

Masum Hasan, PhD

AI Research Scientist

LLMs, Alignment, Human-AI Interaction

 <https://masumhasan.com>

 masumhasan.phd@gmail.com

 Scholar

 [masumhasan6](#)

RESEARCH FOCUS

I design and evaluate large language models that behave more human-like in high-stakes interactions. My work spans LLM alignment (DPO/RLHF), human-AI evaluation, synthetic data generation, and applied ML for healthcare and code intelligence, with deployments in virtual agents and communication training systems.

EDUCATION

University of Rochester

Ph.D. Computer Science

Rochester, NY

Aug 2021 - May 2026

University of Rochester

M.Sc. Computer Science

Rochester, NY

Aug 2021 - Aug 2023

Bangladesh University of Engineering and Technology (BUET)

B.Sc. Computer Science & Engineering

Dhaka, Bangladesh

Feb 2013 - Oct 2018

PROFESSIONAL EXPERIENCE

Independent AI Consultant

Help organizations build and deploy AI strategy & solutions

Rochester, NY

June 2026 - Present

University of Rochester

Graduate Research Assistant, [ROC-HCI Lab](#), Supervised by Prof. [Ehsan Hoque](#)

Rochester, NY

Aug 2021 - May 2026

Research Contributions

- Designed DPO-based alignment methods to induce subtle conversational traits such as human-likeness in 1B–72B LLMs, outperforming OSS and commercial models in large-scale Chatbot Arena-style human evaluations [1]. LLM benchmark to judge human-likeness [2].
- Developed LLM-driven virtual humans for medical communication, role-play, and training, with human-centered design and evaluation [3,4,5,6].
- Built a Unity-based synthetic data pipeline for 2D hand pose estimation, achieving competitive performance with *zero* human annotation [8].
- Developed healthcare ML systems, including video-based ataxia risk prediction [7] and a safety-constrained Parkinson’s education chatbot [6].

Leadership & Execution

- Led and mentored 10+ researchers across 5 projects from ideation to system building, human evaluation, and publication; two undergraduate mentees received outstanding researcher awards.
- Coordinated HCI, ML, healthcare, and clinical collaborations while maintaining IRB and HIPAA compliance.

Bangladesh University of Engineering and Technology (BUET)

Research Assistant, Applied Machine Learning Lab

Supervised by Prof. [Rifat Shahriyar](#) and Prof. [Anindya Iqbal](#)

Dhaka, Bangladesh

Nov 2018 - Jun 2021

Research Contributions

- Conducted NLP-for-code research on source code summarization, refactoring, synthesis, and semantic search, contributing to multiple peer-reviewed publications [10,11,12,13,14].
- Developed Bengali-GPT-2 and state-of-the-art Bengali–English MT systems, including low-resource MT for the endangered Chakma language [15,16,17,18].

Leadership & Execution: As an early Applied ML Lab member, led teams of 9 researchers and helped establish one of Bangladesh’s first academic–industrial ML research labs.

LLMs, Alignment, and Human-Like AI

- [1] [HAL: Inducing Human-likeness in LLMs with Alignment](#) - Masum Hasan, Junjie Zhao, Ehsan Hoque (Under review, EMNLP 2026)
- [2] [Inverse Turing Bench: Evaluating Language Models as Judges of Human vs. AI Dialogue](#) - William Hager, Ishika Rathi, Masum Hasan, Cameron Jones (Under review, EMNLP 2026)
- [3] [SAPIEN: Affective Virtual Agents Powered by Large Language Models](#) – Masum Hasan, Cengiz Ozel, Sammy Potter, Ehsan Hoque. (ACIIW 2023)
- [4] [Poker with Play Money: Exploring Psychotherapist Training with Virtual Patients](#) - Cynthia Baseman, Masum Hasan, Nathaniel Swinger, Sheila Rauch, Ehsan Hoque, Rosa Arriaga (CSCW 2025)

Human-Centered AI and Healthcare

- [5] [AI Standardized Patient Improves Human Conversations in Advanced Cancer Care](#) – Kurtis Haut Masum Hasan, Thomas Carroll, Ronald Epstein, Taylan Sen, Ehsan Hoque (JMIR 2026)
- [6] [A User-Centered Framework to Empower People with Parkinson’s Disease](#) – Wasifur Rahman, Abdelrahman Abdelkader, Sangwu Lee, Phillip Yang, Md Saiful Islam, Tariq Adnan, Masum Hasan, Ellen Wagner, Sooyong Park, E. Ray Dorsey, Catherine Schwartz, Karen Jaffe, Ehsan Hoque. (IMWUT, UbiComp 2024)
- [7] [Auto-Gait: Automatic Ataxia Risk Assessment with Computer Vision on Gait Task Videos](#) – Wasifur Rahman, Masum Hasan, Md Saiful Islam, Titilayo Olubajo, Jeet Thaker, Abdelrahman Abdelkader, Phillip Yang, Tetsuo Ashizawa, Ehsan Hoque. (UbiComp 2023)

Synthetic Data and Computer Vision

- [8] [Hi5 🖐️: Synthetic Data for Inclusive, Robust, Hand Pose Estimation](#) - Masum Hasan, Cengiz Ozel, Nina Long, Alexander Martin, Samuel Potter, Tariq Adnan, Sangwu Lee, Amir Zadeh, Ehsan Hoque (ACII 2025)
- [9] [Hitting your MARQ: Multimodal ARGument Quality Assessment](#) – Md Kamrul Hasan, James Spann, Masum Hasan, Md. Saiful Islam, Kurtis Haut, Rada Mihalcea, Ehsan Hoque. (EMNLP 2021)

NLP for Code and Software Engineering

- [10] [CoDesc: A Large Code–Description Parallel Dataset](#) – Masum Hasan, Tanveer Muttaqeen, Ishtiaq Niloy, Kazi Mehrab, Tahmid Hasan, Mahim Pantho, Wasi Ahmad, Rifat Shahriyar, Anindya Iqbal. (ACL Findings 2021)
- [11] [Text2App: A Framework for Creating Android Apps from Text Descriptions](#) – Masum Hasan, Kazi Sajeed Mehrab, Wasi Uddin Ahmad, Rifat Shahriyar. (NLP4Prog workshop, ACL 2021)
- [12] [Review4Repair: Code Review Aided Automatic Program Repairing](#) – Faria Huq, Masum Hasan, Mahim Pantho, Sazan Mahbub, Anindya Iqbal, Toufique Ahmed. (IST Journal)
- [13] [Using a Balanced Scorecard to Identify Opportunities to Improve Code Review Effectiveness: An Industrial Experience Report](#) – Masum Hasan, Anindya Iqbal, Amiangshu Bosu, Mohammad Rafid Ul Islam, A.J.M. Imtiajur Rahman. (EMSE Journal)
- [14] [BERT2Code: Can Pretrained Language Models be Leveraged for Code Search?](#) – Abdullah Al Ishtiaq, Masum Hasan, Md. Mahim Anjum Haque, Kazi Sajeed Mehrab, Tanveer Muttaqeen, Tahmid Hasan, Anindya Iqbal, Rifat Shahriyar. (Preprint)

Low-Resource NLP

- [15] [ChakmaNMT: A Low-resource Machine Translation On Chakma Language](#) - Aunabil Chakma, Aditya Chakma, Soham Khisa, Chumui Tripura, Masum Hasan, Rifat Shahriyar. (Under review, ACL 2026)
- [16] [LowResource at BLP-2023 Task 2: Leveraging BanglaBert for Low Resource Sentiment Analysis of Bangla Language](#) – Aunabil Chakma, Masum Hasan. (EMNLP 2023, BLP Workshop)
- [17] [Not Low-Resource Anymore: Aligner Ensembling, Batch Filtering, and New Datasets for Bengali-English Machine Translation](#) – Tahmid Hasan, Abhik Bhattacharjee, Kazi Samin, Masum Hasan, Madhusudan Basak, M. Sohel Rahman, Rifat Shahriyar. (EMNLP 2020)

- [18] [Recognition of Bengali Handwritten Digits Using Convolutional Neural Network Architectures](#) – Md Mahmudul Hasan, Md Rafid Ul Islam, Md Tareq Mahmood. (ICBSLP 2018, **Best Student paper award**)

OPEN-SOURCE CONTRIBUTIONS

- **TuringBench-2** — A Benchmark to evaluate LLMs ability to tell humans apart — hf.co/TuringBench-2
- **Human-Aligned LLM (HAL)** – Model weights, dataset, and training code – roc-hci.github.io/HAL
- **Text2App (2020)** – Mobile app from text – github.com/Masum06/Text2App
- **CoDesc** – Code-Description parallel dataset – github.com/Masum06/CoDesc

INVITED TALKS

- **University of Rochester (ROC Talks)** — “Rethinking Capitalism in the Age of AGI,” April 2024. [Video](#)
- **New York Academy of Sciences** — “Future of Interactive Education with Generative AI,” February 2024.
- **University of Rochester** — “Privacy, Security, and Ethics of Generative AI,” February 2024.
- **WXXI News (NPR Affiliate)** — Radio appearance on “Connections with Evan Dawson” discussing the future of human-AI interaction, December 2023.
- **FlowerCity.AI Conference** — Invited speaker, December 2023.
- **University of Rochester** — Speaker, 4-part workshop series on Generative AI, Fall 2023.
- **University of Rochester** — “Introduction to ChatGPT” workshop, Spring 2023.

TECHNICAL SKILLS

Core ML: LLM alignment (DPO, RLHF), transformers, multimodal models, diffusion, ViT, Machine Translation, LLM-Agents

Evaluation & Data: Large-scale human evaluation, synthetic data generation, annotation pipelines

Systems & Tooling: Python, PyTorch, HuggingFace, AWS, LangGraph, CrewAI, SQL, Java, C/C++