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 behavior in **multimodal data**, with applications in **mental health** and **education**. Moreover, I design methods that leverage **causal inference**, data science, and social science theories to derive actionable insights from large-scale behavioral data.

Education

Jan. 2023 Present	**University of Michigan - Ann Arbor**	**Ann Arbor, MI**

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

Sep. 2021 Dec. 2022	**University of California - San Diego**	**La Jolla, CA**

B.S. in **Mathematics-Computer Science**, Minor in **Sociology** | GPA: 3.83/4.0
Coursework: Sociology of Health Care Issues, Introduction to Sociology, Statistics (in Psychology), Languages and Cultures, Linear Algebra, Calculus, Programming and Object-Oriented Design

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 <u>Bowen Yi</u> 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, <u>Bowen Yi</u>, 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*, <u>Bowen Yi</u>*, 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*, <u>Bowen Yi</u>*, 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**
<u>Bowen Yi</u>, 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 explore their 4-year data on intervention messages and 6,000+ medical interns' daily behavioral data, including mood, steps, and sleep.

> Estimate the heterogeneous treatment effects of different categories of intervention messages on various subgroups of patients, employing tools such as EconML and scikit-learn.

> Explore better language representations of intervention messages for valid causal inference, building on Lin et al.'s work on isolated causal effect and sentecon.

> Develop a subgroup-discovery method robust to data perturbations for identifying patient subgroups with significant treatment reactions, utilizing DoubleML and Random Forest.

Human-Centered NLP

Simulating Culturally-Aligned Mental Health Counseling   Sep. 2024 - Present
Advisor: Rada Mihalcea

> Simulate counselor-patient conversations in mental health from Spanish culture using LLMs, based on **visual**, **speech**, and **text** information of the dataset in **[S.3]**.

> Align LLMs with doctor and patient behaviors in Spanish culture by exploring strategies such as prompt-based fine-tuning, 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.

NLP for Enhanced Behavioral Counseling in Spanish   [S.3]   Apr. 2024 - Nov. 2024
Advisors: Verónica Pérez-Rosas, Rada Mihalcea

> Collaborate with John Piette to introduce a new Spanish counseling 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 dataset and a parallel English dataset, such as the word-exchange ratio, language usage (LIWC), and sentiment.

> Designed P-Tuning experiments on multilingual and monolingual LLMs to predict counselor behaviors such as reflections and questions.

Assessing LLMs' Simulation of Human Responses in Dialogue   [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 and associated lexical/stylistic/semantic metrics to assess the alignment of simulated dialogues to real dialogues in multilingual settings.

> Found a low alignment across all three languages among nine tested models, especially in lower-resource languages 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 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 - 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 web scraping tools like Selenium and Beautiful Soup.

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 podcast discussions 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 within podcast transcripts, 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

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 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

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 non-CS fields (such as Medicine and Psychology) and undergraduates with limited experience, building an inclusive community of 180 active members.

Volunteer Tutor for College Applications　　　　　　　　　　　　　　　　　Oct. 2021 - Mar. 2023

> Supported about 10 financially challenged international students from countries such as Mexico and Bangladesh 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), Cantonese (conversational), German (basic), Spanish (learning), Serbian (learning)