Hye Sun Yun

Computer Science PhD Candidate

Boston, MA — USA

✉ yun.hy@northeastern.com | 🏠 www.hyesunyun.com

Education

Northeastern University   *Boston, MA*

PhD in Computer Science   *September 2020 - Present*
- Advisors: Byron Wallace and Timothy Bickmore
- Expected Graduation: Spring 2026

Wellesley College   *Wellesley, MA*

BA in Computer Science and Africana Studies   *September 2013 - May 2017*
- Magna Cum Laude

Research Experience

Northeastern University, NLP Research Group   *Boston, MA*

Research Assistant   *May 2021 - Present*
- Conducted research in the field of NLP and biomedical/clinical text advised by Byron Wallace.
- Worked on NLP multi-document summarization research to improve and update systematic reviews of biomedical literature to help medical professionals make better clinical and policy decisions.
- Explored ways to utilize large generative language models to assist producing medical systematic reviews through user surveys and to automate meta-analyses.

Northeastern University, Relational Agents Group   *Boston, MA*

Research Assistant   *September 2020 - Present*
- Conducted research in the field of HCI and healthcare advised by Timothy Bickmore.
- Assisted in developing a research mHealth app with a conversational virtual agent to promote healthy behaviors for historically Black church members by exploring ways to increase engagement and use through notifications and storytelling.
- Developed a conversational agent system to enhance patient-reported outcomes for individual use.
- Conducted research on using large language models to generate clinical depression survey item variants and testing their validity and reliability.
- Designed and carried out cognitive interviews to better understand how non-medical individuals understand and feel about medical illustrations.

Truveta   *Bellevue, WA*

ML/AI PhD Research Intern   *May 2024 - August 2024*
- Conducted research in the field of LLMs and clinical notes investigating quality and diversity of training data for clinical concept extraction.
- Explored methods to optimize a generalized clinical concept extraction LLM.

Publication

Conferences

Framing Health Information: The Impact of Search Methods and Source Types on User Trust and Satisfaction in the Age of LLMs
Hye Sun Yun, Timothy Bickmore
Extended Abstracts of the CHI Conference on Human Factors in Computing Systems, 2025

Automatically Extracting Numerical Results from Randomized Controlled Trials with LLMs
Hye Sun Yun, David Pogrebitskiy, Iain J Marshall, Byron C Wallace
Machine Learning for Healthcare Conference, 2024

Keeping Users Engaged During Repeated Interviews by a Virtual Agent: Using LLMs to Reliably Diversify Questions

Hye Sun Yun, Mehdi Arjmand, Phillip Sherlock, Michael K Paasche-Orlow, James W Griffith, Timothy Bickmore
24th ACM International Conference on Intelligent Virtual Agents, *2024*

Investigating User Perceptions of Collaborative Agenda Setting in Virtual Health Counseling Session
Mina Fallah, Farnaz Nouraei, **Hye Sun Yun**, Timothy Bickmore
24th ACM International Conference on Intelligent Virtual Agents, *2024*

Appraising the Potential Uses and Harms of LLMs for Medical Systematic Reviews
Hye Sun Yun, Iain Marshall, Thomas Trikalinos, Byron Wallace
Empirical Methods in Natural Language Processing (EMNLP), *2023*

Improving Multiparty Interactions with a Robot Using Large Language Models
Prasanth Murali, Ian Steenstra, **Hye Sun Yun**, Ameneh Shamekhi, Timothy Bickmore
Extended Abstracts of the CHI Conference on Human Factors in Computing Systems, *2023*

Motivating Health Behavior Change with a Storytelling Virtual Agent
Hye Sun Yun, Matias Volonte, Timothy Bickmore
22th ACM International Conference on Intelligent Virtual Agents, *2022*

Techno-Spiritual Engagement: Mechanisms for Improving Uptake of mHealth Apps Designed for Church Members
Hye Sun Yun, Shou Zhou, Everlyne Kimani, Stefan Olafsson, Teresa K. O'Leary, Dhaval Parmar, Jessica A. Hoffman, Stephen S. Intille, Michael K. Paasche-Orlow, Timothy W. Bickmore
Conference on Intelligent User Interfaces Workshop: HEALTHI, *2022*

Journals

Online Health Information Seeking in the Era of LLMs: Cross-Sectional Web-based Survey Study
Hye Sun Yun, Timothy Bickmore
Journal of Medical Internet Research. *2025*

In a pilot study, automated real-time systematic review updates were feasible, accurate, and work-saving
Iain J. Marshall, Thomas A. Trikalinos, Frank Soboczenski, **Hye Sun Yun**, Gregory Kell, Rachel Marshall, Byron C. Wallace
Journal of Clinical Epidemiology *pp. 26–33. 2023*

Media

Church after Sunday: supporting everyday well-being through techno-spiritual health interventions
Teresa K O'Leary, Elizabeth Stowell, Darley Sackitey, **Hye Sun Yun**, David Wright, Michael Paasche-Orlow, Timothy Bickmore, Andrea G. Parker
Interactions *pp. 90–93. 2022*

Preprints

Caught in the Web of Words: Do LLMs Fall for Spin in Medical Literature?
Hye Sun Yun, Karen Y. C. Zhang, Ramez Kouzy, Iain J. Marshall, Junyi Jessy Li, Byron C. Wallace
arXiv. *2025*

Professional Experience

Wayfair *Boston, MA*
Full-stack Software Engineer/Engineering Manager *July 2017 - August 2020*
- Developed and implemented features for a supplier-facing web tool, enabling efficient bulk addition of new products to the catalog.
- Successfully spearheaded two major, multi-phase refactoring initiatives: frontend migration from Tungsten to React and Redux and completely rewriting the team's Excel sheet import service.
- Provided guidance and mentorship during rapid team growth, facilitating the onboarding of 10+ new hires.
- Led team-wide adoption of pair programming to foster collaboration and knowledge sharing and established a program to train new code reviewers.
- Managed and led a team of three software engineers, cultivating a collaborative and productive work environment, and ensuring successful project delivery.

Presentations

Invited Talks

Fall 2024. *Charting My CS Career Path: From Wellesley to Industry and PhD.*
    Invited talk: Wellesley Hackathon (WHACK), Wellesley College

Fall 2024. *Automatically Extracting Numerical Results from RCTs with LLMs.*
    Invited talk: Codametrix, Boston, MA

Fall 2022. *Motivating Health Behavior Change with a Storytelling Virtual Agent.*
    Guest Speaker: CS3950 Intro to CS Research, Northeastern University

Research Presentations

Fall 2024. *Keeping Users Engaged During Repeated Interviews by a Virtual Agent: Using LLMs to Reliably Diversify Questions.*
    Poster: IVA, Glasgow, Scotland.

Summer 2024. *Automatically Extracting Numerical Results from Randomized Controlled Trials with LLMs.*
    Poster: Machine Learning for Healthcare, Toronto, Canada.

Spring 2023. *Using Large Language Models to Reliably Diversify Questions.*
    Poster: Cutting-Edge Connections in PhD Research: Healthcare Innovation Today, Northeastern University.

Spring 2023. *Appraising the Potential Uses and Harms of LLMs for Medical Systematic Reviews.*
    Poster: CRA-WP Grad Cohort for Women, San Francisco, CA.

Spring 2022. *Motivating Health Behavior Change with a Storytelling Virtual Agent.*
    Departmental Poster Session: Khoury College PhD Open House, Northeastern University.

Fall 2022. *Motivating Health Behavior Change with a Storytelling Virtual Agent.*
    Poster: ACM IVA, Faro, Portugal.

Spring 2022. *Techno-spiritual Engagement: Mechanisms for Improving Uptake of mHealth Apps Designed for Church Members.*
    Oral presentation: ACM IUI HEALTHI Workshop, Remote.

Fall 2021. *Longitudinal Analysis of the Attitudes and Conversations Surrounding Vaccines by African Americans on Social Media.*
    Oral presentation: PRISM Health Symposium, San Francisco, CA.

Teaching Experience

Spring 2025	CS4983/6983 Research in Human-Centered NLP, **Instructor of Record** @ Northeastern University
Fall 2024	CS3950 Intro to CS Research, **Teaching Assistant** @ Northeastern University
2019-2020	MVC framework class, **Teaching Assistant** @ Wellesley College
Sep.-Oct. 2019	Quality Assurance to Quality Engineer, **Program Instructor** @ Wayfair
Spring 2016	CS231 Algorithms, **Peer Tutor** and Grader @ Wellesley College
Fall 2016	CS320 Tangible User Interface, **Peer Tutor** @ Wellesley College
Fall 2016	CS235 Languages & Automata, **Peer Tutor** @ Wellesley College
2014-2015	JPN101,102 Beginning Japanese, **Peer Tutor** @ Wellesley College

Mentoring

2025	Karen Y.C. Zhang, BS @ Northeastern University
2024	Patrick Wilson, Master in Public Health Dissertation @ King's College London
2024	David Pogrebitskiy, BS @ Northeastern University
2022	Nikita Salkar, MS @ Northeastern University

Service & Development

Outreach & Volunteering

Mar. 2025	Khoury SHEROs of Color's SHEHack, **GitHub Workshop Instructor & Mentor**
Feb. 2025	Khoury SHEROs of Color's PhD Panel, **Invited Panelist**
Mar. 2024	Northeastern University, Khoury College PhD Open House, **PhD Student Buddy**
Nov. 2023	Queer in AI - Graduate School Application Review Program, **Mentor**
Aug. 2022	Northeastern University, Pathway to PhD Workshop, **PhD Student Panelist**
July 2022	Northeastern University, Wellesley College Research Students' Visit, **PhD Student Panelist**
Jan. 2022	Northeastern University, Khoury College, **Faculty Application Student Reviewer**
Mar. 2021	Northeastern University, Khoury College PhD Open House, **PhD Student Panelist**
Jan. 2021	Northeastern University, Khoury College, **PhD Student Application Reviewer**
2019-2020	Wayfair's Women+ in Software, **Co-founder & Leader**

Development

2024	Northeastern University, National Research Mentoring Network's Research Mentor Training
2023	Computing Resarch Association Widening Participation, Grad Cohort for Women
2022	Association for Computational Linguistics (ACL), Year-Round Mentorship

Peer Review

Conference on Human Factors in Computing Systems (CHI)

Conference on Language Modeling (COLM)

Journal of Medical Internet Research (JMIR)

Machine Learning for Healthcare (MLHC)

ACM Conference on Intelligent Virtual Agents (IVA)

Conference on Autonomous Agents and Multiagent Systems (AAMAS)

IEEE International Symposium on Mixed and Augmented Reality (ISMAR)

IEEE Virtual Reality (IEEE VR)

Workshop on Robustness in Sequence Modeling @ NeurIPS 2022

Human-Centered Large Language Modeling Workshop @ ACL 2024

UserNLP: User-centered Natural Language Processing Workshop @ ACM Web Conference 2022

Cochrane Evidence Synthesis and Methods

Research Synthesis Methods

Skills

Tools/Platforms	Git, HuggingFace, Pytorch, GCP, Docker, Weights & Biases, Prolific, Amazon MTurk
Libraries/Frameworks	Pytorch, HuggingFace, Node.js, REST API, Flask, .NET Core
Front-end	React, Redux, Vue.js, HTML5, SASS, Jekyll
Programming	Python, Java, Javascript, C#, SQL, PHP, LaTeX, Shell, R, C, Racket, Assembly
Databases	MySQL, SQLite, PostgreSQL, MS SQL Server
Languages	English, Korean, Japanese (Basic), Swahili (Basic)