Do, Hyo Jin

Cambridge, MA, US | +1-217-419-7122 | dohyojin90@gmail.com | [Google Scholar](#) | [Homepage](#) | [LinkedIn](#)

RESEARCH INTERESTS

My research lies at the intersection of **human-computer interaction (HCI)** and **artificial intelligence (AI)**. I focus on designing and developing **human-centered AI systems** that enhance human-AI interaction, improve user experience and technology acceptance, and mitigate potential harms and risks. Key research projects include:

- Exploring methods for presenting factuality scores and source attributions of LLM outputs [1, 2, 5, d4]
- Developing EvalAssist, an LLM-as-a-judge tool that improves human-AI alignment [3, 4, d1]
- Inventing AI disclosure approaches to explain AI contributions in human-AI co-created content [w1, d3]
- Designing conversational natural language interfaces to enhance human-AI collaboration [6, 7, 8]
- Directing user research on RAG hallucination detection that shaped an award-winning governance tool [d2]
- Implementing conversational agents for team collaboration [9, 10, 11] and empathetic interactions [12, d5]

EDUCATION

University of Illinois at Urbana-Champaign (UIUC)   Urbana, IL, US

Ph.D. in Computer Science | Interactive Computing Group   Aug 2016 - May 2023

- Advisor: Brian P. Bailey | Co-advisor: Karrie Karahalios
- Dissertation: Designing a Conversational Agent for Promoting Participation of Under-Contributing Members in a Small Group Chat Discussion (Link)
- GPA: 3.91/4.0 (98/100)

Korea Advanced Institute of Science and Technology (KAIST)   Daejeon, South Korea

Master of Science in Computer Science | Artificial Intelligence Group   Sep 2013 - Feb 2016

- Advisor: Ho-Jin Choi
- Thesis: Korean Twitter Emotion Classification and Application (Link)
- GPA: 3.74/4.3 (94/100)

Korea University (KU)   Seoul, South Korea

Bachelor of Science in Computer Science (Minor: Brain and Cognitive Sciences)   Mar 2009 - Aug 2013

- Honor: Valedictorian
- GPA: 4.07/4.5 (95.1/100)

National University of Singapore (NUS)   Singapore

Exchange Student in Computer Science   Aug 2010 - May 2011

PROFESSIONAL EXPERIENCE

IBM Research   Cambridge, MA, US

Research Scientist (Full-time) | Human-centered Trustworthy AI Team   Jan 2023 – Dec 2025

- Managed end-to-end research workflows, from defining problem spaces, designing experimental protocols, and running user studies to analyzing qualitative and quantitative data and publishing findings
- Contributed to open-source projects (e.g., EvalAssist, AI Attribution Toolkit) and products (e.g., Bob, AppConnect, Watsonx.governance)
- Received O-Level Accomplishment Award (2025) and Outstanding Technical Achievement Award (2024)

IBM Research   Cambridge, MA, US

AI Intern | AI Experience Team   May 2022 - Aug 2022

- Designed conversational agents for end-user programming tasks and conducted user studies

IBM Research   Yorktown Heights, NY, US

Research Graduate Intern | Finance Research Team   May 2019 - Aug 2019

- Conducted user studies to understand model risk management practices in financial institutions

Bank of Korea   Seoul, South Korea

Economist   Jan 2016 - Jun 2016

- Administered Economic Statistics System (ecos.bok.or.kr) and managed IT-related projects

Samsung Electronics   Suwon, South Korea

Software Engineer Intern | Android Development Team   Dec 2011 - Feb 2012

- Designed and implemented Android applications, won the Best Team Project Award

TEACHING EXPERIENCE

Teaching Assistant   Urbana, IL, US

University of Illinois Urbana-Champaign   Aug 2016 - Dec 2021

Honors & Awards:
- Graduate Teacher Certificate (Spring 2022)
- List of Teachers Ranked as Excellent (Fall 2021)

Courses:
- CS465 User Interface Design (Spring 2017-19, Fall 2019, Spring 2021, Fall 2021)
- CS498/416 Data Visualization (Summer 2020 & 2021)
- CS125 Introduction to Computer Science (Spring & Fall 2020)
- CS548 Models of Cognitive Processes (Fall 2018)
- CS225 Data Structures (Fall 2017)
- CS440 Artificial Intelligence (Fall 2016)

Teaching Assistant   Daejeon, South Korea

Korea Advanced Institute of Science and Technology (KAIST)   Sep 2013 - Jun 2014

Courses:
- CS476/BiS481 Collective Intelligence in Biomedical Applications (Spring 2014)
- CS360 Introduction to Databases (Fall 2013)

Lecturer (World Friends ICT Volunteer)   Accra, Ghana

University of Ghana   Jul 2012 - Aug 2012

- Lectured Android programming course to university undergraduate and graduate students every weekday

KEY PUBLICATIONS

[1] **Hide or Highlight: Understanding the Impact of Factuality Expression on User Trust**
   Hyo Jin Do and Werner Geyer
   Proceedings of the AAAI/ACM Conference on Artificial Intelligence, Ethics and Society (AIES), 2025

[2]   **Highlight All the Phrases: Enhancing LLM Transparency through Visual Factuality Indicators**
      Hyo Jin Do, Rachel Ostrand, Werner Geyer, Keerthiram Murugesan, Dennis Wei, and Justin Weisz
      Proceedings of the AAAI/ACM Conference on Artificial Intelligence, Ethics and Society (AIES), 2025

[3]   **Generate, Evaluate, Iterate: Synthetic Data for Human-in-the-Loop Refinement of LLM Judges**
      Hyo Jin Do, Zahra Ashktorab, Jasmina Gajcin, Erik Miehling, Martín Santillán Cooper, Qian Pan, Elizabeth M. Daly, Werner Geyer
      Preprint, 2025

[4]   **EvalAssist: Insights on Task-Specific Evaluations and AI-Assisted Judgment Strategy Preferences**
      Zahra Ashktorab, Michael Desmond, Qian Pan, James M. Johnson, Martin Santillán Cooper, Elizabeth M. Daly, Rahul Nair, Tejaswini Pedapati, **Hyo Jin Do**, and Werner Geyer
      Proceedings of the Annual ACM Symposium on User Interface Software and Technology (UIST), 2025

[5]   **Multi-Level Explanations for Generative Language Models**
      Lucas Monteiro Paes, Dennis Wei, **Hyo Jin Do**, Hendrik Strobelt, Ronny Luss, Amit Dhurandhar, Manish Nagireddy, Karthikeyan Natesan Ramamurthy, Prasanna Sattigeri, Werner Geyer, and Soumya Ghosh
      Proceedings of the Annual Meeting of the Association for Computational Linguistics (ACL), 2025

[6]   **Grounding with Structure: Exploring Design Variations of Grounded Human-AI Collaboration in a Natural Language Interface**
      Hyo Jin Do, Michelle Brachman, Casey Dugan, James M. Johnson, Julia Lauer, Priyanshu Rai, and Qian Pan
      Proceedings of the ACM on Human-Computer Interaction (PACM HCI), CSCW, 2024

[7]   **Evaluating What Others Say: The Effect of Accuracy Assessment in Shaping Mental Models of AI Systems**
      Hyo Jin Do, Michelle Brachman, Casey Dugan, Qian Pan, Priyanshu Rai, James M. Johnson, Roshni Thawani
      Proceedings of the ACM on Human-Computer Interaction (PACM HCI), CSCW, 2024

[8]   **Follow the Successful Herd: Towards Explanations for Improved Use and Mental Models of Natural Language Systems**
      Michelle Brachman, Qian Pan, **Hyo Jin Do**, Casey Dugan, Arunima Chaudhary, James M Johnson, Priyanshu Rai, Tathagata Chakraborti, Thomas Gschwind, Jim A Laredo, Christoph Miksovic, Paolo Scotton, Kartik Talamadupula, and Gegi Thomas
      Proceedings of the International Conference on Intelligent User Interfaces (IUI), 2023

[9]   **Inform, Explain, or Control: Techniques to Adjust End-User Performance Expectations for a Conversational Agent Facilitating Group Chat Discussions**
      Hyo Jin Do, Ha-Kyung Kong, Pooja Tetali, Karrie Karahalios, and Brian P Bailey
      Proceedings of the ACM on Human-Computer Interaction (PACM HCI), CSCW, 2023

[10] **To Err is AI: Imperfect Interventions and Repair in a Conversational Agent Facilitating Group Chat Discussions**
      Hyo Jin Do, Ha-Kyung Kong, Pooja Tetali, Jaewook Lee, and Brian P Bailey
      Proceedings of the ACM on Human-Computer Interaction (PACM HCI), CSCW, 2023

[11] **How Should the Agent Communicate to the Group? Communication Strategies of a Conversational Agent in Group Chat Discussions**
      Hyo Jin Do, Ha-Kyung Kong, Jaewook Lee, and Brian P. Bailey
      Proceedings of the ACM on Human-Computer Interaction (PACM HCI), CSCW, 2022

[12] **Do You Have Time for a Quick Chat? Designing a Conversational Interface for Sexual Harassment Prevention Training**
      Hyo Jin Do, Seon Hye Yang, Boo-Gyoung Choi, Wayne T. Fu, and Brian P. Bailey
      Proceedings of the International Conference on Intelligent User Interfaces (IUI), 2021

[13] **Burst Your Bubble! An Intelligent System for Improving Awareness of Diverse Social Opinions**
      Mingkun Gao, **Hyo Jin Do**, and Wai-Tat Fu

Proceedings of the International Conference on Intelligent User Interfaces (IUI), 2018

[14] **An Intelligent Interface for Organizing Online Opinions on Controversial Topics**
Mingkun Gao, **Hyo Jin Do**, and Wai-Tat Fu
Proceedings of the International Conference on Intelligent User Interfaces (IUI), 2017

ORGANIZED WORKSHOP

[w1] **Navigating Generative AI Disclosure, Ownership, and Accountability in Co-Creative Domains**
Hyo Jin Do, Molly Q Feldman, Jessica He, Angel Hsing-Chi Hwang, and Seyun Kim
Adjunct Proceedings of the Annual Symposium on Human-Computer Interaction for Work (CHIWORK), 2025

BOOK CHAPTER

[b1] **Computational Methods for Socio-Computer Interaction**
Wai-Tat Fu, Mingkun Gao, and **Hyo Jin Do**
Oxford University Press, 2018

DEMO, WORKSHOP, AND SHORT PAPERS

[d1] **Synthetic Data for Evaluation: Supporting LLM-as-a-Judge Workflows with EvalAssist**
Martín Santillán Cooper, Zahra Ashktorab, **Hyo Jin Do**, Erik Miehling, Werner Geyer, Jasmina Gajcin, Elizabeth M Daly, Qian Pan, and Michael Desmond
Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP): System Demonstrations, 2025

[d2] **Understanding Industry Practitioners' Experiences in Generative AI Governance**
Hyo Jin Do, Swati Babbar, Wenjing Li, Laura Walks, and Shayenna Misko
Proceedings of the Extended Abstracts of the CHI Conference on Human Factors in Computing Systems, 2025

[d3] **Exploring Industry Practices and Perspectives on AI Attribution in Co-Creative Use Cases**
Jessica He and **Hyo Jin Do**
Joint Proceedings of the ACM Intelligent User Interfaces (IUI) Workshops, 2025

[d4] **Facilitating Human-LLM Collaboration through Factuality Scores and Source Attributions**
Hyo Jin Do, Rachel Ostrand, Justin D. Weisz, Casey Dugan, Prasanna Sattigeri, Dennis Wei, Keerthiram Murugesan, and Werner Geyer
Trust and Reliance in Evolving Human-AI Workflows (TREW) Workshop at CHI, 2024

[d5] **Empathic Virtual Assistant for Healthcare Information with Positive Emotional Experience**
Hyo Jin Do and Wai-Tat Fu
Proceedings of the IEEE International Conference on Healthcare Informatics (ICHI), 2016

[d6] **Analyzing Emotions in Twitter During a Crisis: A Case Study of the 2015 Middle East Respiratory Syndrome Outbreak in Korea**
Hyo Jin Do, Chae-Gyun Lim, You Jin Kim, and Ho-Jin Choi
Proceedings of the International Conference on Big Data and Smart Computing (BigComp), 2016

[d7] **Korean Twitter Emotion Classification Using Automatically Built Emotion Lexicons and Fine-Grained Features**
Hyo Jin Do and Ho-Jin Choi
Proceedings of the Pacific Asia Conference on Language, Information and Computation: Poster, 2015

US PATENTS

[p1] Guiding a user to interact with an intelligent computing system using best practices, 2025
[p2] Computer-based interactive prompt variation generator, 2025