A WEB-BASED WIZARD-OF-OZ PLATFORM FOR COLLABORATIVE AND REPRODUCIBLE HUMAN-ROBOT INTERACTION RESEARCH

by

Sean O'Connor

A Proposal Submitted to the Honors Council

For Honors in the Department of Computer Science

February 5, 2026

Approved: _____

L. Felipe Perrone
Thesis Adviser

_____

Brian King
Second Reader

1   Introduction

To build the social robots of tomorrow, researchers must find ways to convincingly simulate them today. The process of designing and optimizing interactions between human and robot is essential to the Human-Robot Interaction (HRI) field, a discipline dedicated to ensuring these technologies are safe, effective, and accepted by the public. Yet, conducting rigorous research in social robotics remains hindered by complex technical requirements and inconsistent methodologies.

In a typical social robotics interaction, a robot operates autonomously based on pre-programmed behaviors. However, human interaction can be unpredictable. When a robot fails to respond appropriately to a social cue, the interaction can degrade, causing the human partner to lose trust or disengage.

To overcome the limitations of pre-programmed autonomy, researchers often use the Wizard-of-Oz (WoZ) technique to test prototypes of robot behaviors before the underlying technology is fully developed. In this method, a human operator (the "wizard") observes the interaction from a separate room via cameras and microphones, controlling the robot's actions in real-time. To the person interacting with the robot, it appears fully autonomous, creating a convincing simulation that is helpful for rapid prototyping and testing of interaction designs.

Despite its conceptual simplicity, conducting WoZ research presents two challenges. The first is a technical barrier that prevents many non-programmers, such as experts in psychology or sociology, from conducting their own studies. This accessibility problem is compounded by a second challenge: a fragmented hardware landscape. Because different labs use different robot platforms, researchers often must build their own custom control tools for each study. These bespoke systems are rarely shared, making it difficult for scientists to

replicate and build upon each other's findings, which hinders the development of a reliable and verifiable body of knowledge.

To address these challenges, I am developing HRIStudio, a web-based platform for designing, executing, and analyzing WoZ experiments in social robotics. I argue that by lowering technical barriers and providing a common experimental platform, a web-based framework can significantly improve both the disciplinary accessibility and scientific reproducibility of research in social robotics.

2   Context

The challenges of disciplinary accessibility and scientific reproducibility in WoZ research have been explored in HRI literature. In a foundational systematic review of 54 HRI studies, Riek [7] discovered a widespread lack of methodological consistency, noting that very few researchers reported standardized wizard training or measurement of wizard error. This stems from a landscape of specialized, "in-house" systems, where individual labs develop their own custom software for each study, tools that are rarely shared with other researchers. This forces labs to constantly reinvent control interfaces, hindering the replication and verification of scientific findings.

In response, the research community has developed several specialized WoZ platforms. A first wave of tools focused on creating powerful, flexible architectures. Polonius was designed as a robust interface for robotics engineers to create experiments for their non-programmer collaborators, featuring an integrated logging system to streamline data analysis [2]. Similarly, OpenWoZ introduced an adaptable framework that used web protocols to allow different control interfaces to easily connect to the robot, empowering technical users to create deviations from the pre-programmed interaction scripts in real-time [1]. While

architecturally sophisticated, these tools still required significant technical expertise to set up and configure, keeping the accessibility barrier high.

A second wave of tools shifted focus to prioritize usability for a broader audience. WoZ4U was explicitly designed to be an "easy-to-use tool for the Pepper robot" that makes it easier for "non-technical researchers to conduct Wizard-of-Oz experiments" [8]. WoZ4U successfully lowered the accessibility barrier with an intuitive graphical interface. However, this usability was achieved by tightly coupling the software to a single type of robot. This approach creates a significant risk to platform longevity. As Pettersson and Wik note in a review of generic WoZ tools, systems that are too specialized often fall out of use as hardware becomes obsolete [5]. This trade-off between capability, usability, and sustainability reveals a critical gap in the literature. No available tool exists that is simultaneously flexible, accessible, and can endure over time.

In response to this lack of an adequate tool, I designed HRIStudio by combining an intuitive web-based interface with a flexible architecture that allows it to support a wide range of current and future robots. The result is a single, sustainable platform that is both powerful enough for complex experiments and accessible enough for interdisciplinary research teams.

3   Description

I created HRIStudio as an integrated, web-based platform designed to manage the entire lifecycle of a WoZ experiment in social robotics: from interaction design, through live execution, to final analysis. I designed the platform around three core principles: making research accessible to non-programmers, ensuring the experiments are reproducible, and providing a time-enduring tool for the HRI community.

To solve the challenge of accessibility, I provide researchers with tools to visually map out an experiment's flow, much like creating a storyboard for a film. This intuitive approach allows a social scientist, for example, to design a complex HRI without writing a single line of code. The platform provides different interfaces to facilitate collaboration between the members of a team: A researcher gets a design canvas to build the study, the wizard gets a streamlined control panel to run the experiment, and an observer gets a tool for taking timestamped notes.

To enable experiment reproducibility, I designed HRIStudio to mitigate key methodological challenges inherent in WoZ research. The first challenge is inconsistent wizard behavior; a tired or distracted human operator can unintentionally introduce errors, compromising a study's validity. HRIStudio's wizard interface acts as a "smart co-pilot," guiding the operator through the pre-designed script with clear prompts for what to do and say next. This minimizes human error and increases the likelihood of a standardized experience for every participant. The second challenge lies in the complex task of managing experimental data. A typical study generates multiple streams of data that are difficult to synchronize manually, including video, audio, robot sensor logs, and wizard actions. The platform acts as a central recorder, automatically capturing and timestamping every data stream into a single, unified timeline. This simplifies analysis and allows another researcher to "replay" the entire experiment to verify and analyze the findings of another's study.

Finally, to ensure the platform will be a time-enduring tool for the community, I designed the system to be robot-agnostic. Rather than being constrained to operate with a single kind of robot, the platform uses a system of standardized "connectors," like a universal remote programmable for any television. This flexible architecture ensures that the platform will remain a valuable tool for the community long after any specific robot becomes obsolete, providing a stable, lasting foundation for future research.

4   Significance

This work is significant because it accelerates the foundational research needed to deploy social robots in critical societal roles, such as providing companionship for the elderly in assisted living facilities or acting as classroom aides for children with autism. My tool directly enables and accelerates the rigorous, human-centered research on which the success and public acceptance of these technologies depend.

The primary significance of HRIStudio is its potential to lower the barrier to entry for HRI research. By allowing for visual programming, the platform removes technical barriers that have traditionally limited this research to engineering disciplines. It invites the domain experts who should be leading these studies to design and execute their own experiments, leading to better research questions and effective robot behaviors.

My goal with HRIStudio is to elevate the scientific rigor of the field. By promoting a common structure for designing experiments and support for data collection, HRIStudio allows researchers to more easily replicate, verify, and build upon each other's work. This work supports the ongoing effort to make HRI a more cumulative science, where findings can be more easily verified and built upon by other researchers.

Ultimately, my work contributes a piece of critical, open-source infrastructure to the HRI community that directly addresses the documented challenges of accessibility, reproducibility, and sustainability. Beyond its immediate utility, the platform's architecture also serves as a tangible blueprint for web-based scientific tools, demonstrating a successful model for bridging the gap between an intuitive user interface and the complexity of controlling live robotic hardware.

The foundational concepts of this work have already been reported in two peer-reviewed publications at the IEEE International Conference on Robot and Human Interactive Com-

munication [3, 4]. This work represents the culmination of that research, delivering the platform's full implementation, a critical evaluation by real users, and its release as a tool for the community.

5   Independent Contribution

This work builds upon a foundational collaboration with my adviser that led to two publications and the initial high-level design of the HRIStudio platform. For this work, my primary intellectual contribution is the independent execution of the project; I am the sole developer responsible for the complete software implementation and for the design and execution of the user study.

6   Methods

The foundational concepts and early architecture of HRIStudio have been established in prior work [3, 4]. The primary goal of this work is to translate that foundation into a complete, stable, and usable platform, and then rigorously evaluate its success. Therefore, the work is divided into two key phases: first, the final implementation of the platform's core features as outlined in the project timeline, and second, a formal user study to validate its impact on experimental consistency and efficiency.

The study will involve recruiting approximately 10-12 participants from non-engineering fields (e.g., Psychology, Education) who have experience designing experiments but little to no programming background. The core task will be to recreate a well-documented experiment from the HRI literature using the NAO6 robot. To ensure a level playing field, all participants will first attend a workshop on the software package they are assigned. The participants will

be divided into two groups: a control group will use the manufacturer-provided Choregraphe software [6], and an experimental group will use HRIStudio.

My evaluation will focus on two primary outcomes. The first is methodological consistency: I will quantitatively assess the accuracy of each group's recreated experiment by comparing their final implementation against the original study's protocol. This will involve a detailed scoring rubric that measures discrepancies in robot behaviors, trigger logic, and dialogue. The second outcome is user experience: after the task, participants will complete a survey to provide qualitative and quantitative feedback on their assigned software. This mixed-methods approach will provide robust evidence to assess HRIStudio's effectiveness in making HRI research more accessible and reproducible.

A detailed project schedule, outlining all key milestones and deadlines, is provided in Appendix A.

7 Conclusion

This work addresses a significant bottleneck in HRI research. By creating HRIStudio, a web-based platform for Wizard-of-Oz experimentation, this work confronts the interconnected challenges of disciplinary accessibility and scientific reproducibility. The platform provides publicly available infrastructure that empowers non-technical domain experts to conduct rigorous HRI studies. Ultimately a common, accessible, and sustainable tool does more than just simplify experiments. It fosters a more collaborative and scientifically robust approach to the entire field of HRI.

References

[1] Guy Hoffman and Cynthia Breazeal. OpenWoZ: A Runtime-Configurable Wizard-of-Oz Framework for Human-Robot Interaction. *Proceedings of the 11th ACM/IEEE International Conference on Human-Robot Interaction*, pages 117–124, 2016.

[2] David V. Lu and William D. Smart. Polonius: A Wizard of Oz Interface for HRI Experiments. *Proceedings of the 6th International Conference on Human-Robot Interaction*, pages 77–78, 2011.

[3] Sean O'Connor and L. Felipe Perrone. HRIStudio: A Framework for Wizard-of-Oz Experiments in Human-Robot Interaction Studies (Late Breaking Report). 2024 33rd IEEE International Conference on Robot and Human Interactive Communication (RO-MAN), 2024.

[4] Sean O'Connor and L. Felipe Perrone. A Web-Based Wizard-of-Oz Platform for Collaborative and Reproducible Human-Robot Interaction Research. 2025 34th IEEE International Conference on Robot and Human Interactive Communication (RO-MAN), 2025.

[5] Pettersson, John Sören and Wik, Malin. The longevity of general purpose Wizard-of-Oz tools. In *Proceedings of the Annual Meeting of the Australian Special Interest Group for Computer Human Interaction*, OzCHI '15, page 422–426, New York, NY, USA, 2015. Association for Computing Machinery.

[6] E. Pot, J. Monceaux, R. Gelin, and B. Maisonnier. Choregraphe: a graphical tool for humanoid robot programming. In *RO-MAN 2009 - The 18th IEEE International Symposium on Robot and Human Interactive Communication*, pages 46–51, 2009.

[7] Laurel D. Riek. Wizard of Oz studies in HRI: a systematic review and new reporting guidelines. *J. Hum.-Robot Interact.*, 1(1):119–136, July 2012.

[8] Frank Rietz and Maren Bennewitz. WoZ4U: An Open-Source Wizard-of-Oz Interface for Human-Robot Interaction. *Proceedings of the 16th ACM/IEEE International Conference on Human-Robot Interaction*, pages 95–103, 2021.

Appendix A: Project Timeline

Timeframe	Milestones & Key Tasks
Fall 2025: Development and Preparation	
September	Finalize and submit this proposal (Due: Sept. 20). Submit IRB application for the user study.
Oct – Nov	Complete final implementation of core HRIStudio features. Conduct extensive testing and bug-fixing to ensure platform stability.
December	Finalize all user study materials (consent forms, protocols, etc.). Begin recruiting participants.
Spring 2026: Execution, Analysis, and Writing	
Jan – Feb	Upon receiving IRB approval, conduct all user study sessions.
March	Analyze all data from the user study. Draft Results and Discussion sections. Submit "Intent to Defend" form (Due: March 1).
April	Submit completed thesis draft to the defense committee (Due: April 1). Prepare for and complete the oral defense (Due: April 20).
May	Incorporate feedback from the defense committee. Submit the final, approved thesis by the university deadline.