

Participant Information Sheet for Exempt Research

Northeastern University, Khoury College of Computer Sciences

Name of Investigator(s): Jonathan Bell, Gwenyth Lincroft, Emma Bortz, Avery Yen

Title of Project: Understanding User Experiences with Mechanistic Interpretability

Funded by: The National Science Foundation, IIS-2408455

Version date: 11/3/2025

We are inviting you to participate in a research study. Participating is voluntary; you do not have to participate if you do not want to. You can withdraw from the study at any time.

The purpose of this study is to experiences and pain points (both existing and future) with, perceptions of, and use cases and research projects involving mechanistic interpretability (MI) techniques that current MI, interdisciplinary, and other domain researchers have. Participating in this research study will include one interview discussing your research and your experience with AI and interpretability. The interview will take about 15 to 30 minutes to complete.

You can skip questions that you do not want to answer or stop the interview at any time.

We are asking for your permission to audio record the interview. If you agree to being recorded, please let us know.

Your part in this study will be confidential. Only the researchers on this study will see the information about you. Personal identifiers will not be published or presented. Participation or non-participation in this study will have no effect on your ability to participate in any class or on your standing in any class. If you have any questions about this study, please contact Gwenyth Lincroft (lincroft.g@northeastern.edu) or Emma Bortz (e.bortz@northeastern.edu), the people mainly responsible for the research. You can also contact Jonathan Bell (j.bell@northeastern.edu), the Principal Investigator, or Avery Yen (yen.av@northeastern.edu).

If you have any questions about your rights in this research, you can contact the Northeastern University Department of Human Research at Tel: (773) 396-2327, or Email: IRBReview@northeastern.edu . You may call anonymously if you want.