically reflect on the use of LLMs as tools in both research and applied settings, including their promises and limitations.

Resources

Course Website: https://umd.instructure.com/courses/1393988
Readings: No required textbook or assigned readings. Recommended readings are listed in the syllabus for optional reference.
Software: LLM access via Hugging Face; Google Colab will be used for experimentation and analysis.
Estimated Cost: $0.00 (assuming use of free-tier or open-source LLMs; no paid tools are required.)

Workshop Structure

This is a hands-on, research-oriented workshop, including:

- A short conceptual lecture introducing foundational ideas
- Interactive tutorials and/or group work
- Reflection and discussion

This workshop does not include exams. Instead, there will be weekly assignments that serve as preparation for data analysis and coding. Completing these in advance will allow us to dive straight into coding during class sessions. The final assignment will be a short survey on potential topics for the Spring semester. Students are encouraged to actively participate in discussions and activities during the workshop and to make use of office hours to further strengthen their skills.

Schedule

#	Date	Topic	Description
1	9/23	Foundations of Survey Measurement	– Introduction to quantitative measurement in survey – Overview of key concepts: consistency, validity
2	9/25	Operationalizing Abstract Constructs	– Turn abstract constructs into measurable variables – Evaluate measurements using sample data
3, 4	9/30, 10/2	Prompting LLMs	– Introduction to prompting LLMs
5	10/7	Evaluating LLM Outputs	– Simulate human responses by prompting LLMs – Apply survey evaluation metrics to LLM outputs
6	10/9	Human-LLM Collaboration	– Explore Human-LLM collaborative designs – Discuss use cases and limitations

Assignment

To-Do
Set up Google Drive folder — before Session 2
Prepare HuggingFace token — before Session 4
Survey on potential research topics — after Session 6

Tips for Success

- **Engage during sessions:** In this workshop, your participation and willingness to share ideas will directly shape what you take away from it.
- **Approach each task as a research exercise:** Rather than aiming for "right answers," treat each activity as an opportunity to experiment, reflect, and build skills. Be comfortable with uncertainty and ambiguity!
- **Use office hours:** If you encounter any issues with the code (or anything), please make active use of office hours. I will be glad to work with you individually to resolve problems. If you are unable to attend during the scheduled office hours, you may arrange a one-on-one meeting with me via email.
  - Please note that, due to limited class time, individual troubleshooting cannot be provided during sessions. To support independent work, I have included detailed comments throughout the code, but difficulties may still arise. In such cases, I strongly encourage you to make use of office hours or request a 1:1 meeting.

Recommended Readings

GenAI and Social Science Survey

- Bail, C. A. (2024). Can Generative AI improve social science? Proceedings of the National Academy of Sciences, 121(21), e2314021121. https://doi.org/10.1073/pnas.2314021121
- Bisbee, J., Clinton, J. D., Dorff, C., Kenkel, B., & Larson, J. M. (2024). Synthetic Replacements for Human Survey Data? The Perils of Large Language Models. Political Analysis, 32(4), 401–416. https://doi.org/10.1017/pan.2024.5
- Dominguez-Olmedo, R., Hardt, M., & Mendler-Dünner, C. (2025). Questioning the survey responses of large language models. In *Proceedings of the 38th International Conference on Neural Information Processing Systems (NeurIPS '24)*, Vol. 37, Article 1458, 45850–45878. Curran Associates Inc., Red Hook, NY. (link)
- And recent articles from the *Sociological Methods & Research* Special Issue on *Integrating Generative AI into Social Science Research* (2025). https://journals.sagepub.com/toc/smra/54/3

Persuasiveness of LLMs

- Matz, S. C., Teeny, J. D., Vaid, S. S., Peters, H., Harari, G. M., & Cerf, M. (2024). The potential of generative AI for personalized persuasion at scale. Scientific Reports, 14(1), 4692. https://doi.org/10.1038/s41598-024-53755-0
- Durmus, E., Lovitt, L., Tamkin, A., Ritchie, S., Clark, J., & Ganguli, D. (2024). Measuring the persuasiveness of language models. *Anthropic*. https://www.anthropic.com/news/measuring-model-persuasiveness