

Bowen Yi

☎ +1 858-776-0900 🌐 bowenyi-pierre.github.io ✉ bowenyi@umich.edu 🐦 @BowenYi2003

Research Interests

Human-Centered NLP/ML, Multimodal AI, AI for health, Behavioral Data Science, Computational Social Science.

Education

Aug. 2025 Present	University of Southern California PhD in Computer Science GPA: - Advisor: Shrikanth Narayanan	Los Angeles, CA
Jan. 2023 May. 2025	University of Michigan - Ann Arbor B.S. in Computer Science GPA: 3.79/4.0	Ann Arbor, MI
Sep. 2021 Dec. 2022	University of California - San Diego B.S. in Mathematics-Computer Science, Minor in Sociology GPA: 3.83/4.0	La Jolla, CA

Publications

S=In Submission, C=Conference, W=Working Paper

- [C.1] **Examining Spanish Counseling with MIDAS: a Motivational Interviewing Dataset in Spanish** [🔗]
Aylin Gunal*, Bowen Yi*, John Piette, Rada Mihalcea, Veronica Perez-Rosas gens (* = Equal Contribution)
Conference of the Nations of the Americas Chapter of the Association for Computational Linguistics [NAACL 2025]
- [C.2] **Causally Modeling the Linguistic and Social Factors that Predict Email Response** [🔗]
Yinuo Xu*, ..., Bowen Yi*, ..., David Jurgens (* = Equal Contribution with Random Order)
Conference of the Nations of the Americas Chapter of the Association for Computational Linguistics [NAACL 2025]
- [C.3] **The Generation Gap: Exploring Age Bias Underlying in the Value Systems of Large Language Models** [🔗]
Siyang Liu and Trish Maturi and Bowen Yi and Siqi Shen and Rada Mihalcea
Conference on Empirical Methods in Natural Language Processing, Miami, USA [EMNLP 2024]
- [S.1] **NLP for Social Good: A Survey of Challenges, Opportunities and Responsible Deployment** [🔗]
Antonia Karamolegkou, ..., Bowen Yi, ..., Daryna Dementieva
[Under review at EMNLP 2025]
- [S.2] **Real or Robotic? Assessing Whether LLMs Accurately Simulate Qualities of Human Responses in Dialogue** [🔗]
Jonathan Ivey*, ..., Bowen Yi*, ..., David Jurgens (* = Equal Contribution with Random Order)
[Under review at EMNLP 2025]
- [W.1] **Unveiling Behavioral Differences in Bilingual Information Operations: A Network-Based Approach** [🔗]
Bowen Yi
[Preprint on arXiv]
- [W.2] **Uncovering the Impact of Intervention Messages on Diverse Population Groups**
Bowen Yi, Rada Mihalcea, Fang Yu, Elena Frank, Joan Zhao, Srijan Sen, Maggie Makar

Research Experience

Behavioral Data Science

Causal Effect of Intervention Messages on Medical Interns' Well-being [W.2] Feb. 2024 - Apr. 2025

Advisors: [Maggie Makar](#), [Rada Mihalcea](#)

- Collaborate with [Srijan Sen's Lab](#) at Michigan Psychiatry to analyze 4 years of intervention message data and behavioral data from **wearable devices** for 6,000+ medical interns, including metrics such as mood, steps, and sleep.
- Measure CATE and isolated causal effects of intervention message categories on patient subgroups, utilizing tools like EconML and scikit-learn.
- Develop interpretable language representations for 939 intervention messages to enable valid causal inference, based on 14 expert-encoded psychotherapy categories and lexicons like LIWC.
- Design robust subgroup-discovery methods that accommodate data perturbations and time-varying interventions, employing DoubleML and Random Forest to identify patient subgroups with significant treatment responses, demonstrating improved stability compared to baseline methods.

Human-Centered NLP

Simulating Culturally-Aligned Mental Health Counseling

Sep. 2024 - Apr. 2025

Advisor: [Rada Mihalcea](#)

- > Simulate counselor-patient conversations in mental health from Hispanic culture using LLMs, leveraging **visual** information (interpreted using multimodal LLMs like video-ChatGPT), **speech** features (such as tone and pitch), and **text** data from the dataset in [C.1].
- > Align LLMs with doctor and patient behaviors in Spanish culture by exploring strategies such as prompt engineering, fine-tuning, **reinforcement learning**, and **multi-agent frameworks**.
- > Develop a human-LLM collaborative framework to evaluate cultural sensitivity, behavioral alignment, and counseling quality of the simulated counselor, following criteria like [ENACT](#) while reducing reliance on existing professional knowledge-based benchmarks.

Assessing Alignment of LLM Simulated Human Responses [S.1]

July. 2024 - Sep. 2024

Advisors: [David Jurgens](#)

- > Evaluated alignment of LLM simulations with human interactions on 100,000 English, Chinese, and Russian dialogues from [the WildChat dataset](#).
- > Co-designed an evaluation framework with metrics to assess lexical, stylistic, and semantic alignment between simulated human dialogues and real human-LLM multilingual interactions.
- > Found a low alignment across all three languages among nine tested models even when prompted in the target language, particularly in lower-resource settings like Russian.

Inspecting Age Bias in the Value Systems of LLMs [C.3]

April. 2024 - Jun. 2024

Advisors: [Rada Mihalcea](#)

- > Analyzed the alignment of social, economic, and other 11 categories of world values across six age groups in 62 countries on 6 leading LLMs, leveraging data from the [World Values Survey](#).
- > Responsible for evaluating the age bias in [the Mistral model](#) and studying the impact of age identity prompts on value misalignment.
- > Findings suggested a general inclination of LLM values towards younger demographics, especially when tested on the US population.

Computational Social Science

Climate Change and Socio-Political Stability in Sub-Saharan Africa [🌍]

May. 2024 - May. 2025

Advisor: [Verónica Pérez-Rosas](#)

- > Collaborate with social scientists from [Arun Agrawal](#)'s lab to examine causal relationships between climate or demographic changes and sociopolitical stability as reflected in scientific literature.
- > Develop data infrastructure, analytical models, and an automated pipeline to collect and process over 20,000 relevant scientific articles, leveraging tools like Grobid, Selenium, and Beautiful Soup.

Examining Enhanced Behavioral Counseling in Spanish [C.1]

Apr. 2024 - Apr. 2025

Advisors: [Verónica Pérez-Rosas](#), [Rada Mihalcea](#)

- > Collaborate with [John Piette](#) to introduce the first Spanish Motivational Interviewing dataset from 74 public videos, annotated by experts on counseling strategies.
- > Analyzed culture-driven behavioral and psycholinguistic differences between the Spanish and English therapists, such as the word-exchange ratio, language style matching, language usage (LIWC), and sentiment.
- > Designed [P-Tuning experiments](#) on multilingual and monolingual LLMs to predict counselor behaviors such as reflections and questions.
- > Investigating the impact of visual captioning in counseling videos on model performance for classifying therapist strategies and generating personalized feedback.

Uncovering the Impact of George Floyd Incident on Podcast Ecosystem [Report, Dataset]

Aug. 2023 - Jun. 2024

Advisors: [David Jurgens](#), [Dallas Card](#)

- > Modeled the topical trends (using Mallet), [political bias](#), and conversation dynamics in 600,000+ transcribed podcast episodes published from May to June 2020.
- > Investigated the impact of George Floyd's death on podcasts across 79 themes, identifying named entities mentioned with George Floyd that reflect themes of social justice and police violence.
- > Fine-tuned and calibrated language models (e.g., RoBERTa, MiniLM) to classify news content and sentiment in podcasts, with plans to incorporate **audio** features such as pitch into classifiers for better performance.

Modeling Conversational Dynamics in Email Exchanges [C.2]Mar. 2023 - Jun. 2024

Advisor: David Jurgens

- > Analyzed 11.3M emails from the GMANE corpus, identifying and measuring key conversational dynamics and social-linguistic factors of email exchanges including intimacy, formality, and cogency.
- > Created a dataset of 1,800 emails annotated for intents, expectations, and 14 pragmatic features; benchmarked models including logistic regression and zero-shot LLMs.
- > Conducted causal analysis revealing that social status, argumentation quality, and social connection influence response rates, with social status being the most significant.

Academic Service

Reviewer EMNLP 2025, GenAI4Health @ NeurIPS 2025
Program Committee GenAI4Health @ NeurIPS 2025

Industry Experience

Software Engineering Intern | Boardx.us Jun. 2023 - Aug. 2023

Mentor: Mr. Feng Zhang

- > Contributed to designing collaborative/interactive digital whiteboards with a focus on improving user experience and managing company data on Microsoft Azure.
- > Assisted in integrating AI chatbots to inspire creativity in user design and writing.
- > Introduced accessibility features, including color-blind-friendly design, to make the product more inclusive.

Presentations

“Examining Spanish Counseling with MIDAS”

- > NAACL Main Conference 2025 [Poster] | [Recording] May. 2025 (Albuquerque, NM)
- > MSLD 2025 [Poster] Apr. 2025 (Notre Dame, IN)

“Causally Modeling the Linguistic and Social Factors that Predict Email Response”

- > NAACL Main Conference 2025 [Poster] May. 2025 (Albuquerque, NM)

“NLP for Enhanced Behavioral Counseling in Spanish”

- > BoF Session on Large Multimodal Models for Biomedical Research, EMNLP 2024 Nov. 2024 (Miami, FL)
- > Michigan AI Symposium, Embodied AI [Poster] Oct. 2024 (Ann Arbor, MI)
- > e-HAIL Symposium, Generative AI in Healthcare [Poster] Sep. 2024 (Ann Arbor, MI)

“Can LLMs Simulate Human Subject Studies?”

- > NLP Reading Group, Michigan AI Lab [Slides] | [Note] Sep. 2024 (Ann Arbor, MI)

“Podcast Study: Uncovering News in Non-News Space”

- > David Jurgens Lab Meeting, UMSI [Slides] Feb. 2024 (Ann Arbor, MI)

“Tutorial: Topic Modeling with Mallet”

- > David Jurgens Lab Meeting, UMSI [Slides] Nov. 2023 (Ann Arbor, MI)

Honours and Awards

Viterbi School of Engineering Fellowship, 2025 - 2026 | University of Southern California [🌐] \$44,290 per year.
Departmental Fellowship, 2025 | University of Utah [🌐] Declined to join USC.
Excellence Fellowship, 2025 - 2028 | Stony Brook University [🌐] \$10,000 for 3 Years. Declined to join USC.
James B. Angell Scholar, 2025 | University of Michigan [🌐] For 2 consecutive terms of all A’s.
University Honors, 2023 - 2025 | University of Michigan [🌐] For achieving a GPA above 3.5 in every full-time term.
IEEE-Eta Kappa Nu, 2022 & 2024 | UC San Diego & University of Michigan [🌐] Invited membership for ranking in the top 25% of the junior class and top 33% of the senior class
Provost Honors, 2021 - 2022 | UC San Diego [🌐] For achieving a GPA above 3.5 in every full-time term.

Teaching and Leadership

- Grader, EECS 486 Information Retrieval @ UMich**

Jan. 2025 - Apr. 2025
- > Grade and provide detailed feedback on the coding assignments, exam, and final project for 100+ students.
- NLP Reading Group, University of Michigan** *Co-Organizer* [🌐]

Sep. 2024 - Apr. 2025
- > Organize and host weekly paper presentations by designing event setups and providing speaker tutorials.
- > Communicate NLP research to attendees from CS and non-CS fields (such as Medicine and Psychology) and undergraduates with limited experience, building an inclusive community of 180 active members.
- Volunteer Mentor for College Applications**

Oct. 2021 - Mar. 2023
- > Supported about 10 financially challenged domestic and international students through their first-year or transfer college applications.
- > Provided constructive feedback on essay brainstorming, standardized test preparation, major and college choice, etc.

Skills

DL/ML Programming	PyTorch, TensorFlow, PyTorch-Transformers, Scikit-Learn
Programming Languages	Python, C/C++, Java, HTML/CSS
Natural Languages	Hunanese (native), Mandarin (native), English (fluent), Greek (learning), Serbian (learning)