

Bowen Yi

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

Research Interests

My research interests lie in **Human-Centered NLP**, **Behavioral Data Science**, and **Computational Social Science**. I focus on modeling social interactions and how **social roles** and **cultural differences** impact individual behaviors in **multimodal** contexts, with applications in **mental health** and education. Moreover, I design methods that leverage **causal inference**, **ML**, and social science theories to derive actionable insights from large-scale behavioral data.

Education

Jan. 2023 Present	University of Michigan - Ann Arbor B.S. in Computer Science GPA: 3.89/4.0 Coursework: Natural Language Processing, Human-Centered ML, Machine Learning, Computer Vision, Information Retrieval, Human-Centered Software Design and Development	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 Coursework: Sociology of Health Care Issues, Introduction to Sociology, Statistics, Languages and Cultures, Linear Algebra, Calculus, Programming and Object-Oriented Design	La Jolla, CA

Publications

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

- [C.1] **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.3] **Examining Spanish Counseling with MIDAS: a Motivational Interviewing Dataset in Spanish** [🔗]
Aylin Gunal, Bowen Yi, John Piette, Rada Mihalcea, Veronica Perez-Rosas
[Under Review at NAACL 2025]
- [S.2] **Real or Robotic? Assessing Whether LLMs Accurately Simulate Qualities of Human Responses in Dialogue** [🔗]
Jonathan Ivey*, Shivani Kumar*, Jiayu Liu*, Hua Shen*, Sushrita Rakshit*, Rohan Raju*, Haotian Zhang*, Aparna Ananthasubramaniam*, Junghwan Kim*, Bowen Yi*, Dustin Wright*, Abraham Israeli*, Anders Giovanni Møller*, Lechen Zhang*, David Jurgens (* = Equal Contribution with Random Order)
[Under Review at NAACL 2025]
- [S.1] **Causally Modeling the Linguistic and Social Factors that Predict Email Response** [🔗]
Yinuo Xu*, Sushrita Rakshit*, Aparna Ananthasubramaniam*, Omkar Yadav*, Mingqian Zheng*, Michael Jiang*, Lechen Zhang*, Bowen Yi*, Kenan Alkiek*, Abraham Israeli*, Bangzhao Shu*, Hua Shen*, Jiaxin Pei*, Haotian Zhang*, Miriam Schirmer*, David Jurgens (* = Equal Contribution with Random Order)
[Under Review at NAACL 2025]
- [W.1] **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 [S.4] Feb. 2024 - Present

Advisors: *Maggie Makar, Rada Mihalcea*

- Collaborate with **Srijan Sen's Lab** at Michigan Psychiatry to analyze 4 years of intervention message data and daily behavioral data from 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 and **sentecon**.
- 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 - Present

Advisor: [Rada Mihalcea](#)

- > Simulate counselor-patient conversations in mental health from Hispanic culture using LLMs, leveraging **visual** information (interpreted using video-ChatGPT), **speech** features (such as tone and pitch), and **text** data from the dataset in [S.3].
- > Align LLMs with doctor and patient behaviors in Spanish culture by exploring strategies such as prompt engineering, fine-tuning, **reinforcement learning** (e.g. [DPO plus mechanistic understanding](#)), and **multi-agent frameworks**.
- > Develop a human-LLM collaborative **evaluation** framework to assess cultural sensitivity, behavioral alignment, and potential counseling quality of simulated personas, following criteria like [ENACT](#) while minimizing reliance on existing professional knowledge-based benchmarks.

Assessing Alignment of LLM Simulated Human Responses [S.2]

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.1]

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 [C.2]

May. 2024 - Present

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 [S.3]

Apr. 2024 - Nov. 2024

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. The dataset is used in ongoing work to simulate culturally aware doctor-patient interaction.
- > 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.

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 [S.1]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.

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

“NLP for Enhanced Behavioral Counseling in Spanish”

- > BoF Session on Large Multimodal Models for Biomedical Research, EMNLP 2024Nov. 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

University Honors, 2023 - 2024 | 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 to 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

NLP Reading Group, University of Michigan Co-Organizer Sep. 2024 - Present

- > 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 ApplicationsOct. 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 ProgrammingPyTorch, TensorFlow, PyTorch-Transformers, Scikit-Learn

Programming LanguagesPython, C/C++, Java, HTML/CSS

Natural LanguagesHunanese (native), Mandarin (native), English (fluent), Cantonese (conversational), German (basic), Spanish (learning), Serbian (learning)