Bhavya **Chopra**

EECS PhD Student, University of California, Berkeley

bhavyac16.github.io | bhavyachopra@berkeley.edu | Google Scholar | github.com/BhavyaC16 | Berkeley, CA, USA

Education

Present – August 2024 | **University of California, Berkeley** — **Berkeley, CA, USA**
PhD in Electrical Engineering and Computer Science
Advisor: Prof. Aditya Parameswaran
GPA: 4.000/4.000

June 2022 – August 2018 | **Indraprastha Institute of Information Technology Delhi (IIIT-Delhi)** — **New Delhi, India**
Bachelor of Technology *(with Honors)*, Computer Science and Design
GPA: 9.14/10.00

May 2018 – May 2017 | **Delhi Public School, R.K. Puram** — **New Delhi, India**
CBSE AISSCE Class XII
Aggregate: 95.6/100.0

Research Interests

My research interests lie at the intersection of Artificial Intelligence, Data Management, and Human-Computer Interaction. I design and build mixed-initiative data systems that combine human insight with LLM capabilities to help end-users transform and interpret data at scale.

Work Experience

May 2025 – August 2025 | **Research Intern, Tableau Research** — **Seattle, WA, USA**
Advisor: Dr. Arjun Srinivasan
- Developed Forge, a tool for AI-assisted ad-hoc tabular data derivation, supporting generation of new columns based on formulae, semantic data processing, and information retrieval.

July 2024 – July 2022 | **Research Fellow, PROSE Team at Microsoft** — **Bengaluru, India**
Advisors: Dr. Austin Z. Henley, Dr. Gustavo Soares
- Conducted research with developers, designed conversational debugging for GitHub Copilot Chat.
- Designed improved user experience for Flash Fill, shipped with Data Wrangler.
- Developed dataset generation pipelines for Responsible AI evaluation of Excel Copilot.

July 2021 – May 2021 | **Software Development Intern, MathWorks** — **Bengaluru, India**
- Developed *MatSight*, a visual analysis tool for the UX Team; won **Best Hack** & **Most Voted Hack** prizes.
- Refactored & increased functional test coverage for MATLAB (R2021b) Support Package for Raspberry Pi.

August 2019 – June 2019 | **Software Development Intern, IIIT-Delhi** — **New Delhi, India**
- Designed, developed, & hosted TechTree for the academic department and faculty to maintain and share descriptions of over 350 courses with students.
- Helps 3000+ students every semester in planning courses through visualization of course pre-requisites via interactive tables and directed acyclic graphs.

Publications

C=Conference, S=In Submission, D=Demo, W=Workshop, T=Thesis, *=Equal Contribution

[C.7] **Why Do Multi-Agent LLM Systems Fail?** [🔗]
Mert Cemri*, Melissa Z. Pan*, Shuyi Yang*, Lakshya A Agrawal, Bhavya Chopra, Rishabh Tiwari, Kurt Keutzer, Aditya G. Parameswaran, Dan Klein, Kannan Ramchandran, Matei Zaharia, Joseph E. Gonzalez, Ion Stoica
NeurIPS Datasets and Benchmarks Track
☆ *Spotlight* — [**NeurIPS 2025**]

[C.6] **Steering Semantic Data Processing with DocWrangler** [🔗]
Shreya Shankar*, Bhavya Chopra*, Mawil Hasan, Stephen Lee, Björn Hartmann, Joseph M. Hellerstein, Aditya G. Parameswaran, Eugene Wu
ACM Symposium on User Interface Software and Technology
Best Paper Honorable Mention — [**UIST 2025**]

[C.5] **Rethinking Dataset Discovery with DataScout** [🔗]
Rachel Lin*, Bhavya Chopra*, Wenjing Lin, Shreya Shankar, Madelon Hulsebos, Aditya G. Parameswaran
ACM Symposium on User Interface Software and Technology **[UIST 2025]**

[C.4] **Let's Fix this Together: Conversational Debugging with GitHub Copilot** [🔗]
Yasharth Bajpai, Bhavya Chopra, Param Biyani, Cagri Aslan, Dustin Coleman, Sumit Gulwani, Chris Parnin, Arjun Radhakrishna, Gustavo Soares
IEEE Symposium on Visual Languages and Human-Centered Computing
🏆 *Best Paper Award* **[VL/HCC 2024]**

[C.3] **Detangler: Helping Data Scientists Explore, Understand, and Debug Data Wrangling Pipelines** [🔗]
Nischal Shrestha, Bhavya Chopra, Austin Z. Henley, Chris Parnin
IEEE Symposium on Visual Languages and Human-Centered Computing
🏆 *Best Paper Award* **[VL/HCC 2023]**

[C.2] **StoryBox: Independent Multi-modal Interactive Storytelling for Children with Visual Impairment** [🔗]
Bhavya Chopra, Richa Singh
Late-Breaking Work, ACM SIGCHI Conference on Human Factors in Computing Systems **[CHI 2022]**

[C.1] **Reality Tales: Facilitating User-Character Interaction with Immersive Storytelling** [🔗]
Bhavya Chopra*, Khushali Verma*, Sonali Singhal*, Utsav Singla*
Student Research Competition, ACM SIGCHI Conference on Human Factors in Computing Systems
🏅 *Second Position at CHI Student Research Competition (Undergraduate Category)* **[CHI 2021]**

[W.4] **Why Do Multi-Agent LLM Systems Fail?** [🔗]
Mert Cemri*, Melissa Z. Pan*, Lakshya A Agrawal*, Shuyi Yang, Bhavya Chopra, Rishabh Tiwari, Kurt Keutzer, Aditya G. Parameswaran, Dan Klein, Kannan Ramchandran, Matei Zaharia, Joseph E. Gonzalez, Ion Stoica
Building Trust in Language Models and Applications Workshop at ICLR 2025 **[ICLR 2025]**

[W.3] **Challenges in using Conversational AI for Data Science** [🔗]
Bhavya Chopra, Ananya Singha, Anna Fariha, Sumit Gulwani, Chris Parnin, Ashish Tiwari, Austin Z. Henley
Workshop on Human-In-the-Loop Data Analytics, SIGMOD 2025 **[SIGMOD 2025]**

[W.2] **Exploring Interaction Patterns for Debugging: Enhancing Conversational Capabilities of AI-assistants** [🔗]
Bhavya Chopra*, Yasharth Bajpai*, Param Biyani, Gustavo Soares, Arjun Radhakrishna, Chris Parnin, Sumit Gulwani
3rd Workshop on Bridging Human-Computer Interaction and Natural Language Processing, NAACL 2024 **[NAACL 2024]**

[W.1] **Semantically Aligned Question and Code Generation for Automated Insight Generation** [🔗]
Ananya Singha, Bhavya Chopra, Anirudh Khatry, Sumit Gulwani, Austin Z. Henley, Vu Le, Chris Parnin, Mukul Singh, Gust Verbruggen
LLM4Code Workshop 2024, IEEE/ACM International Conference on Software Engineering
🏆 *Best Paper Award* **[ICSE 2024]**

[D.1] **CoWrangler: Recommender System for Data Wrangling Scripts** [🔗]
Bhavya Chopra, Anna Fariha, Sumit Gulwani, Austin Z. Henley, Daniel Perelman, Mohammad Raza, Sherry Shi, Danny Simmons, Ashish Tiwari
Demo Track, ACM International Conference on Management of Data **[SIGMOD 2023]**

[T.1] **Study of Assertions: Understanding Assertion Use in Java Projects on GitHub** [🔗]
Bhavya Chopra
Bachelor Thesis, IIIT-Delhi, 2022

Selected Software Projects

Quill: Runtime C++ Work-Stealing Library February 2021

Developed a light-weight thread pool based work-stealing runtime for async-finish task-parallelism in C++ to support flat-finish scopes using the Pthread library. (Technologies Used: **C++, Pthread Library**)

Plants versus Zombies: Strategy Video Game October 2019 – January 2020

Re-implemented Plants vs Zombies, following object oriented programming principles. The project is well received in the open source community, with widespread attention from enthusiasts in various online forums. (Technologies Used: **Java, JavaFX**) (GitHub, Video)

LiveErr0r: Real-time feedback for elementary learners January 2019 – February 2019

Designed and developed an interactive system to check the correctness of handwritten mathematical expressions in real time, providing instant haptic feedback to elementary learners. (Technologies Used: **Python, Android, Arduino**)

Honours and Awards

Best Paper Honorable Mention, UIST 2025 [🔗] Awarded by ACM Symposium on User Interface Software and Technology for the paper titled *"Steering Semantic Data Processing with DocWrangler"*

Best Paper Award, VL/HCC 2024 [🔗] Awarded by IEEE Symposium on Visual Languages and Human-Centered Computing for the paper titled *"Let's Fix this Together: Conversational Debugging with GitHub Copilot"*

Best Paper Award, LLM4Code 2024 [🔗] Awarded by IEEE/ACM International Conference on Software Engineering for the paper titled *"Semantically Aligned Question and Code Generation for Automated Insight Generation"*

Best Paper Award, VL/HCC 2023 [🔗] Awarded by IEEE Symposium on Visual Languages and Human-Centered Computing for the paper titled *"Detangler: Helping Data Scientists Explore, Understand, and Debug Data Wrangling Pipelines"*

Second Position, Student Research Competition at CHI 2021 [🔗] Awarded by ACM Student Research Competition in the undergraduate category for the paper titled *"Reality Tales: Facilitating User-Character Interaction with Immersive Storytelling"*

Dean's List Award for Excellence in Research (2020–21) and (2021–22) Awarded for 2 consecutive years by Dean of Innovation, Research & Development (DIRD), IIIT-Delhi for research on inclusive storytelling experiences for children

Dean's List Award for Excellence in Academics (2020–21) Awarded by Dean of Academic Affairs (DoAA), IIIT-Delhi

GHC Student Scholarship 2021 Awarded by AnitaB.org for attending the 2021 Virtual Grace Hopper Celebration

Ishwar Chandra Memorial Award, DPS R.K. Puram, 2017 [🔗] Awarded by Ms. Joanna Kempkers, the High Commissioner of New Zealand, for excellent academic performance

Academic Research and TA Positions

Graduate Student Researcher, EPIC Data Lab January 2025—Present

Founding Member and Undergraduate Researcher, Accessibility & Inclusive Design Lab August 2021—June 2022

Undergraduate Researcher, Program Analysis Group (PAG) November 2020—May 2022

Teaching Assistant, Introduction to Human-Computer Interaction January 2022—May 2022

Undergraduate TA for freshman year course Introduction to HCI (DES102) at IIIT-Delhi, taught by Prof. Rajiv Ratn Shah

Teaching Assistant, Program Analysis August 2021—December 2021

Undergraduate TA for graduate level course Program Analysis (CSE503) at IIIT-Delhi, taught by Prof. Rahul Purandare

Academic Service

Reviewer CHI 2025, UIST 2025, CHI 2024 Late-Breaking Work, CHI 2023, CHI 2022 Late-Breaking Work
Volunteer CSCW 2022, CHI PLAY 2022, CSCW 2021, IndiaHCI 2021

Leadership Positions

Delhi Chapter Head, Women in Machine Learning and Data Science (WiMLDS) March 2021—July 2024

Organised the chapter's 4-week Mentorship Programs and Knowledge Series talks on data science, ML and AI

President, Women in Tech, IIIT-Delhi June 2020—June 2022

Delivered talks on HCI research & web-development, organized hackathons for diversity, and deployed the club website

Skills

Programming Languages Python, Java, C++, C#, TypeScript, JavaScript, F*, MATLAB
Tools and Technologies Git, shell scripting, React, openai-python, FastAPI, Flask, PySpark, pandas, Hadoop, Heroku, Spoon, OpenGL, Habanero C, OpenMP, MPI, JavaFX, Figma, Miro, LaTeX
Relevant Coursework **Computer Science:** Data Structures, Analysis & Design of Algorithms, Database Management Systems, Operating Systems, Computer Networks, Machine Learning, Big Data Analytics*, Program Verification*, Decision Procedures*, Parallel Programming*, Computer Graphics*, Artificial Intelligence Systems*, Data-Centric LLMs*
Human-Computer Interaction: Human Computer Interaction, Engineering Design, Visual Design, Design Processes, Advanced Topics in Human Centered Computing*, Design of Interactive Systems*, Inclusive Design and Accessibility*, Building Usable Programming Tools*

(* — *Graduate Level Courses*)