Ang Li

Research Scientist at Spotify

linkedin.com/in/angli432100 github.com/LittleRabbitHole http://www.pitt.edu/~anl125/
520-247-8820 ANL125@pitt.edu or li.ang198744@gmail.com

About

Experienced quantitative researcher with a Ph.D. in Information Science, specializing in Human-Computer Interaction and data science. With over 2 years of industry expertise, I've successfully led research projects focused on user understanding and metric development. At Spotify, I played a pivotal role in leading impactful research initiatives, collaborating with cross-functional teams of data scientists, researchers, and engineers. My work spans various missions, including developing metrics to estimate the impact of recommendation interventions, providing insights to shape recommendation strategies, and offering valuable perspectives on limitations and improvements. Eager to continue contributing insights and expertise to dynamic projects in the intersection of research, data science, and technology.

Education

2016 – 2021 **University of Pittsburgh**, Pittsburgh, PA
PhD in Information Science; 3.9/4.0 GPA
Thesis : Collaborative Contribution of News Related Open Content in Social Media
Committee : Rosta Farzan (Chair), Yu-Ru Lin, Daqing He, Brian Keegan, Jenn Thom

2013 - 2016 **DePaul University**, Chicago, IL
Master of Science in Predictive Analytics, 3.9/4.0 GPA

Research Experience

01/04/2022 – Present

Spotify, New York City, NY
Research Scientist, Supervised by Dr. Sam Way

Develop causal metrics to drive audience growth for long-tail podcast creators : Led the development of causal metrics to drive audience growth for long-tail podcast creators by spearheading collaborative efforts with researchers, data scientists, and engineers.

- Utilized causal modeling and offline observational data to measure the impact of diverse recommendation and promotion strategies.
- Collaborated closely with data scientists and engineers to design and implement these causal metrics, ensuring robust validation methods.
- Successfully integrated the developed causal metrics into production as one of primary success indicators for assessing the effectiveness of recommendation and promotion strategies.
-

Audiobook recommendation : Collaborated with researchers, data scientists, and engineers to pioneer the creation of Spotify's inaugural personalized audiobook recommendation system.

- Conducted research dedicated to comprehending user listening habits across various content types and deciphering signals associated with audiobook interactions.
- Derived crucial insights from the research, shaping the design of audiobook recommendations.
- Results from online tests demonstrated a remarkable enhancement, with the recommendation algorithms boosting audiobook interaction metrics by over 20%
-

Develop financial metrics for estimating diverse recommendation interventions : Led research projects by collaborating with researchers, data scientists, product managers, and engineers to develop methodologies for quantifying the financial impact of diverse recommendation and promotion interventions across various missions.

- Established key analysis frameworks to measure and quantify financial impacts, including ads revenue, premium revenue, and associated costs, arising from different recommendation and promotion strategies.
- Uncovered valuable insights into the advantages and limitations of the analysis framework, offering recommendations for immediate next steps.
- The insights and recommendations served as crucial guidance for mission leads when making decisions on the prioritization of developing tools for quantifying financial impact in mission-level initiatives.

June 2019 – August 2019

Spotify, New York City, NY

Research Scientist Internship, Supervised by Dr. Benjamin Carterette

Research on understanding users' Music versus Podcast consumption habits.

- Adopting propensity score matching and difference-in-difference methods, we assessed the causal influence of adding podcasts listening on user music consumption behavior by using large scale observational data collected over one year
- Utilized statistical methods based on the large-scale log data, we examined users' consumption habits for podcasts vs. music and uncovered the differences.
- Utilized statistical methods and machine learning techniques based on the large-scale log data, we examined users' behavior to infer users mindset as they approach a music search
- A machine learning model was then developed to predict users' listening content based on their listening habit, and the model can achieve high accuracy rate.
- The research was used to guide Spotify business strategy and was mentioned by CTO at 2021 Spotify Investor Day. The research paper was presented in the Web Conference 2020. [C.7]

June 2018 – August 2018

Spotify, Boston, MA

Research Scientist Internship, Supervised by Dr. Jenn Thom

Research on the music search mindset project : This project intends to understand how users seek information in the domain of the music search.

- Utilized a mixed-method approach that combines both qualitative user studies with quantitative statistical analyses to uncover how users search music within the music streaming platform.
- Guided by the usability test and interview study, we designed and conducted a user survey to gather data directly from users about their mindsets when they search music using mobile devices.
- Utilized statistical methods and machine learning techniques based on the large-scale log data, we examined users' behavior to infer users mindset as they approach a music search. The model achieve good accuracy level.
- The research paper was presented in the Web Conference 2019. [C.6]

June 2020 – August 2020

Google, Mountain View, CA

UX Research Internship, Supervised by : Dr. Tao Dong

Conducted quantitative research to support Flutter open source software (OSS) community.

- Utilized a mixed-method approach, the project aims to improve the current pull request (PR) triaging and reviewing process while maintainers interacting with the PR contributors.
- Utilized qualitative research method, we interviewed project managers and evaluated the project documents to understand the ideal attributes of good PRs maintainers would like to see and draw hypotheses on what maintainers are expected to do during PR reviewing and accepting process.
- Utilized statistical methods based on the large-scale log data, we investigated the maintainers' current practices in triaging and reviewing PRs.
- The results identify different gaps existed in the current system, and also provide recommendations on how the system can be better designed to support project maintainers

September 2016 – November 2011

University of Pittsburgh, Pittsburgh, PA

Research Assistant, Research Supervisor : Dr. Rosta Farzan, Dr. Yu-Ru Lin

My dissertation study aims to understand the content production process by users in various social media platforms including Wikipedia and Twitter. The results provide insights on how to design social computing systems that can provide more inclusive and less polarized user-generated content (see publications [C.2, C.4, C.5, C.8, C.9]). I have rich experiences in both qualitative research approach such as content analysis and human annotation as well as quantitative methods such as analyzing and modeling the large-scale human generated online behavior data including

- Generalized regression analysis to evaluate the relationship between users' content production process and the content quality/bias as outcomes;
- Survival analysis to evaluate the member retention in the platform;
- Develop coding scheme and utilize qualitative content analysis to understand how different types of the communications could help to engage current users;
- Mediation analysis to examine the social interactions as mediator factors that influence members' production and retention;
- Network analysis to discover social interaction patterns;
- NLP techniques (e.g. topic modeling, word-embedding, sentiment analysis, etc.) to process the user-generated content and extract the linguistic features and topics.
-

Selected Publications

Peer-Reviewed Conference and Journal Papers :

[C.10]. **Li, A.**, Farzan, R., & López, C. (2022). Let's Work Together! Wikipedia Language Communities' Attempts to Represent Events Worldwide. Interacting with Computers, 2022

[C.9]. **Li, A.**, Farzan, R., Lin, Y. R., Zhou, Y., Teng, X., & Yan, M. (2022). Identifying and Understanding Social Media Gatekeepers : A Case Study of Gatekeepers for Immigration Related News on Twitter. Proc. ACM Hum.-Comput. Interact. 6, CSCW2, Article 304 (November 2022)

[C.8]. **Li, A.**, Yao, Z., Yang, D., Kulkarni, C., Farzan, R., & Kraut, R. E.. (2020, May). Successful Online Socialization : Lessons from the Wikipedia Education Program. In Proceedings of ACM Hum.-Comput. Interact. 4, CSCW1, Article 050 (May 2020) **Honorable Mention award in CSCW 2020, Top 5%.**

[C.7]. **Li, A.**, Wang, A., Nazari, Z., Chandar, P., & Carterette, B. (2020, April). Do podcasts and music compete with one another? Understanding users' audio streaming habits. In Proceedings of *The Web Conference 2020* (pp. 1920-1931).ACM.

[C.6]. **Li, A.**, Thom, J., Chandar, P., Hosey, C., Thomas, B. S., & Garcia-Gathright, J. (2019, May). Search Mindsets : Understanding Focused and Non-Focused Information Seeking in Music Search. In Proceedings of *The World Wide Web Conference 2019* (pp. 2971-2977). ACM.

[C.5]. Ertugrul, A. M., Lin, Y. R., Chung, W. T., Yan, M., & **Li, A**. (2019). Activism via attention : interpretable spatiotemporal learning to forecast protest activities. *EPJ Data Science*, 8(1), 1-26.

[C.4]. **Li, A.**, & Farzan, R. (2018, September). Keeping up on Current Events! A Case Study of Newcomers to Wikipedia. In *International Conference on Social Informatics* (pp. 348-369). Springer, Cham.

[C.3]. Chung, W. T., Lin, Y. R., **Li, A.**, Ertugrul, A. M., & Yan, M. (2018, September). March with and Without Feet : The Talking About Protests and Beyond. In *International Conference on Social Informatics* (pp. 134-150). Springer, Cham.

[C.2]. Zheng, K., **Li, A.**, & Farzan, R. (2018, March). Exploration of Online Health Support Groups Through the Lens of Sentiment Analysis. In *International Conference on Information* (pp. 145-151). Springer, Cham.

[C.1]. Birnholtz, J., Davison, J., & **Li, A.** (2017). Attending to attention : How do people attract, manage, and negotiate attention using mobile devices? *Mobile Media & Communication*, 2050157917714504.

Academic Services

- Associate Chair and committee member of ACM Conference on Computer Supported Cooperative Work and Social Computing 2023 (CSCW 2023)
- Reviewer of ACM CHI Conference on Human Factors in Computing Systems (CHI) 2023
- Reviewer of ACM Conference on Computer Supported Cooperative Work and Social Computing (CSCW) 2018 till 2022
- Reviewer of ACM Transactions on Social Computing Journal
- Student volunteer for CSCW 2018

Skills

Statistical Analysis	R, Python, Matlab
Database	Relational Database, SQL
Data Visualization	Gephi, Tableau, D3.js