



# HUMAN-COMPUTER INTERACTION LAB



## 42<sup>nd</sup> Annual Symposium

**THURSDAY, MAY 29, 2025**

.....  
**PRESENTATION SUMMARIES  
AND OTHER INFORMATION**



## **Human-Computer Interaction Lab (HCIL) University of Maryland, College Park**

---

The Human-Computer Interaction Lab (HCIL), launched in 1983 at the University of Maryland, has a rich history of transforming the experience people have with new technologies. From understanding user needs to developing and evaluating these technologies, the lab's faculty, staff, and students have been leading HCI research for more than 40 years.

The HCIL was established as an interdisciplinary effort within the University of Maryland's Institute for Advanced Computer Studies (UMIACS). Today, HCIL participants include faculty, staff, and students from the following units on campus: INFO College, Computer Science, Psychology, Education, English, Engineering, Journalism, and American Studies. In addition, HCIL faculty are members in a number of centers and institutes on campus, including UMIACS, the Maryland Institute for Technology in the Humanities (MITH), the Applied Research Laboratory for Intelligence and Security (ARLIS), the Maryland Cybersecurity Center (MC2), the Social Data Science Center (SoDa), the Institute for Trustworthy AI in Law & Society (TRAILS), and the Dingman Center for Entrepreneurship.

This booklet contains Symposium presentation summaries and selected highlights of our faculty and students' news and accomplishments over the past year.

Please visit our website (<https://hcil.umd.edu>) for more information about the research happening in the HCIL.

If you would like more information, please contact:

Human-Computer Interaction Lab  
2117 Hornbake Building, South Wing  
4130 Campus Drive  
College Park, Maryland 20742  
[hcil-info@umd.edu](mailto:hcil-info@umd.edu)

## HCIL Director's Letter

HCIL's 42<sup>nd</sup> Annual Symposium | May 2025



I'm writing this letter as the spring semester wraps up after a very long few months. For many researchers in the US – including those doing HCI and HCI-adjacent work – 2025 has been filled with stress and uncertainty. Some of our HCIL colleagues have lost federal funding for their grants due to connections to DEI, misinformation, or other important topics. International students and faculty traveling for conferences take on additional risks when traveling out of country to share research findings. And academia itself is facing numerous threats from the current administration as it tries to control who we admit, what we teach, and what we research.

These are grim times for researchers, and it is more important than ever that we in the HCI community support each other as we continue to do important work. Recent events in the government and in developments related to AI/LLMs have highlighted the importance of **humans** in technology use and adoption, and we are at the forefront of evaluating usability and accessibility of technology as well as potential risks and harms arising from the (mis)use of technology. We are building new tools to assist people communicate and interact with the world in meaningful ways. And we work closely with partner communities to improve people's lives.

HCIL's faculty and students are doing important work across a wide range of topics, and today we celebrate work done in the last year. Many HCIL members have recently returned from sharing their work at the ACM Conference on Human Factors in Computing Systems (CHI), held this year in Yokohama, Japan. UMD had an impressive showing this year; by my count, 45 students and faculty presented 20 papers, as well as contributions to SIGs, panels, workshops, interactivity sessions, late-breaking work, and AltCHI submissions. Notably, HCIL authors are included on two best paper awards (top 1% of papers) and two honorable mentions (top 5% of papers). An impressive showing indeed!

This year we also welcomed several new faculty to campus, including Mohit Iyyer, Fumeng Yang, and Gabe Kaptchuk. At the same time, we wish a very fond farewell and happy retirement to Carol Boston, who has been a cornerstone of the INFO College for many years, a regular attendee to HCIL events, and most recently supporting Sheena Erete's work.

Today you'll hear from many of our students and faculty on a wide range of research topics. In total, there will be three plenary talks, one panel, and 37 lightning talks spanning human-centered AI, accessibility, usable privacy and security, ubiquitous computing, and much more.

There is much to celebrate this year, and I encourage you to check out the back of the program, which includes a list of some of the awards earned by HCIL faculty and students. I'd also like to call out a few extra-noteworthy items. First, a big congrats to Joel Chan, HCIL's Associate Director, who was recently promoted to Associate Professor and awarded tenure. I also want to congratulate other faculty who were recently promoted, including Naeemul Hassan (Associate Professor with tenure) and Tammy Clegg (full professor). Promotion in academia is a very long and involved process, and promotions are a testament to the research impact they have in HCI. I also wish a hearty congratulations and best wishes to the graduating PhD and master's students – some of them are listed in the back of the program.

It takes a village to run this symposium, so I want to offer my gratitude to those who have supported us, including through financial support (VEX Robotics, funding agencies, campus units), as well as the facilities team in the School of Public Policy, who have helped us set up this space after Iribe became unavailable, and the events and communications teams in the iSchool. Special thanks goes to Paris Lane, the HCIL Coordinator, who has been helping me for the last two years while pursuing her master's degree in Games, Entertainment, and Media Analytics (GEM) – a degree she'll have received shortly before the symposium!

Finally, I want to thank the students and faculty who make up the HCIL. You inspire me every day with your curiosity, ingenuity, and perseverance. I'm excited to share with you some of the amazing work they've been up to over the last year.

Welcome to the 42<sup>nd</sup> annual HCIL symposium!

A handwritten signature in black ink, reading "Jessica Vitak". The signature is fluid and cursive, with the first name "Jessica" and last name "Vitak" clearly legible.

Jessica Vitak  
Director, Human-Computer Interaction Lab (HCIL)  
Professor, College of Information  
University of Maryland, College Park





# 42<sup>nd</sup> Annual Human-Computer Interaction Lab (HCIL) Symposium

May 29, 2025

Thurgood Marshall Hall (TMH-0301)

University of Maryland

*All times listed in Eastern Time (ET)*

## Detailed Program

Time	Details
9:00am	<b>Registration Opens (atrium)</b> <ul style="list-style-type: none"><li>• Check in and grab your name tag and some swag. Check out lunch options (themed tables, trivia), and get caffeinated.</li></ul>
9:30am	<b>Welcome Address</b> <ul style="list-style-type: none"><li>• Jessica Vitak, Director, HCIL</li></ul>
9:50am	<b>Plenary Talks</b> <ul style="list-style-type: none"><li>• <b>Huaishu Peng, Assistant Professor, CS:</b> Small (Computational) Artifacts for Augmented (Human) Ability</li><li>• <b>Hawra Rabaan, Postdoctoral Scholar, iSchool:</b> Incorporating Social Justice Theories into Design: Introducing Survivor-Centered Transformative Justice and Healing Structures</li><li>• <b>Michelle Mazurek, Associate Professor, CS:</b> Revisiting what it means to be Usable: Human-Centered Security Beyond End Users</li></ul>
10:50am	<b>Lightning Talks: Session 1</b> <ul style="list-style-type: none"><li>• 12 short talks on topics spanning human-centered AI and accessibility</li></ul>
12:30pm	<b>Lunch</b> <p>Grab a drink and lunch and chat at one of our themed tables or outside (weather permitting).</p>
1:30pm	<b>Panel</b> <ul style="list-style-type: none"><li>• <b>Topic:</b> Peer Review Hereafter: Purpose, Science, and HCI</li><li>• <b>Participants:</b> Jay Patel (moderator), Melinda Baldwin, AJ Boston, Raj Paletti, Lu Sun, and Ronen Tamari</li></ul>



# 42<sup>nd</sup> Annual Human-Computer Interaction Lab (HCIL) Symposium

May 29, 2025

Thurgood Marshall Hall (TMH-0301)

University of Maryland

*All times listed in Eastern Time (ET)*

## Detailed Program (continued)

Time	Details
<b>2:10pm</b>	<b>Lighting Talks: Session 2</b>
	<ul style="list-style-type: none"><li>• 11 short talks on topics spanning ubiquitous computing and usable privacy &amp; security</li></ul>
<b>3:35pm</b>	<b>Coffee Break</b>
	<ul style="list-style-type: none"><li>• Coffee, tea, and light snacks</li></ul>
<b>4:00pm</b>	<b>Lightning Talks: Session 3</b>
	<ul style="list-style-type: none"><li>• 14 short talks on topics spanning XR, making, &amp; design and search, sensemaking, and learning</li></ul>
<b>5:45pm</b>	<b>Reception (atrium)</b>
	<ul style="list-style-type: none"><li>• Stay for cake, hors d'oeuvres, and conversation!</li><li>• Check out student posters</li></ul>

## **Sponsorships**

The research that happens in the lab—and the lab itself—cannot happen without financial support, and we are grateful to our corporate, campus, and research sponsors, especially those who have contributed to HCIL and its members in the last year.

### **Industry Sponsor:**

VEX Robotics

### **Campus Sponsors:**

College of Information

Philip Merrill College of Journalism

Computer Science Department

College of Computer, Mathematical, and Natural Sciences

Applied Research Laboratory for Intelligence and Security (ARLIS)

University of Maryland Institute for Advanced Computer Studies (UMIACS)

### **Research Sponsors:**

Army Research Office

Chan-Zuckerberg Initiative

Google Research

Meta

National Aeronautics and Space Administration (NASA)

National Institutes of Health (NIH)

National Science Foundation (NSF)

Navigation Fund

Sloan Foundation

Teaching and Learning Transformation Center (TLTC)

University of Maryland

US Defense Advanced Research Projects Agency (DARPA)

US Department of Defense (DoD)

US Food & Drug Administration (FDA / M-CERSI)



**PLENARY SPEAKERS**  
**&**  
**PANEL DESCRIPTION**



# Human-Computer Interaction Lab

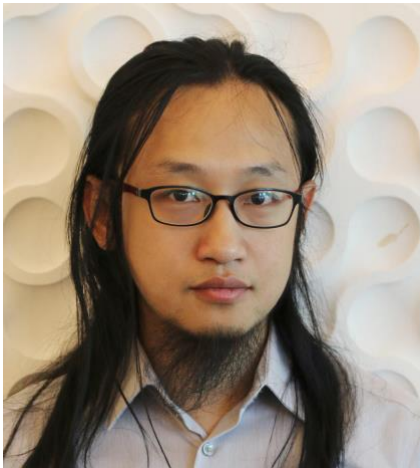
## 42<sup>nd</sup> Annual Symposium



### Plenary Speaker: Huaishu Peng, Assistant Professor, CS Department

#### Talk Title: Small (Computational) Artifacts for Augmented (Human) Ability

**Abstract:** In this talk, I will discuss a set of our recent and ongoing work from the Small Artifacts Lab on human-computer interaction. Specifically, I would like to draw connections between small, personal, and physical devices and a new frontier of embodied AI, presenting three research prototypes: a camera-equipped, motorized measuring tape that can help blind and low-vision people use inaccessible public kiosks; a set of tabletop tangible bricks for blind users to design the graphics and layout of a personal website; and a small, movable wearable that can quickly move across one's body, presumably sensing the host's activities. I will discuss the following questions: How can these embodied artifacts augment human ability? What are the benefits and risks of extending digital computation to a physical form that is personal or interpersonal (i.e., beyond industrial context)? What are the research opportunities in interdisciplinary research across multiple fields?



**Bio:** Huaishu Peng is an Assistant Professor in the Department of Computer Science at the University of Maryland, College Park. He aims to advance interactive technologies by designing, prototyping, and evaluating novel tangible artifacts that are personal, hands-on, and often compact in form factor. He is interested in the methods for building these personal artifacts (e.g., through design and interactive fabrication), the scenarios in which they are used (e.g., in mixed reality), and the users who can benefit from them (e.g., through assistive and enabling technologies). His work has been frequently published in prestigious venues such as CHI, UIST, IMWUT, and SIGGRAPH, and has received several best paper and demo awards. Additionally, his research has been featured in media outlets including Wired, MIT Technology Review, TechCrunch, and Gizmodo.





## Plenary Speaker: Hawra Rabaan, Postdoctoral Scholar, College of Information

### Talk Title: Incorporating Social Justice Theories into Design: Introducing Survivor-Centered Transformative Justice and Healing Structures

**Abstract:** While domestic violence (DV) affects all socioeconomic groups, identity deeply shapes how individuals experience and recover from abuse. Marginalized survivors often face significant barriers when accessing government and community resources, which frequently overlook their autonomy, needs, and choices. Using the experiences of US-based Muslim women as a case study, this work derives a survivor-centered transformative justice (SCTJ) framework which discerns individual and systemic, supports designing *alongside* victim-survivors, and centers their autonomy. Also, it introduces healing structures that promote justice by addressing harmful practices, discriminatory laws, and enabling collective and survivor-centered interventions. We outline three interconnected sociotechnical processes that constitute these healing structures when implemented collectively and simultaneously:

1. **Survivor advocacy**, which ensures safety, empowerment, informed decision-making, and continuity of care.
2. **Community accountability**, which fosters reparation and behavioral change of abusers through collective intervention and support.
3. **Institutional transformation**, which reshapes the broader systemic conditions that perpetuate abuse through structural reform and prevention.

This work contributes to HCI theoretically, empirically, and through design by centering those most affected by harm while addressing the community and structural dimensions that sustain violence—particularly in underrepresented and underserved communities.



**Bio:** Hawra Rabaan is a Postdoctoral Researcher at the Human-Computer Interaction (HCI) Lab in the College of Information at the University of Maryland. Her work is grounded in feminist and justice-oriented frameworks, focusing on community-driven AI education, research, and design. She critically examines the intersection of technology, gender, and social justice, addressing systemic inequities. She is the first to introduce Islamic feminism as a theoretical and analytical lens in HCI, expanding conversations on gender, power, and technology design beyond Western paradigms. She developed the *Survivor-Centered Transformative Justice* approach, a framework for designing alongside victim-survivors of domestic violence, emphasizing autonomy, collective accountability, and systemic transformation.

Rabaan's research has been recognized internationally, with publications in top-tier venues such as the Conference on Human Factors in Computing Systems (CHI) and Computer-Supported Cooperative Work & Social Computing (CSCW), earning an honorable mention award. She has also contributed to *Interactions* magazine through the Association for Computing Machinery (ACM), bridging feminist perspectives with computing and design research. She holds a BS in Information Technology, an MS in Information and Communication Technologies for Development (ICTD), and a PhD in Human-Computer Interaction (HCI) with a minor in Social Work.



# Human-Computer Interaction Lab

## 42<sup>nd</sup> Annual Symposium



**Plenary Speaker: Michelle Mazurek, Associate Professor, CS Department**

**Talk Title: Revisiting what it means to be Usable: Human-Centered Security Beyond End Users**

**Abstract:** The human-centered security community has made significant progress in making security and privacy tools, notifications, and warnings more legible and usable for end users. However, many critical security and privacy problems remain out of the hands of end users, or -- even when simplified -- require more knowledge, time, or effort to manage than is reasonable or fair to expect from most users. As such, the next important challenge in human-centered security is to go beyond end users and explore how to make security and privacy more usable for the professionals whose decisions directly or indirectly affect end users at larger scale. These professionals include not only software developers, vulnerability analysts, and security operations personnel, but also social scientists who publish research data, product reviewers, and even YouTube influencers. In this talk, I will briefly highlight recent research from my lab addressing some of these constituencies.



**Bio:** Michelle Mazurek is an Associate Professor in the Computer Science Department and the Institute for Advanced Computer Studies at the University of Maryland, College Park, where she also directs the Maryland Cybersecurity Center. Her research aims to understand and support the human elements of digital security- and privacy-decision making. Recent projects include examining how and why developers make security and privacy mistakes; examining how security and privacy information is distributed via YouTube influencers; and analyzing how users learn about and decide whether to adopt security advice. Dr. Mazurek has served as Program Chair for the Symposium on Usable Privacy and Security (SOUPS) and the Privacy Enhancing Technologies Symposium (PETS). She has received a number of awards, including the NSF CAREER award, DARPA Young Faculty Award, the NSA's Best Scientific Cybersecurity Paper award, and several distinguished paper awards. She received her PhD in Electrical and Computer Engineering from Carnegie Mellon University.



# Human-Computer Interaction Lab

## 42<sup>nd</sup> Annual Symposium



## PANEL: Peer Review Hereafter: Purpose, Science, and HCI

Jay Patel, Joel Chan

University of Maryland – College Park (INFO)

patel93@hey.com, joelchan@umd.edu

"The current "crisis" of peer review arguably has its origins in this moment in the 1970s, when the process was cast as **the only acceptable method of evaluating scientific quality**. The more we have expected of peer review, the more its opportunities to disappoint have expanded." – Melinda Baldwin

"...publication is the **beginning of peer review**" – Richard McElreath

### ABSTRACT

Whether visible or not, peer review is more *nuanced than supposed* and *metamorphosing faster* than in past generations. Formal peer review, as practiced within journal and conference ecosystems, emerged earliest and solidified as a mechanism to triage manuscripts for overwhelmed editors. Today, scholarly publishers are experimenting with novel work models and AI tools to enhance its fairness and efficiency. Informal peer review, which we call the practice of critically evaluating scholarly reports in public and digital communities, is more recent. It is the inverse of formal peer review in that it is more emergent, underground, and inclusive. The most vocal of these reviewers, sometimes called *sleuths*, act as checks and balances on the credibility of formal peer review by surfacing errors, fraud, and questionable decisions in studies. Tools and funding schemes to support informal peer review are proliferating, extending the time scale along which reviewing operates.

In this panel, we will discuss the *purpose* of peer review variants according to history, metascience, and HCI. We will also discuss the meaning of *peer*, the role of *automation*, and opportunities to overcome the fixation on evaluative *reviews* as the sole desirable imprints of scholarly discourse. Our discussion will surface *scientific perspectives from researchers* and solutions by *tool-builders* who are envisioning a future for research evaluation that is untethered to our inherited scholarly infrastructures.

### MODERATOR

Jay Patel (Ph.D. Candidate, INFO) studies research evaluation in formal and informal settings and evidence synthesis.

### PANELISTS

**Melinda Baldwin** is Associate Professor of History and the AIP Endowed Professor in History of Natural Sciences at UMD. Her expertise is in the history of science with emphasis on scientific publishing, grant-making, and peer review. She is the author of *Making "Nature": The History of a Scientific Journal* and her upcoming book documents the history of peer review.

**Arthur J (AJ) Boston** is Scholarly Communication Librarian and Associate Professor at Murray State University (MSU) Libraries. There, he coordinates the Office of Research and Creative Activity, manages the institutional repository, and teaches. He blogs about open science, peer review, and scholarly communications at [aj-boston.pubpub.org](http://aj-boston.pubpub.org). His posts are exquisite.

**Raj Paletti** is co-founder of the preprint discussion platform alphaXiv. His past research applied deep learning to predict heart attacks and supported assistive teleoperation of robots. He is an MS student at Stanford University studying CS and AI.

**Lu Sun** is a Ph.D. Candidate in CS at UC San Diego and applies machine learning methods to scaffold academic peer review. She has developed and presented prototypes like ReviewFlow for assisting peer reviewers and MetaWriter for assisting editors.

**Ronen Tamari** is a researcher and entrepreneur building a distributed intelligence system, *Cosmik*, to aid scholars in curating, evaluating, and making sense of scholarly research together. He completed his Ph.D. in CS at Hebrew University of Jerusalem.



# **LIGHTNING TALKS: ABSTRACTS**

## **[talks listed in order of presentation]**

### **Clusters:**

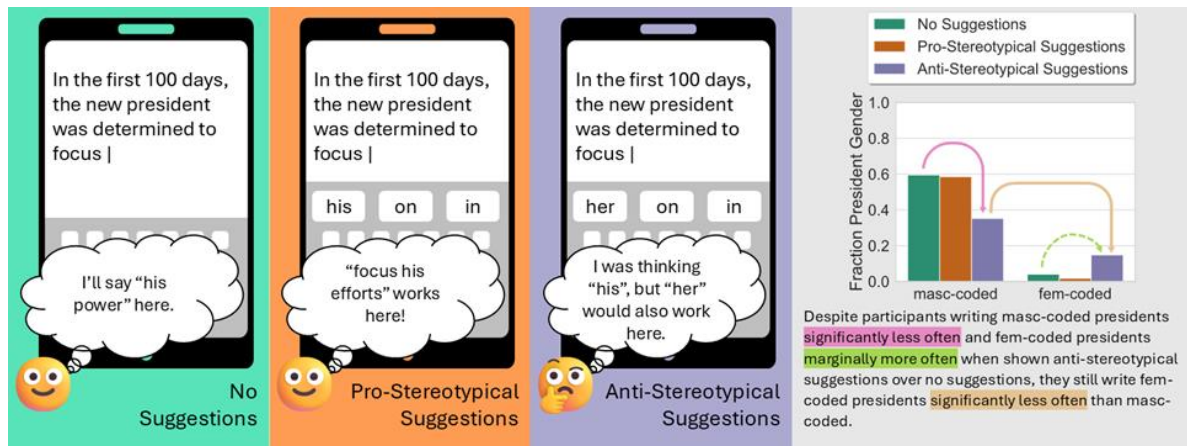
- Human-Centered AI
- Accessibility
- Usable Privacy & Security
- Ubiquitous Computing
- XR, Making, & Design
- Search, Sensemaking, and Learning

# Anti-stereotypical Predictive Text Suggestions Do Not Reliably Yield Anti-stereotypical Writing

Connor Baumler and Hal Daumé III

University of Maryland

Contact: [baumler@umd.edu](mailto:baumler@umd.edu)



**Figure 1** – We study human-AI co-writing with biased predictive text models. In settings like the one pictured, we find that though anti-stereotypical suggestions significantly decrease the amount of pro-stereotypical stories written, this is not enough to remove (let alone reverse) the pro-stereotypical bias in the co-written stories.

AI-based systems such as language models can replicate and amplify social biases reflected in their training data. Among other questionable behavior, this can lead to LM-generated text—and text suggestions—that contain normatively inappropriate stereotypical associations. Prior work has found that co-writing with a language model can affect content—for example, influencing sentiment in reviews. In this work, we consider the question of how “debiasing” a language model impacts stories that people write using that language model in a predictive text scenario.

In a pre-registered and IRB-approved online study ( $n = 414$ ), we asked participants to write short English stories with (treatment condition) or without (control condition) the help of a predictive text system. In the treatment condition, when the participants were provided with text predictions, these predictions were generated—on a per-story basis—by either a language model that was designed to make suggestions that aligned with social stereotypes, or one that was designed to challenge social stereotypes surrounding gender and sexuality. These suggestions represent extreme cases where the model *only* suggests pro-stereotypical or anti-stereotypical continuations—instead of being “realistically” biased or “debiased.”

We find that, in certain scenarios, language model suggestions that align with common gender and sexuality stereotypes are more likely to be accepted by human authors. Conversely, although anti-stereotypical language model suggestions sometimes lead to an increased rate of anti-stereotypical stories, this influence is far from sufficient to lead to parity in the gender and sexuality-related character traits expressed in stories.

For example, as shown in Figure 1, when the predictive text model *only* suggests fem-coded language to describe a president character, participants write significantly fewer masc-coded presidents and marginally more fem-coded presidents than in the control condition (with no suggestions). However, even with exclusively fem-coded suggestions, participants still produce significantly more masc-coded presidents overall. This suggests that if a developer's goal is to achieve gender parity in stories co-written by humans and AI, providing an anti-stereotypical predictive text model—let alone a “debiased” one—may be insufficient to fully counteract existing biases in human writing.





# (Dis)placed Contributions: Uncovering Hidden Hurdles to Collaborative Writing Involving Non-Native Speakers, Native Speakers, and AI-Powered Editing Tools

Yimin Xiao<sup>1</sup>, Yuewen Chen<sup>1</sup>, Naomi Yamashita<sup>2</sup>, Yuexi Chen<sup>1</sup>, Zhicheng Liu<sup>1</sup>, Ge Gao<sup>1</sup>

<sup>1</sup>University of Maryland, <sup>2</sup>Kyoto University

Contact: [yxiao@umd.edu](mailto:yxiao@umd.edu)

## OVERVIEW

Content creation today often takes place via collaborative writing. A longstanding interest of CSCW research lies in understanding and promoting the coordination between co-writers. However, little attention has been paid to individuals who write in their non-native language and to co-writer groups involving them. While language diversity in a group may enhance joint content production, writing and expressing ideas in a non-native language can be challenging, especially for people with limited proficiency.

To enhance collaborative writing involving non-native speakers, we explore coordination mechanisms that encourage non-native speakers' ideational input while alleviating their concerns about expressing their ideas. Specifically, we considered workflow structure as a key coordination component: how would non-native speakers contribute to collaborative writing when they act before vs. after native speakers? We adopted Goffman's theory of footing to inform our assessment of each co-writer's contribution along several aspects of language production: lexical, ideational, and expressional. We also considered how AI-powered language tools mediate non-native speakers' participation in collaborative tasks.

## METHOD

We conducted a mixed-method study to explore how people write in a non-native language in collaboration with native speakers. Our participants included 32 co-writer groups, each consisting of one native speaker of English and one non-native speaker with limited proficiency. They performed collaborative writing through two different workflows: half began with non-native speakers taking the first editing turn and the other half began with native speakers. We collected each group's writing history to measure their writing process and interviewed and surveyed participants to understand their task experiences.

## KEY RESULTS

**Non-native speakers' contribution to ideational aspects of collaborative writing.** Non-native speakers in our study clearly separated expressional and ideational aspects of

writing. They tried to offer concrete ideas as their contributions, despite challenges to express them.

**A late-mover disadvantage exclusively experienced by non-native speakers.** Non-native speakers' ideational contributions were reduced when they acted after their native speaker partner compared to before. When acting as a late mover, they found it difficult to identify space for making contributions without messing up the content that had been edited by their partners.

**The unexpected effects of using AI tools on the collaboration process.** Non-native speakers' usage of AI-powered editing tools required substantial effort and left them with limited time to focus on the writing task. It also led native speakers to overestimate non-native speakers' language capability and agency displayed in the writing. This caused unexpected tension in their collaboration.

**Main takeaways from this study.** In collaborative writing involving non-native speakers, fair interpretation of a person's contribution must consider 1) each person's inputs beyond the lexical level and 2) the interpersonal effects of using AI-powered editing tools. Collaborative writing workflow should be structured to facilitate coordination of ideational and expressional aspects of the writing process.

## ACKNOWLEDGEMENT

This work was funded by National Science Foundation under grant # 1947929.

## PUBLICATION

Yimin Xiao, Yuewen Chen, Naomi Yamashita, Yuexi Chen, Zhicheng Liu, and Ge Gao. 2024. (Dis)placed Contributions: Uncovering Hidden Hurdles to Collaborative Writing Involving Non-Native Speakers, Native Speakers, and AI-Powered Editing Tools. In *Proceedings of the ACM on Human-Computer Interaction*. 8 (CSCW2), 1-31. <https://doi.org/10.1145/3686942>

# Sustaining Human Agency, Attending to Its Cost: An Investigation into Generative AI Design for Non-Native Speakers' Language Use

Yimin Xiao<sup>1</sup>, Cartor Hancock<sup>1</sup>, Sweta Agrawal<sup>2</sup>, Nikita Mehandru<sup>3</sup>,  
Niloufar Salehi<sup>3</sup>, Marine Carpuat<sup>1</sup>, Ge Gao<sup>1</sup>

<sup>1</sup>University of Maryland, <sup>2</sup>Instituto de Telecomunicações, Portugal, <sup>3</sup>University of California, Berkeley

Contact: [yxiao@umd.edu](mailto:yxiao@umd.edu)

## OVERVIEW

AI systems and tools today can generate human-like expressions on behalf of people. It raises the crucial question of how to sustain human agency in AI-mediated communication. We investigated this question in the context of machine translation (MT) assisted conversations.

We adopted Bandura's social cognitive theory of agency to guide our design exploration of sustaining human agency by moderating people's resource involvement to influence MT outputs. Specifically, we designed three human-MT interfaces for non-native speakers to interact with MT that use increasing levels of language resources: Labeling, Regular Post-Editing, and Augmented Post-Editing.

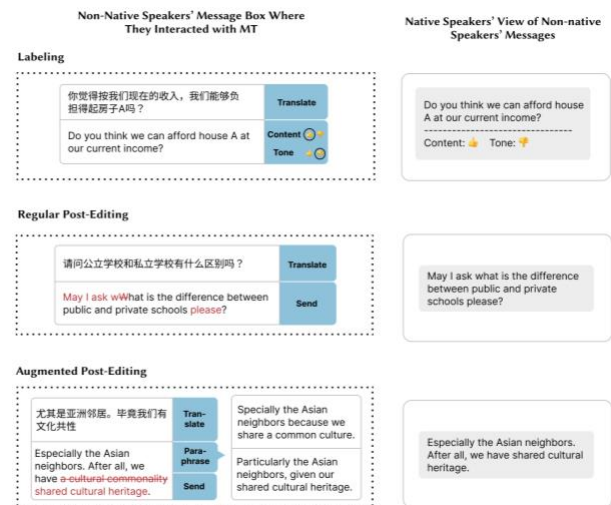
## METHOD

We conducted an online between-subject experiment to examine non-native speakers' agency and communication performance across the three interfaces.

Our participants included 45 dyads. Each dyad consisted of one new immigrant in the United States, who leveraged MT for English information seeking as a non-native speaker, and one local native speaker, who acted as the information provider. Non-native speakers could influence the English production of their message in one of three ways: labeling the quality of MT outputs, regular post-editing without additional hints, or augmented post-editing with LLM-generated hints.

## KEY RESULTS

**The resource-agency link to operationalize human agency preservation.** We found that all three human-MT interfaces sustained non-native speakers' agency to a moderate to high extent. Non-native speakers' agency was higher in the post-editing conditions, where they used more language resources than the labeling condition.



**Figure 1** – We explored three human-MT interfaces: Labeling, Regular Post-Editing, and Augmented Post-Editing. These three conditions allowed non-native speakers to leverage increasing amounts of language resources to influence MT outputs and sustain their agency.

**The costs associated with human agency.** The relatively high levels of agency non-native speakers exercised in the two post-editing conditions came at the cost of less in-depth conversation and lower communication outcome.

## ACKNOWLEDGEMENT

This work was funded by National Science Foundation under grants #2147292 and #2229885.

## PUBLICATION

Yimin Xiao, Cartor Hancock, Sweta Agrawal, Nikita Mehandru, Niloufar Salehi, Marine Carpuat, and Ge Gao. 2025. Sustaining Human Agency, Attending to Its Cost: An Investigation into Generative AI Design for Non-Native Speakers' Language Use. Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems. 1–16. <https://doi.org/10.1145/3706598.3713626>

# Training with Less: How People Select Data with Higher Value for AI

Farnaz Zamiri Zeraati, Jonggi Hong, Kyungjun Lee, Hernisa Kacorri

University of Maryland, College Park

Contact: [farnaz@umd.edu](mailto:farnaz@umd.edu)

## OVERVIEW

People are increasingly made aware of the importance of data for AI. They are often called to make conscious decisions around the use of their own photos, text, and interactions for improving models overall or fine-tuning them to their needs. Understanding the value of their data can play a critical role in these decisions. Yet, it is unclear how those who may not have machine learning expertise can tell which of their data are of value for training models. To study this, we conduct a crowdsourcing study with 100 participants and create a dataset called CrowdTeaMa, which contains 22500 images. We developed a web-based testbed for a mobile teachable object recognizer, inviting participants to train and evaluate it using three objects of their choice within a specific object category (e.g., cereal, drink, snack, spice, etc.). After testing their model, we asked them to consecutively select smaller data subsets of 20, 5, and 1 from the 30 images that they took of each object. We asked them to select the subset of photos that make the model more robust.

**Why machine teaching?:** Machine teaching aims to reduce the number of examples a teacher needs to provide to a learner. In certain contexts, this can be viewed as the task of optimal data subset selection. Aligned with our research questions, machine teaching serves as an effective proxy by providing an opportunity to gain deeper insights into how people perceive and interact with these systems, and what types of data they consider most valuable for training models.

Alongside the human-selected subsets we also implemented 3 different algorithms for data selection: Random selection as the baseline, Facility Location-based selection, and generative AI assistant selection (Copilot). By calculating image similarity metrics across these different subsets, and model's performance across different subsets, we gain insights into how humans select data with higher value for AI.

Our results highlight six unique patterns in participants' subset selection strategies, which outperform random. However, when comparing them to other computational methods for selection that employ submodular functions and

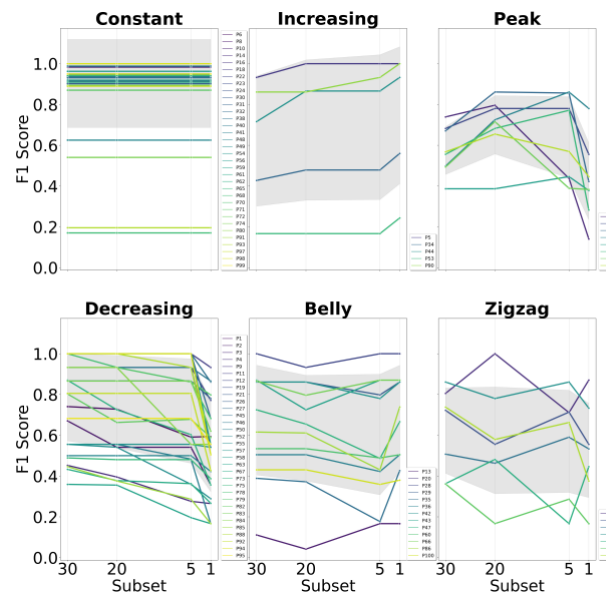


Figure 1 – F1 score trends across clusters of participants

generative AI, it isn't always a clear win. These findings demonstrate the potential of computational methods for supporting people in decisions around their data and their value.

This paper presents two main contributions. First, the primary contribution is empirical which arises from an examination of how people select subset of data that they consider of higher value for training machine learning algorithms. As another empirical contribution, we present a collection of qualitative findings about different user groups' strategies in selecting data and their comprehension (or lack of comprehension) of the data value. From these insights, we derive a set of implications for the design of Interactive Machine Learning systems tailored to novice users.

## ACKNOWLEDGEMENTS

This research was supported by the U.S. National Science Foundation (NSF).



## Characterizing How Users Audit Fairness and Bias in AI Chatbots

Jan Tolsdorf<sup>1</sup>, Alan F. Luo<sup>2</sup>, Monica Kodwani<sup>1</sup>, Junho Eum<sup>1</sup>, Mahmood Sharif<sup>3</sup>, Michelle L. Mazurek<sup>2</sup>, Adam J. Aviv<sup>1</sup>

<sup>1</sup>Department of Computer Science, The George Washington University

<sup>2</sup>Department of Computer Science, University of Maryland

<sup>3</sup>Blavatnik School of Computer Science, Tel Aviv University

Contact: jan.tolsdorf@email.gwu.edu

### OVERVIEW

The increasing role of generative AI in economic and societal contexts has placed greater demands on the research community to articulate and mitigate potential risks posed by adoption of these technologies. In the case of conversational agents, particularly chatbots, public and academic discourse typically centers on fairness, trust, and risk. However, there is limited insight into how users assess for and develop perceptions of these traits in AI chatbots.

To better understand how end users assess for and develop perceptions of fairness, trust, and risk in conversational AI chatbots, we asked the following research questions:

- What kinds of prompting strategies do end users employ and why do they choose them?
- What topics do they explore?
- What kinds of bias, unfairness, or problematic content do they look for? What do they identify?
- How does the outcome of their audit change their perceptions of AI chatbots?

We developed and deployed an observational platform in January 2025 on which participants could interact with and audit an AI chatbot. We recruited a U.S. representative sample of 260 participants and tasked them with assessing how well an AI chatbot upheld ten principles related to fairness, trust, and risks sourced from frameworks established by prior literature [1, 2]. Participants were required to have two conversations and after each one, indicate if the outputs contained any bias, unfairness, or problematic content.

We collected a total of 562 conversations which we qualitatively analyzed, capturing the topics that participants discussed with the AI chatbot, what they asked the chatbot to do or generate, how they interacted with the chatbot, and how they assessed the outputs. We also conducted interviews with 14 randomly selected participants to better understand the rationales behind their strategy selection,

their expectations of the chatbot, and in what ways—if at all—their perceptions on chatbots changed after the assessment.

Our preliminary findings indicate that most (75%) of our participants did not find any bias, unfairness, or problematic content. Across our sample, participants employed a variety of assessment strategies, with an emphasis on factual correctness and neutrality on subjective and divisive topics. Some participants utilized more “deceptive” approaches, such as taking a biased position or and checking whether the chatbot would conform to or reinforce those views.

Preliminary analysis of our interview data indicate that the reasons for strategy selection is highly informed by their personal backgrounds and prior experiences with AI chatbots. When it came to choosing a topic of conversation, many simply did so because it was front of mind or highly visible in the public discourse. Many participants asked the chatbot to do something that they typically ask for in their everyday usage with other AI or had prior exposure to known failure cases.

Our participants had a high degree of confidence in their assessment quality, rating their methodology as comprehensive and effective. Our quantitative analyses indicate that participants who did not find any biases, unfairness, or problematic outputs had increased positive perceptions of the reliability and trustworthiness of AI chatbots. This poses a potentially hazardous scenario in which insufficiently comprehensive auditing regimens lead users to place greater trust in AI systems and potentially adopt or reinforce the biases present in model itself.

### REFERENCES

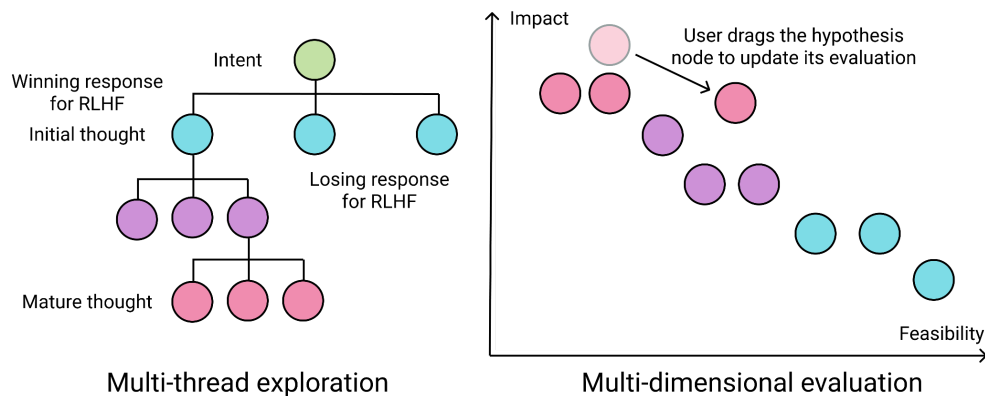
1. Narayanan, D., Nagpal, M., McGuire, J., Schweitzer, S., and De Cremer, D. Fairness Perceptions of Artificial Intelligence: A Review and Path Forward. *International Journal of Human-Computer Interaction*, 40(1):4–23, 2023.
2. Starke, C., Baleis, J., Keller, B., & Marcinkowski, F. Fairness perceptions of algorithmic decision-making: A systematic review of the empirical literature. *Big Data & Society*, 9(2):205395172211151, 2022.

# Towards Human-AI Alignment on the Conceptual Level: GenAI-assisted Hypotheses Exploration and Evaluation

Zijian Ding, Joel Chan

College of Information, University of Maryland

Contact: [ding@umd.edu](mailto:ding@umd.edu)



**Figure 1** Illustrations of GenAI-assisted hypotheses exploration view (left) and hypothesis evaluation view (right).

Current research in Generative AI has shown remarkable progress in models' capacity to "do things right"—reliably executing well-specified tasks—, a trend evident in recent saturated benchmarks. Yet conceptual-level "doing the right thing", such as formulating and evaluating testable hypotheses, remains stubbornly difficult for fully autonomous systems because it requires a nuanced appreciation of why certain approaches are inappropriate. Human experts develop fine-grained right-wrong classifiers for "doing the right thing" through years of practice, iterative successes and failures, and peer exchange. Much of this tacit knowledge is not externalized in documentation or publications and therefore does not reach GenAI training data. Therefore, a workflow that fuses human expertise with GenAI is essential for navigating complex decision processes that demand parallel exploration and multi-dimensional evaluation.

We study this alignment problem in the context of hypothesis exploration and evaluation, a complex conceptual task that requires expert intuition and control. We are prototyping a GenAI-assisted hypothesis exploration and evaluation system with two key

components (Figure 1): 1) In the *multi-thread exploration view*, a generative model proposes candidate hypotheses from an initial intent or research question; the system visualizes them as an ordered node-link diagram that situates each branch relative to the overarching intent and to one another, enabling the user to perceive and control the breadth and depth of exploration [1]. 2) In the *multi-dimensional evaluation view*, these hypothesis nodes are placed on an XY plane whose axes represent context-specific criteria—such as feasibility, clarity, novelty, significance, and relevance. Users can flexibly refine the hypothesis evaluation or its scores by simply dragging the node within the plane. Interaction traces between experts and the model can then feed RLHF post-training loops, steadily tightening the model's alignment with expert judgment.

## REFERENCES

Z. Ding, M. Brachman, J. Chan, W. Geyer. Structuring GenAI-assisted Hypotheses Exploration with an Interactive Shared Representation. In *Companion Proceedings of the 30th International Conference on Intelligent User Interface (IUI'25 demo)*.



# Virtual Co-presenter: Connecting Deaf and Hard-of-Hearing Livestreamers and Hearing Audience in E-Commerce Livestreaming

Yuehan Qiao<sup>1,2</sup>, Zhihao Yao<sup>2</sup>, Meiyu Hu<sup>3</sup>, Qianyao Xu<sup>2</sup>

<sup>1</sup>College of Information, University of Maryland   <sup>2</sup>Tsinghua University   <sup>3</sup>Tongji University  
yhqiao@umd.edu

## OVERVIEW

Deaf and Hard-of-Hearing (DHH) individuals are increasingly participating as livestreamers in China's growing e-commerce industry. However, their reliance on sign language (SL) communication significantly limits their reach and engagement with hearing audiences, negatively impacting their potential earnings. Existing SL translation solutions struggle to adequately address the unique demands of real-time e-commerce livestreaming, and the high cost of hiring SL interpreters further restricts their ability to expand their audience base.

Through semi-structured interviews with 4 SL livestreaming team members and 15 hearing audiences, we explored barriers that cause hearing audiences to lack the intention to continue watching, ultimately leading to their absence in SL livestreaming. Beyond the difficulties in understanding SL, hearing audiences struggled with visually overwhelming presentations, including simultaneous signing, textual product information, and demonstrations. Additionally, unrelated background audio to prevent platform penalties for silent streams disrupted audiences and reduced their trust. Emotional expressions from DHH livestreamers, conveyed through signing and facial gestures, were frequently misunderstood as aggressive or rushed, lowering audience engagement and purchase intent.

To address these issues, we developed a virtual co-presenter through a design workshop involving six designers. The virtual co-presenter, designed as a simple, human-like cartoon character, provides synchronized voice broadcasting paired with expressive gesture animations. Specific gestures were crafted to naturally highlight important product details and amplify positive emotional cues at key moments, improving clarity and viewer engagement. Initial evaluations with both SL livestreaming team members and hearing audiences suggested that the virtual co-presenter shows potential in enhancing information comprehension and emotional resonance. They emphasized the importance of naturalistic and relatable animations and suggested integrating symbolic elements representing DHH identities, such as hearing aids, to better communicate identity.



Figure 1 – Virtual Co-presenter

Ongoing evaluations with the improved virtual co-presenter focus on its application in real-world livestreaming scenarios. Early results indicate that the virtual co-presenter improves audience engagement, comprehension, and trust while preserving the primary role and visibility of DHH livestreamers. These insights contribute to designing accessible and inclusive technologies that enhance participation and interaction in e-commerce livestreaming.

## ACKNOWLEDGEMENTS

We acknowledge the contributions of Yun Wang and Weiwei Zhang in facilitating the workshop and interview, and the valuable suggestions from Dr. Hernisa Kacorri.

## PAPERS

1. Yuehan Qiao, Zhihao Yao, Meiyu Hu, and Qianyao Xu. 2025. Virtual Co-presenter: Connecting Deaf and Hard-of-hearing Livestreamers and Hearing audience in E-commerce Livestreaming. In Extended Abstracts of the CHI Conference on Human Factors in Computing Systems (CHI EA '25), April 26-May 1, 2025, Yokohama, Japan. ACM, New York, NY, USA, 9 pages. <https://doi.org/10.1145/3706599.3720130>.



## Strategies to Enhance Technology Learnability For Older Adults: A Systematic Review

Nupur Wagle, Caro Williams-Pierce, Amanda Lazar

College of Information

Contact: [nwagle@umd.edu](mailto:nwagle@umd.edu)

### INTRODUCTION

Older adults are increasingly engaging with digital technologies such as smartphones, web applications, and online services, but they can encounter barriers because most systems are not designed with their needs in mind. There are commercially available products that are advertised as being designed specifically for older adults. However, this approach may not be practical at the scale needed for our increasingly digitized world due to cost, complexity, and potential stigma. A more scalable approach may be to investigate the best approaches to provide supportive materials for existing technologies.

We are conducting a systematic literature review to investigate how researchers are providing information or cues for existing technologies or web-based content (e.g., hardware, apps, webpages) to make them more learnable, usable, and accessible for older adults. By synthesizing design patterns and implementation approaches, we aim to surface actionable insights that reduce barriers and promote more equitable digital inclusion for aging populations.

### RESEARCH QUESTIONS

This review is guided by the following research questions:

1. What strategies or design patterns have researchers used to add information or cues to interfaces?
2. What stage of development do these projects reach?
3. How are those strategies/design patterns implemented?
4. What interfaces or technologies are these strategies designed to support?

### METHOD

We conducted a systematic literature review following the PRISMA guidelines. After defining our research questions and reviewing pilot papers to identify key terminology, we performed structured keyword searches in the ACM Digital Library. We screened 245 papers by title and abstract, selecting 26 that described strategies for adding or layering support onto existing digital interfaces to improve usability for older adults. We are currently conducting a thematic

analysis of strategies, development stages, and technologies, using open and axial coding to surface patterns across design approaches.

### FINDINGS

Our early analysis reveals a variety of strategies used to support older adults through added layers of guidance. These fall into the following core themes:

1. **Contextual Overlays and Visual Cues:** Systems that layer information directly on top of interfaces (e.g., tooltips, ghost hands, floating labels) to support recognition and reduce cognitive load.
2. **Step-by-Step Instructional Guidance:** Progressive or scaffolded learning through sequential prompts, audio narration, or guided tasks.
3. **Error-Tolerant Exploration:** Designs that reduce fear of failure by allowing users to make mistakes safely.
4. **Search Space Reduction and Feature Narrowing:** Systems that minimize complexity by filtering or directing attention to relevant features.
5. **Emerging Strategies:** Some papers also explored **multimodal cueing** (e.g., visual + audio + tactile feedback) and **social scaffolding** (e.g., peer support via video calls or shared walkthroughs) as promising directions.

### NEXT STEPS

Our analysis of design strategies (RQ1) has revealed a rich set of themes, and we are currently in the process of examining how these strategies are implemented (RQ3), the types of technologies they support (RQ4), and the stages of development they reach (RQ2). Ultimately, the goal is to inform future research and design efforts that move us closer to inclusive, age-friendly digital ecosystems.

### ACKNOWLEDGEMENTS

The contents of this abstract were developed under a grant from the National Institute on Disability, Independent Living, and Rehabilitation Research, ACL (NIDILRR grant #90REGE0024).

# Foreseeing Quality of Life in AAC: A Systematic Literature Review on Interpersonal Communication in AAC

Ruotong Gao, Claire O'Connor, Lei Mao, Stephanie Valencia

College of Information, University of Maryland

Contact: ruotongg@umd.edu

## OVERVIEW

“We desperately need improvements and have advanced so far in technology [that] there’s really no reason research hasn’t yet expanded to improve [our] quality of life.”

—Sarah Price Hancock, CommunicationFIRST.

Interpersonal communication is central to multiple dimensions of quality of life, yet it remains unclear how Augmentative and Alternative Communication (AAC) technologies can best support quality of life for individuals with communication access needs. As AAC research increasingly embraces user-centered values, there is a growing need to identify both current gaps and emerging opportunities related to AAC users’ quality of life. In response to the community call from CommunicationFIRST—an advocacy group led by AAC users—underscoring ongoing concerns among AAC users and advocates about quality of life and the need for greater attention from researchers, we conducted a systematic literature review on 67 peer-reviewed articles on AAC and interpersonal communication. We focus on how these works address quality-of-life outcomes and we ask:

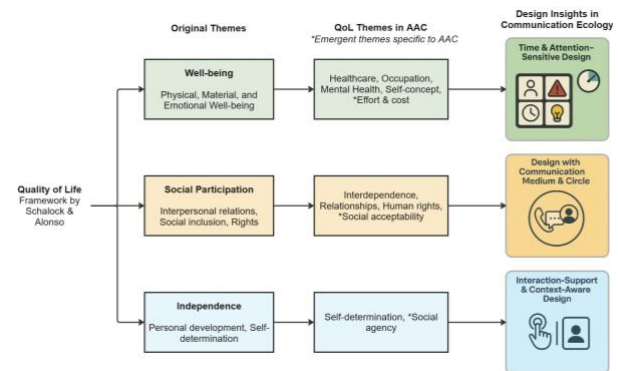
**RQ1:** How do current and emerging AAC research efforts contribute to improving quality of life and interpersonal communication of AAC users?

**RQ2:** To what extent does the research contribute to quality of life, and what aspects are underrepresented?

**RQ3:** What design strategies in AAC technology enhance users’ lives and communication, and what challenges arise?

## FINDINGS

Quantitatively, we examined publication trends, venues, geographic distribution, the types of technologies studied, disability communities represented, and the extent of stakeholder involvement. Qualitatively, we conducted a thematic analysis using a widely recognized quality-of-life framework and identified emerging themes and gaps in the three core themes: **well-being**, **social participation**, and **independence**, as depicted in Figure 1. We further interpret these findings through a communication ecology lens to



**Figure 1** – Structure of the findings linking QoL and design insights.

offer recommendations for research and design that better support AAC users’ quality of life within multilayered communication ecologies, pointing to design implications that prioritize **interdependent** processes and position AAC users as **co-creators** of communication systems.

## HIGHLIGHTS

In our analysis, the QoL framework serves as a critical lens for identifying gaps and emerging themes in AAC research. We identify five key gaps in the areas of **employment**, **education**, **leisure**, **rights-based participation**, and **reciprocal community relationships**, which remains underexplored in relation to communication access. Emerging themes in AAC research include the **cost and effort** involved in operating AAC systems, the **social acceptability** of AAC use, and the **social agency** AAC users experience in conversations. By identifying both underexplored areas and emerging dimensions in the literature, we aim to encourage a more comprehensive and user-centered approach in AAC research, one that accounts for the full spectrum of communicative, emotional, social, and structural factors that shape the lives of AAC users.



# I'm Sorry, Let Me Rephrase: AI-Powered Tools for AAC Users to Communicate Intent Effectively

Claire O'Connor

College of Information

coconno5@umd.edu

## OVERVIEW

Augmentative and Alternative Communication (AAC) is widely recognized as an assistive tool for individuals with speech and language impairments, supporting their communication needs [1]. AAC devices may consist of high technology devices which include touch and eye gaze interaction with a tablet while low technology devices consist of selected pictured icons from a physical board (PECS). As AAC use grows among individuals with disabilities, common challenges prevail such as time delays and high effort required to compose responses with AAC solutions [2,3,4]. Language models offer potential solutions by generating responses instantly with minimal input, whether written or voice based. Through a codesign workshop with AAC users and their caregivers, alongside focus groups with speech-language pathologists (SLPs), this study aims to (1) understand how AAC users and caregivers work together to create prompts to help AAC users leverage LLMs as a communication tool, (2) understand the benefits and challenges of using LLMs to facilitate conversational agency between AAC users and their communication partners and (3) evaluate the feasibility of using LLMs for AAC communication from the AAC user and SLP's perspective. The codesign workshop will engage AAC users in both physical and digital activities designed to develop prompts for ChatGPT, while the focus groups will provide SLPs with an opportunity to share insights on the practicality and implications of integrating ChatGPT into AAC communication. The findings from this research will inform future codesign workshops with additional populations, such as individuals with developmental (e.g. autism) and cognitive disabilities (e.g. aphasia), and support the development of specialized use cases within a fine-tuned model. By expanding this research to address communication challenges for other underrepresented

groups, this study fosters a more inclusive approach to AI innovation.

## ACKNOWLEDGEMENTS

Full funding was provided by Google for this research.

## PAPERS

1. [n.d.]. Augmentative and Alternative Communication (AAC). <https://www.asha.org/practice-portal/professional-issues/augmentative-andalternative-communication/> Publisher: American Speech-Language-Hearing Association.
2. Seray Ibrahim, Michael Clarke, Asimina Vasalou, and Jeff Bezemer. 2024. Common ground in AAC: how children who use AAC and teaching staff shape interaction in the multimodal classroom. *Augmentative and Alternative Communication* 40, 2 (April 2024), 74–85. doi:10.1080/07434618.2023.2283853 Publisher: Taylor & Francis eprint: <https://doi.org/10.1080/07434618.2023.2283853>.
3. Jennifer Seale, Ann Bisantz, and Jeffery Higginbotham. 2020. Interaction symmetry: Assessing augmented speaker and oral speaker performances across four tasks. *Augmentative and Alternative Communication* 36 (July 2020), 1–13. doi:10.1080/07434618.2020.1782987
4. Stephanie Valencia, Amy Pavel, Jared Santa Maria, Seunga (Gloria) Yu, Jeffrey P. Bigham, and Henny Admoni. 2020. Conversational Agency in Augmentative and Alternative Communication. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems (CHI '20)*. Association for Computing Machinery, New York, NY, USA, 1–12. doi:10.1145/3313831.3376376



## Approaches for Improving Color Contrast in Authoring Documents

Alexa Siu<sup>1</sup>, Rajiv Jain<sup>1</sup>, Abhinav Kannan<sup>2</sup>, Jose Echevarria<sup>1</sup>, Mary Ann (MJ) Jawili<sup>3</sup>, Yalpi Shiva Prasad<sup>3</sup>, Rick Treitman<sup>3</sup>, Garreth W. Tigwell<sup>4</sup>, Jonathan Lazar<sup>2</sup>

Adobe Research<sup>1</sup>, University of Maryland<sup>2</sup>, Adobe Inc.<sup>3</sup>, Rochester Institute of Technology<sup>4</sup>

Contact: [abhinavk@umd.edu](mailto:abhinavk@umd.edu)

### OVERVIEW

Effective use of colors is important in documents so that they are more accessible, especially to people with color vision deficiencies (CVD). A number of challenges exist for document authors and designers to ensure that their content is color accessible, in addition to visually appealing. Several guidelines exist, in addition to tools that help authors meet color contrast requirements. However, users often lack awareness or clarity in incorporating them into their authoring and workflow process [1].

In this talk, we present a set of prototypes and approaches that can support designers by providing real time color accessibility recommendations using a born-accessible workflow. Previous work on remediation of documents for accessibility uses a remediation approach [2]. In contrast, our system provides just-in-time feedback that alerts users *during* the design process, as opposed to later remediation. We explore approaches that address complex backgrounds such as gradients, images and text backgrounds, while providing born-accessible recommendations that meet WCAG contrast guidelines [3]. Our system recommendations include adjusting text color, adding backgrounds or outlines, and changing opacity levels.

We have conducted two studies that evaluate the effectiveness of our system – the first being a survey among 40 participants on user preferences on accessibility and aesthetics – across a set of text-background combinations. The second study was an interview study with 8 participants evaluating the integration of our system into authors' design workflow, compared to a baseline system.

Participants' feedback suggests that designers find value from a directly integrated, just-in-time system that respects their design autonomy while meeting accessibility requirements. Our hope is that future work builds on born-accessible approaches for design, and addresses more gaps between requirements for accessibility and creative authoring.

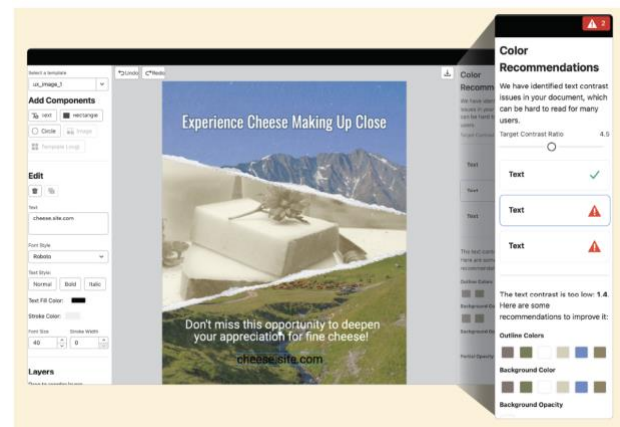


Figure 1 – Color contrast prototype.

### ACKNOWLEDGEMENTS

This project is funded by Adobe Research and the UMD Grand Challenges Impact Award.

### PAPERS

1. Sarah Andrew and Garreth W. Tigwell. 2022. Accessible Design is Mediated by Job Support Structures and Knowledge Gained Through Design Career Pathways. *Proc. ACM Hum.-Comput. Interact.* 6, CSCW2 (November 2022), 1–24. <https://doi.org/10.1145/3555588>
2. Sparsh Paliwal, Joshua Hoeflich, J Bern Jordan, Rajiv Jain, Vlad I Morariu, Alexa Siu, and Jonathan Lazar. 2024. FormA11y - A tool for remediating PDF forms for accessibility. *ACM Trans. Comput.-Hum. Interact.* (November 2024), 3702317. <https://doi.org/10.1145/3702317>
3. World Wide Web Consortium. Understanding Contrast Minimum (WCAG 2.1). Retrieved April 24, 2025 from <https://www.w3.org/WAI/WCAG21/Understanding/contrast-minimum.html>



## Flash Forward: Towards Safe Video Tools for Photosensitive Viewers

J. Bern Jordan

College of Information, University of Maryland

[bjordan@umd.edu](mailto:bjordan@umd.edu)

### OVERVIEW

Flashing in the environment, video content, and electronic games can cause seizures in people with photosensitive epilepsy (PSE). To be considered hazardous, flashing must simultaneously be (1) intense enough (either changes in luminance or changes in chromaticity from saturated red), (2) cover a sufficiently large area, and (3) be within a range of flashing rates. There are five guidelines and standards that are not fully harmonized (described in detail in [1]).

### TOOLS FOR VIDEO SAFETY

Automatic tools for hazard screening are important because manual evaluation of videos for hazards is laborious.

Following on after the success of the longstanding free (but not open source) Photosensitive Epilepsy Analysis Tool (PEAT, Figure 1), we want to build new, open-source tools to help developers, publishers, and content creators improve safety for all viewers.

#### For developers of safety algorithms

In the last few years there have been a number of projects for PSE safety. Some projects have followed published standards and have done limited testing. Other projects (including all current open-source tools) were not developed to meet any specific safety standards.

To ensure that algorithms are properly catching hazardous content, there needs to be better test videos that are not just simple flashing frames. We are developing tools and a set of benchmarking videos that cover a broader range of potentially harmful stimuli [2].

#### For video publishers & platforms

There is growing interest in content safety among social media and video streaming services. We are developing standards-compliant algorithms that can flag hazardous materials in a variety of workflows. The first algorithms will analyze content according to the exact thresholds of the respective standard. Later algorithms may take multiple passes, for example, a conservative screening pass followed by a resource-intensive compliance pass.

#### For content creators

Content creators both large and small are also becoming more interested in content safety. We are developing user-

friendly tools that can analyze content for hazards. For some types of failures, the tool could also give guidance on potential fixes.



Figure 1 – PEAT analysis of hazardous content.

### PROJECT WEBSITE

The current PEAT software is available at <https://trace.umd.edu/peat/>

### ACKNOWLEDGEMENTS

The contents of this extended abstract were developed in part with support under Grant #90REGE0024 from the National Institute on Disability, Independent Living, and Rehabilitation Research (NIDILRR), Administration for Community Living (ACL), Department of Health and Human Services (HHS).

### PAPERS

1. Jordan, J. B. & Vanderheiden, G. C. (2024). International Guidelines for Photosensitive Epilepsy: Gap Analysis and Recommendations. *ACM Trans. Access. Comput.* 17(3), Article 17, 1-35. <https://doi.org/10.1145/3694790>
2. Jordan, J. B. (accepted 2025). Evaluating Conformance of Video Safety Tools for Photosensitive Epilepsy. In M. Antona & C. Stephanidis (eds.), *Universal Access in HCI, HCI International 2025*, 15 pages



# Perceptions, barriers, and Integration of LLMs Across the Cybersecurity Workforce

Miuyin M. Yong Wong, Alan F. Luo, Clarissa E. Gullion, Michelle L. Mazurek

Department of Computer Science, University of Maryland

miuyin@umd.edu

## OVERVIEW

The public release of commercial Large Language Models (LLMs) such as ChatGPT has changed the technology landscape almost overnight. Organizations in nearly every sector appear to be racing to integrate these tools in pursuit of dramatic productivity gain. In cybersecurity, controlled research studies have demonstrated that even with their limitations, LLMs can assist in security tasks such as reverse engineering, malware detection and capture the flag exercises [1]. At the same time, other research has shown how the same models can be weaponized to generate malicious code or write convincing phishing emails [2], making defenders job harder.

Despite these promising and worrying findings, it remains unclear whether these lab driven successes translate into everyday practice. While prior work have explored how LLMs are used in practice for software development [3] for example, we know little about if or how cyber professionals actually use LLMs for defensive tasks. Given the ongoing staff shortages in the cybersecurity space and the overwhelming amount of cyber attacks, this workforce stands to benefit greatly. That makes it crucial to understand the current LLM usage, identify barriers of adoption and uncover the challenges practitioners face when integrating these tools into their workflows. This study aims to fill this gap.

## METHODS

We began the study by requesting participants to fill out a prescreening questionnaire that collected information about their job role, years of cybersecurity experience, basic demographics, and frequency usage of LLMs in cybersecurity related job tasks. Any respondent currently working in a cybersecurity role was then invited to an hour long semi-structured interview. During the interviews, we branched by self-reported LLM use: (1) Current LLM users were asked about their adoption history, the specific security daily tasks they use LLMs for, custom integration, their process for verifying outputs, and best practice they have developed (2) Non users were asked to describe the factors that have kept them from using LLMs and what factors may encourage future adoption. In both tracks we asked

participants about their perceived risks, concerns and desired improvements to understand how LLMs are or aren't integrated into their cybersecurity workflows.

## PRELIMINARY RESULTS

We have completed three interviews so far, each with a security practitioner who actively incorporates LLMs into their daily work.

### Perceived Benefits

- Time Saving: Participants report a significant reduction in analysis time, especially when writing scripts that decrypt or unpack malware samples
- Pattern matching: Participants view LLMs biggest strength is identifying patterns. In fact, two participants predict that future models will replace signature-based detection and even reshape the traditional SOC roles as we know them today.

### Key Concerns

- Data leakage: Risk of exposing customers proprietary data when sending prompts to cloud-hosted models. Their companies have written policies that prohibit the use of commercial LLMs for sensitive data but rely on individual judgement rather than technical guard-rails.
- Required human supervision: Participants stressed that for the foreseeable future, LLM output must be reviewed by humans before action or escalation.

Overall, participants see productivity advantages and are optimistic about future improvements but remain cautious.

## REFERENCES

1. Pearce, Hammond, et al. "Pop quiz! can a large language model help with reverse engineering?." *arXiv* 2022
2. Roy, Sayak Saha, et al. "From chatbots to phishbots?: Phishing scam generation in commercial large language models." *2024 IEEE SP*
3. Klemmer, Jan H., et al. "Using ai assistants in software development: A qualitative study on security practices and concerns." *ACSAC* 2024



## A Design Probe on De-Identification Tools for Researchers

Wentao Guo, Wellington Barbosa,\* Adam J. Aviv,\* Michelle L. Mazurek

University of Maryland, \*George Washington University

[wguo5@umd.edu](mailto:wguo5@umd.edu)

### OVERVIEW

Social scientists, clinical trialists, and other researchers collectively generate extensive data about human subjects. Publishing this data supports reproducibility, meta-analysis, and transparency. However, sharing data about humans risks stigma, legal consequences, violence, and other harms. Thus, researchers *de-identify* data: modifying data to make it harder to re-identify or learn information about individuals.

Our prior work with researchers [1] shows that in practice, de-identification is time-consuming and provides unclear privacy protections. Researchers use tools to scaffold the process, but they are wary of the impact on utility of more rigorous de-identification methods. We believe the right tools could unlock more rigorous de-identification methods that provide stronger privacy protections; improve usability with visualizations and interactive features; and help researchers understand trade-offs between privacy and utility on the fly. To this end, we are conducting an open-ended user study with researchers at UMD and GWU that uses three different de-identification tools as design probes to explore the following research questions:

1. What kinds of workflows and capabilities would researchers want in a de-identification tool?
2. How do researchers assess the acceptability of computational de-identification tools and methods?

### METHODS

We are recruiting graduate students, research staff, and faculty at UMD and GWU. Participants must meet one of three criteria related to their coursework or experience working with quantitative human-subjects data. Selected participants will take be invited to a 75-minute study session in person on campus, for which they will receive \$60.

During the user study, participants use a tool to de-identify a modified version of the 2023 Detroit Metro Area Communities Study dataset [2], which contains information about Detroit residents' demographics, finances, health, employment, and more. We have modified the dataset to reduce its scope (given limited time with participants) and to simulate a less de-identified dataset (e.g., by synthesizing fake values to replace redactions in the published data). De-identifying this dataset is broken into three phases: (1)

*understanding initial data characteristics*, (2) *applying de-identification techniques*, and (3) *understanding the impact of de-identification*. Each phase begins with a co-design activity in which participants imagine how they would like a de-identification tool to support them, before we guide participants through using their assigned tool to accomplish the task at hand. The study concludes with a semi-structured interview portion in which we ask participants to reflect broadly on their experience using the de-identification tool.

Participants are assigned one of three tools, each featuring different de-identification workflows:

- *sdApp* is a GUI interface that provides access to some features from *sdMicro*, an oft-recommended R package. Participants use generalization and suppression methods to attain *k*-anonymity.
- *ARX* is a GUI-based application that is commonly recommended and featured in numerous peer-reviewed papers. Participants use generalization and suppression to attain *k*-anonymity. Unlike *sdApp*, *ARX* does not support our study's phase *understanding initial data characteristics*; therefore, participants instead review a codebook generated by external software such as *Stata*.
- *MST*, a Python script, is based on the algorithm that won NIST's 2018 Differential Privacy Synthetic Data Competition. Participants use *MST* to generate differentially private synthetic data. As *MST* supports neither *understanding initial data characteristics* nor *understanding the impact of de-identification*, participants review a codebook generated by external software in the former phase and use a separate application for synthetic data evaluation in the latter (tentatively, participants are using *SynthRO*, a dashboard for benchmarking synthetic datasets presented in two recent peer-reviewed papers).

### ACKNOWLEDGEMENTS

This work is supported in part by NSF grant OAC-2232863.

### LINKS

1. <https://wentaoquo.com/publications/de-id-interviews/>
2. <https://doi.org/10.3886/ICPSR38892.v1>



# Stakeholders' Perspectives on Privacy Negotiation to Protect Non-Primary Users' Privacy in Smart Environments

Sunyup Park  
Information  
sypark@umd.edu

## OVERVIEW

My dissertation explores the communication opportunities between primary and non-primary users to prevent and resolve privacy-related issues in smart environments.

Smart environments are spaces and places that are equipped with devices that have computational abilities to collect, analyze, and share user data. Since smart devices may collect sensitive data, smart environments threaten users' privacy.

Users in smart environments can be categorized into primary and non-primary users, depending on their access and control of smart devices. Primary users (PUs) have access and control of the devices, and are interested and enthusiastic about adopting smart technologies. Non-primary users (NPU), on the other hand, have limited or no access and control of smart devices, but are exposed to data collection. The privacy implications of smart environments are especially concerning for NPUs, as they are exposed to data collection but have limited means to protect themselves due to their lack of access and control of smart devices.

Research in this space has largely focused on technical approaches PUs take to mitigate privacy risks, many of which are inaccessible to NPUs; more recently, however, some researchers have highlighted the possibility of mitigating NPUs' privacy concerns through non-technical measures, such as communication (i.e., privacy negotiation). Yet, privacy negotiation as a concept is ill-defined and has been approached from a perspective that emphasizes efficiency (e.g., time, outcome).

My dissertation explores privacy negotiation as a way to mitigate NPUs' privacy concerns from stakeholders' perspectives. The first study investigates PUs' consideration of NPUs' privacy, as a starting point of privacy negotiation.

I interviewed 15 short-term rental (e.g., Airbnb) hosts who use smart devices in their rental properties about their privacy-protecting practices (e.g., usage, management, and communication of smart devices) [1]. The second study investigates NPUs' willingness to engage in privacy negotiation. I interviewed 25 incidental users (i.e., people who interact with smart devices that are not theirs) about their privacy-protecting behaviors (e.g., (re)actions to smart devices) [2]. The third study brings primary and non-primary users together to engage in privacy negotiation. I use focus groups with co-design activities to facilitate privacy negotiation and communication. Participants will play a privacy negotiation game to engage in privacy negotiation, reflect on the activity to identify challenges, and ideate on ways to overcome them. By analyzing the interactions between primary and non-primary participants, I hope to identify the values and processes to gain insights for design and guidelines to support privacy negotiation.

## ACKNOWLEDGEMENTS

Funding for this research was provided by National Science Foundation CNS-2232656.

## PAPERS

1. **Park, S.**, He, W., Deldari, E., Emami-Naeini, P., Huang, D. Y., Vitak, J., Yao, Y., & Zimmer, M. (2024). Well-intended but half-hearted: Hosts' consideration of guests' privacy in rental properties. *Proceedings of the Symposium on Usable Privacy and Security (SOUPS)*. USENIX Association.
2. **Park, S.**, Nellore, N., Zimmer, M., & Vitak, J. (Under Review). In Search of "a Way to Level the Playing Field": Helping Incidental Users Navigate Privacy Risks in Smart Environments. Submitted to *Proceedings of the ACM on Human-Computer Interaction, CSCW*.





## A Qualitative Analysis of Fuzzer Usability and Challenges

Yunze Zhao<sup>1</sup>, Wentao Guo<sup>1</sup>, Harrison Goldstein<sup>1</sup>, Daniel Votipk<sup>2</sup>, Kelsey Fulton<sup>3</sup>, Michelle Mazurek<sup>1</sup>

<sup>1</sup>Department of Computer Science, University of Maryland

<sup>2</sup>Department of Computer Science, Tufts University

<sup>3</sup>Department of Computer Science, Colorado School of Mines

[yunze@umd.edu](mailto:yunze@umd.edu)

### OVERVIEW

Fuzzing has become a cornerstone in vulnerability discovery due to its automation and effectiveness in uncovering unexpected software behaviors. Fuzzing is a dynamic testing technique that feeds programs mutated inputs to trigger unexpected behavior. Most modern fuzzers use feedback like code coverage to guide input generation and explore new execution paths. However, most fuzzing research has emphasized technical innovations while overlooking the human factors that influence real-world usage. This paper addresses this gap by conducting 18 semi-structured interviews with experienced fuzzing practitioners across academia and industry to understand the usability challenges developers face when using fuzzers and integrating fuzzers into their workflows, and to identify actionable opportunities for tool improvement. Our work provides a foundational exploration of fuzzing usability from the perspective of real users and calls for a reorientation of fuzzer design toward human-centric, workflow-aware tools.

### Methods

We conducted 18 semi-structured interviews with security engineers, software developers, and researchers who use fuzzing tools in practice. Participants were recruited via community outreach, targeted invitations based on fuzzing activity, and snowball sampling. To ensure broad coverage, we included participants with varying experience levels and roles. Interviews were conducted remotely and transcribed using local automated tools. Transcripts were analyzed using collaborative thematic analysis, resulting in a codebook organized around three core stages of fuzzing workflows: (1) configuring fuzzers and preparing targets, (2) running and monitoring fuzzing campaigns, and (3) integrating fuzzers into broader development processes.

### Users' Understanding of Fuzzing

Participants have diverse yet inconsistent understandings of how fuzzers work. Most learned through self-guided experimentation and informal resources, leading to conceptual gaps. Fuzzers were widely viewed as essential

for exposing unpredictable behavior, yet users often described them as opaque and hard to reason about. Coverage and crash metrics were treated as rough signals rather than reliable indicators, and trust in fuzzers often stemmed from intuition or experience, not transparency.

### Configuring Fuzzers and Preparing Targets

Setup is a one of the major obstacles. Participants struggle with instrumentation, harness creation, and identifying good entry points—especially in complex or legacy codebases. Documentation often assumes prior knowledge, forcing users into trial-and-error workflows. Many rely on intuition or repurpose test cases to prepare fuzzers. There is a strong demand for automated harness generation, guided setup, and better support for target selection and seed evaluation.

### Monitoring Fuzzing Campaigns:

Runtime feedback is widely perceived as insufficient. Participants find coverage metrics and crash reports hard to interpret, and deduplication often unreliable. Many resort to manual triage and custom scripts. Deciding when to stop is also arbitrary, typically based on time limit, resources consumption, and personal preference. Users want clearer signals about progress, richer output, and human-in-the-loop feedback to improve visibility and efficiency.

### Integration of Fuzzing Workflows:

Participants often adapt fuzzers through self-built external wrappers or CI integration to fit their testing environments. Fuzzers are often described as rigid or overcomplicated, especially hybrid fuzzers. Usability, not performance, was seen as the key barrier to broader adoption. Participants desire more modular architectures, smoother integration with development workflows, and better support for reproducibility and collaboration.

### Paper

Yunze Zhao, Wentao Guo, Harrison Goldstein, Daniel Votipk, Kelsey Fulton, Michelle Mazurek. A Qualitative Analysis of Fuzzer Usability and Challenges. CCS, Under Submission (2025).



## Breach Notifications at Scale: Perspectives from the Breached

Peter Hong, Miuyin Yong, Nathan Reiting, Michelle L. Mazurek  
Department of Computer Science, University of Maryland, College Park  
phong12@umd.edu

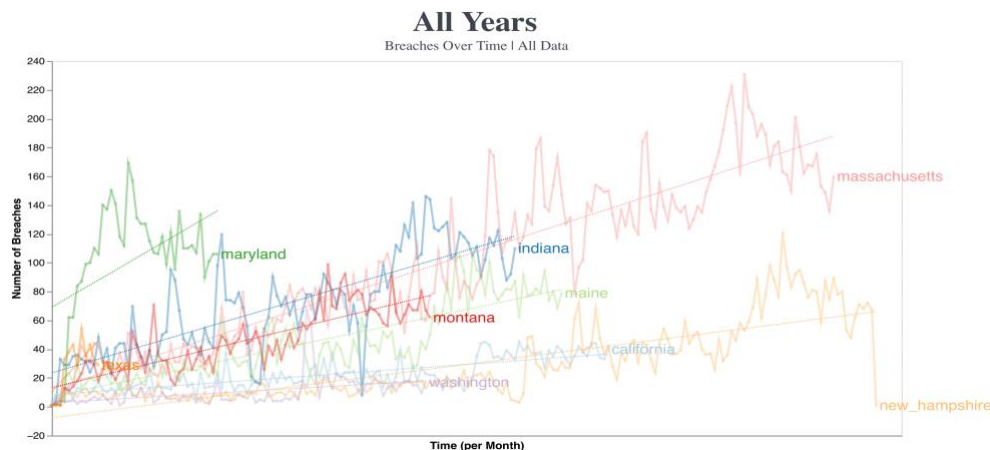


Figure 1 – Monthly count of reported data breaches by state, plotted from each state’s initial reporting period

### OVERVIEW

Every year, as Figure 1 demonstrates, more and more data breaches are occurring. Despite the rising frequency of data breaches, insight into how organizations perceive, experience, and respond to these incidents remains limited.

To address this growing issue, we develop the Breach Notification Database (BNDB) as a contact source for employee interviews, thereby investigating the realities of data breaches from the perspective of the breached. Our goal is to gain insights that will help companies prevent future breaches, grouping our aims along the following research questions: (RQ1) **How was the company affected by the data breach?** (RQ2) **Why are recurring data breaches occurring?** and (RQ3) **What resources would help prevent data breaches from occurring?**

### METHODS

We will reach out to companies that have been affected by a data breach within the last five years, focusing first on those that have experienced more than one breach within the same state in a single year. Subjects will participate in confidential, qualitative-style video call interviews where they will be asked a series of questions relevant to their

organization’s breach experience. Participants will be compensated \$50 for participation in what we expect to be an hour-long interview.

We will assess the participants’ understanding of data breaches and recollection of a specific company-wide breach incident, discussing the onset and general sentiment during the breach. We will then identify perceived barriers to preventing data breaches, as well as any patterns observed, before examining the organization’s internal response protocols and potential technical and legal solutions.

Ultimately, while quantitative metrics such as breach counts, number of individuals affected, and dates of breaches are routinely documented, the human perspectives of those within repeatedly breached organizations are often overlooked. We aim to understand the resources and strategies companies believe would be most effective in mitigating future breaches, develop a comprehensive understanding of the current data breach landscape in the United States, and support the development of practical solutions to combat this proliferating issue.

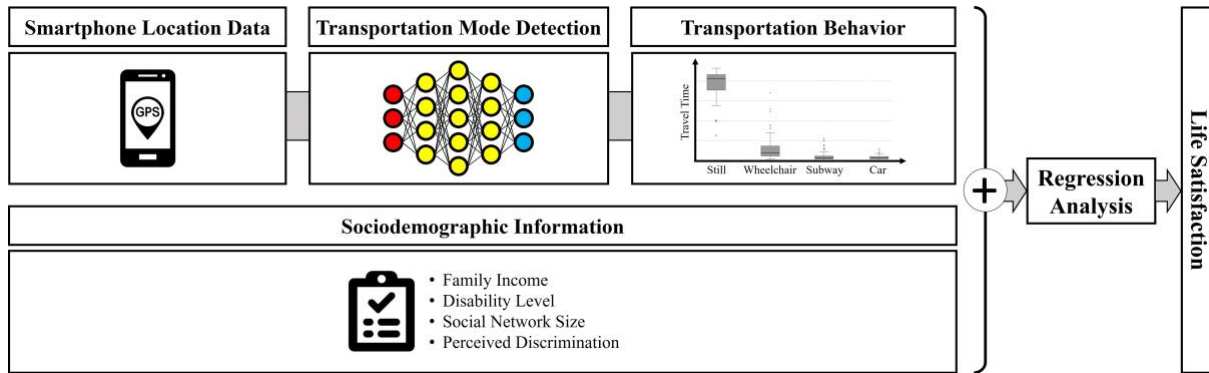
### PROJECT WEBSITE

<https://bndb.umiacs.umd.edu>

# Predicting Wheelchair Users' Life Satisfaction from Smartphone Data Using Transportation Mode Detection

Sungjin Hwang<sup>1</sup>, Jiwoong Heo<sup>2</sup>, Youngwug Cho<sup>2</sup>, Jucheol Moon<sup>3</sup>, Yushin Lee<sup>2</sup>, Hansung Kim<sup>2</sup>, Jaehyuk Cha<sup>2</sup>, Kwanguk Kim<sup>2</sup>

University of Maryland, College Park<sup>1</sup>, Hanyang University<sup>2</sup>, California State University Long Beach<sup>3</sup>  
 sungjin@umd.edu<sup>1</sup>



**Figure 1** – A transportation mode detection model can infer wheelchair users' travel behavior from smartphone location data. Regression analysis shows that transportation behaviors and sociodemographic factors predict life satisfaction.

## OVERVIEW

In this study, we present a novel approach to understanding the impact of mobility on life satisfaction among wheelchair users by applying transportation mode detection technology on smartphone location data. While previous studies mainly focused on improving detection model performances and people without disabilities, we aimed to uncover wheelchair users' travel time using each transportation by applying a transportation mode detection technology on wheelchair users' smartphone location data. Moreover, we investigated the impact of transportation behavior on wheelchair users' life satisfaction through regression analysis. Figure 1 shows the overview of the study.

To achieve our goal, we trained a convolutional neural network (CNN)-based transportation mode detection model using 24,585 minutes of smartphone location data collected from both wheelchair and non-wheelchair users. The model classified location data collected from unseen users into four transportation modes (still, wheelchair, subway, and car) with an 86.76% of the F1-score.

We applied this detection model to an in-the-wild smartphone location dataset collected from wheelchair users in Seoul, South Korea and inferred travel time using each transportation mode. Regression analyses revealed that travel time ratio of subway and four sociodemographic

variables can predict life satisfaction of wheelchair users (adjusted  $R^2 = .171$ ;  $p$ -value  $< .05$ ) and a proportion of subway usage was positively associated with life satisfaction (Coef = 2.541;  $p$ -value  $< .05$ ).

Our study demonstrates how transportation mode detection for wheelchair user technology can inform social science research by providing objective mobility data with minimal user attrition. Moreover, such mobility information can be used for urban planning and accessibility-aware routing algorithms, making the world more accessible for wheelchair users.

## ACKNOWLEDGEMENTS

This work was supported by the National Research Foundation of Korea (NRF) Grant funded by the Korea government [Ministry of Science and ICT(MSIT)] (No.2018R1A5A7059549).

## PUBLICATION

Sungjin Hwang, Jiwoong Heo, Youngwug Cho, Jucheol Moon, Yushin Lee, Hansung Kim, Jaehyuk Cha, and Kwanguk Kim. 2024. Transportation Mode Detection Technology to Predict Wheelchair Users' Life Satisfaction in Seoul, South Korea. *Proc. ACM IMWUT*, 8, 1, Article 9 (March 2024), 20 pages. <https://doi.org/10.1145/3643506>



# From Verbal Reports to Personalized Activity Trackers: Challenges of Data Collection with Older Adults in the Wild

Hossein Khayami<sup>1</sup>, Lining Wang<sup>2</sup>, Young-Ho Kim<sup>3</sup>, Bongshin Lee<sup>4</sup>, David E. Conroy<sup>5</sup>, Amanda Lazar<sup>6</sup>, Eun Kyoung Choe<sup>6</sup>, Hernisa Kacorri<sup>6</sup>

<sup>1</sup>Department of Electrical and Computer Engineering, UMD, <sup>2</sup>Department of Computer Science, UMD, <sup>3</sup>NAVER AI Lab, <sup>4</sup>Yonsei University, <sup>5</sup>University of Michigan, <sup>6</sup>College of Information, UMD

Contact: [khayami@umd.edu](mailto:khayami@umd.edu)

Activity tracking systems offer great potential to support older adults' health and well-being. Yet, their accuracy remains uncertain for this group, especially in real-world settings. Validating such systems requires high-quality ground-truth data—typically collected in labs or labeled by researchers. In this study, we explore the feasibility of collecting ground truth from older adults “in the wild” using two sources: verbal self-reports and a thigh-worn activity tracker, while collecting smartwatch IMU data.

Our findings highlight the need for personalized activity tracking. Cadence thresholds distinguishing stepping intensities—low, moderate, and vigorous—varied significantly across participants, with moderate thresholds ranging from 80 to 105 steps per minute. Verbal reports often failed to clarify intensity levels, as participants commonly used the term “walking” for both low and moderate activity. This variation challenges the use of universal thresholds and supports a personalized approach. Though vigorous activities like running were rare, participants reliably used distinct language for them, with cadence data reinforcing the difference. To improve personalization, we propose combining cadence data with contextual information such as heart rate, elevation, and environmental conditions.

We also examined the costs and benefits of consensus labeling, combining both data sources. Two major costs emerged: **data shrinkage**, with only one-third of IMU data labeled due to incomplete or conflicting labels; and **participant and researcher burden**. Older adults faced challenges with wearing devices and recalling activities, while researchers spent over 300 hours aligning and annotating data.

Despite these challenges, consensus labeling led to notable improvements in model performance, especially for higher-intensity activities like running and cycling. It also surfaced valuable insights into error sources, such as model misclassifications of atypical postures or misinterpretation of tool vibrations as vehicle movement—issues less detectable with a single data source.



**Figure 1** - Overview of the consensus labeling process combining thigh-worn sensor data with verbal reports from older adults, recorded via smartwatch.

Overall, our study highlights both the promise and complexity of collecting real-world ground-truth data from older adults. Personalized thresholds and multimodal data integration are essential for building accurate, user-centered activity recognition systems for this diverse population.

## ACKNOWLEDGEMENTS

We thank our participants for their time and feedback. We're also grateful to Bonnie McClellan and Explorations On Aging for recruitment support. This work was funded by the NSF (Award IIS-1955568), with additional support from NIDILRR (Grant #90REGE0024) and IITP (Grant RS-2020-II201361), MSIT, Republic of Korea, through the AI Graduate School Program at Yonsei University.

## PUBLICATION

Hossein Khayami, Lining Wang, Young-Ho Kim, Bongshin Lee, David E. Conroy, Amanda Lazar, Eun Kyoung Choe, and Hernisa Kacorri. 2025. From Verbal Reports to Personalized Activity Trackers: Understanding the Challenges of Ground Truth Data Collection with Older Adults in the Wild. PACM IMWUT 1, 1 (January 2025), 33 pages. <https://doi.org/10.1145/3731749>



# Toward Enabling Older Adults to Provide High-Quality Activity Labels: Unpacking Accuracy, Precision, and Granularity in Activity Labeling

Yiwen Wang<sup>1</sup>, Hossein Khayami<sup>2</sup>, Bongshin Lee<sup>3</sup>, Amanda Lazar<sup>1</sup>, Hernisa Kacorri<sup>1</sup>, Eun Kyoung Choe<sup>1</sup>

<sup>1</sup>College of Information, University of Maryland, College Park

<sup>2</sup>Department of Electrical and Computer Engineering, University of Maryland, College Park

<sup>3</sup>Yonsei University, Seoul, Republic of Korea

Contact: [yw7615@umd.edu](mailto:yw7615@umd.edu)

Project Website: <https://mymove-collective.github.io/>

## OVERVIEW

High-quality labels with broad representations and real-world variability are key to developing activity recognition models tailored to the needs and characteristics of older adults. However, real-world data labeling presents significant challenges, placing a heavy burden on users to provide high-quality labels while staying engaged in their activities. This paper investigates older adults' perceptions of providing high-quality labels in the context of training their personalized activity trackers. We conducted a co-design study with 12 older adults to envision the labeling process—describing activity names and time spans—using the teachable machines paradigm as a scaffold. We unpack the contextualized definitions of accuracy, precision, and granularity through a thematic analysis of older adults' perspectives on activity labeling. Our findings present participants' preferred strategies, including user-initiated labeling and machine-initiated prompting for acquiring high-quality activity labels while mitigating the labeling burden and the prompting intrusiveness. We discuss design considerations of future data labeling tools that address discrepancies between user perceptions and technical standards when involving older adults in training their personalized activity trackers.

## METHODS

We conducted a co-design study with 12 older adults (aged 64 to 93). Participants shared their understanding and preferred strategies for capturing high-quality activity labels through interacting with a teachable activity tracker. We adopted a scenario-based design method to explore how participants interact with activity labeling systems in their daily context. The activity scenarios include walking, sitting, standing, and other activities they are interested in tracking.

Following that, we provided design probes to elicit participants' feedback regarding user-initiated labeling and machine-initiated prompting strategies in a semi-structured format. Our probes include five topics: (1) interaction modalities and devices, (2) frequency and duration of labeling, (3) when to provide labels, (4) reviewing and correcting, and (5) receiving prompts for labeling.

## CONCLUSION

Our findings revealed that contextual differences and personal significance (e.g., goals and intentions) influence the accuracy and precision with which participants label their activity names. To further alleviate the labeling burden and enhance label quality, participants suggested that machines initiate prompts to ensure and verify completeness, suggest new activities for labeling, and probe for fine granular activity semantics and contexts. Meanwhile, participants desired to maintain control by adjusting label granularity and customizing prompt timing. Understanding participants' perceptions and preferred strategies for collecting high-quality labels can inform the design of future activity labeling systems that older adults can effectively and comfortably use.

## ACKNOWLEDGEMENTS

This work was supported by the National Science Foundation awards #1955568.

## PAPER

Yiwen Wang, Hossein Khayami, Bongshin Lee, Amanda Lazar, Hernisa Kacorri, Eun Kyoung Choe. Toward Enabling Older Adults to Provide High-Quality Activity Labels: Unpacking Accuracy, Precision, and Granularity in Activity Labeling. *IMWUT*. Under submission. (2025).



# Identifying Design Opportunities for Behavior Displacement for Intervening Sedentary Time Among Older Adults

Mengying Li, Eun Kyoung Choe

College of Information

Contact: [myl99629@umd.edu](mailto:myl99629@umd.edu)

## OVERVIEW

Prolonged sedentary behavior poses significant health risks, calling for interventions that promote active lifestyles. For older adults, every physical activity, no matter how small or significant, plays a vital role in their quality of life. However, many interventions aimed at reducing sedentary behavior have overlooked the unique needs and preferences of older adults. In this study, we explore design opportunities for supporting behavior displacement—replacing sedentary time with active movements—as a potential strategy for intervening sedentary time among older adults. Through a 7-day diary study and interviews with 13 participants, we aim to understand how older adults make decisions around displacing sedentary time and identify design opportunities for supporting the displacement in their daily life.

## METHODS

We conducted a 7-day diary and debriefing interviews with 13 older adults (ages 65–83) in the U.S. Participants completed a pre-study questionnaire based on the International Physical Activity Questionnaire (IPAQ), followed by a guided tutorial on diary procedures. During the diary period, participants submitted three or more diaries per day to document sedentary or screen-based activities, indicating context, posture, social setting, devices used, and whether displacement was considered or implemented. Diaries also captured participants' reasoning and any cues or strategies related to movement. Afterward, we conducted one-hour debriefing interviews to reflect on diary entries and explore displacement decision-making. Data were analyzed through codebook thematic analysis.

## FINDINGS

### Considerations on Whether to Displace

Participants described their decisions on displacement in response to physical discomfort, health awareness, and routine responsibilities like household chores or pet care. However, displacement was often avoided when tasks demanded full attention such as work-related activities. Many also preferred to remain sedentary during immersive

leisure moments or when under time pressure. Physical fatigue and social norms further limited displacement, as participants rationalized rest after exertion or chose not to interrupt formal settings. These decisions reflect a deliberate negotiation between well-being, productivity, and the social and cognitive demands of daily life.

### Sequential and Concurrent Displacement

Participants adopted sequential displacement by inserting movement between or within sedentary tasks, especially when the physical and cognitive goals of the activities were competing. Many took intentional breaks during computer work, reading, or watching TV to stand, stretch, or walk. Some scheduled short physical activities between sedentary sessions, or used household chores as a natural interruption.

In contrast, concurrent displacement involved integrating light physical movement into sedentary activities without disrupting them. This was common during low-attention tasks, such as stretching during Zoom calls, walking while on phone calls, or light exercises while watching TV.

## DISCUSSION

Our findings highlight the need for displacement interventions that are adaptive, context-aware, and respectful of older adults' routines and agency. Rather than prompting generic physical activity, systems should detect low-attention moments or routine transitions and suggest subtle movements that align with users' ongoing tasks. Embedding prompts into familiar tools like calendars or multiple devices and apps can reduce disruption while supporting sustained behavior change. We suggest designing with older adults' values and autonomy, enabling displacement as a flexible, integrated part of daily life.

## PAPER

Mengying Li, Eun Kyoung Choe. Identifying Design Opportunities for Behavior Displacement for Intervening Sedentary Time Among Older Adults. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*. Under submission. (2025).



# An Object-Centered Approach To Study Technology Use with People with Cognitive Concerns

Ruipu Hu, Eun Kyoung Choe, Amanda Lazar

College of Information

Contact: rhu12@umd.edu

## INTRODUCTION

Designing with objects may have promise for individuals with cognitive concerns. For instance, prior work found tracking object states can help people recall interactions with them [2]. In this study, we study the role of objects in the context of videoconferencing (VC), a medium through which important activities such as social communication, exercise, and lectures are accessed. We ask, what object qualities support the cognition of people with cognitive concerns during a video call?

## METHODS

We recruited ten older adults with cognitive concerns. To understand how objects come into play, we collected data through interviewing participants, observing their VC use in the usual locations where they participated in VCs, and supporting participants in completing a modified diary study that captures the objects in VCs. We used reflexive thematic analysis to analyze our data [1].

## FINDINGS

Two qualities of objects that support different kinds of cognitive functioning are *grabbability* and *glanceability*.

### Grabbability to Stay On Track.

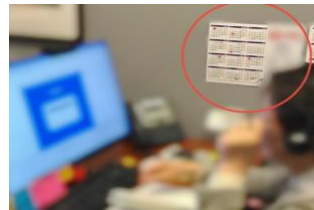
Getting up and searching for an object, like a bottle of water in the middle of a VC, can interrupt the flow of a call, which can result in more significant disruption for people with cognitive concerns. *Grabbable objects* are within arm's reach, often prepared in advance to be in the right location for a participant during a call.

A participant described a time she went upstairs to retrieve something before a call, but then lost track of her original purpose midway due to distractions. P10 now prepares essential objects (e.g., charger) in advance so that they are *grabbable* during a call. Other tools she might need are easily reachable in a bag beside her chair.

### Glanceability to Support Temporal Disorientation.

Time-related information, such as dates and days of the week can be challenging to retain for individuals with cognitive concerns. Placing glanceable objects within the visual field become participants' extended practice to support recalling

time in the face of temporal disorientation during VCs. A participant who had recently begun experiencing more frequent memory difficulties about time, continued to use VC by placing a calendar in a glanceable location.



*The participant discreetly glances at the date while discussing agendas or coordinating future commitments.*

## DISCUSSION

Our findings on object qualities lead to design directions for future VC systems that can better support some types of cognitive needs that can arise in VC use. One example is that temporal-related objects (e.g., clocks, calendars) must remain consistently glanceable throughout the VC session whether on-screen or physically to support temporal disorientation.

## ACKNOWLEDGEMENTS

This work was supported in part by the National Science Foundation (NSF number IIS-2045679) and the National Institute on Disability, Independent Living, and Rehabilitation Research, ACL (NIDILRR grant number 90REGE0024).

## PAPERS

1. Virginia Braun and Victoria Clarke. 2019. Reflecting on reflexive thematic analysis. *Qualitative research in sport, exercise and health* 11, 4 (2019), 589–597.
2. Franklin Mingzhe Li, Di Laura Chen, Mingming Fan, and Khai N. Truong. 2019. FMT: A Wearable Camera-Based Object Tracking Memory Aid for Older Adults. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 3, 3, Article 95 (September 2019), 25 pages. <https://doi.org/10.1145/3351253>



# From Awareness to Action: Ambient Display for Self-regulated Smartphone Usage

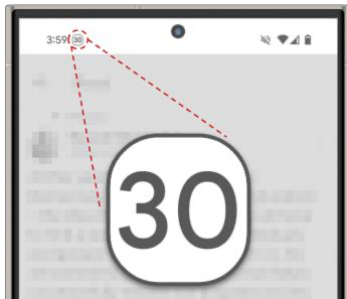
Jian Zheng, Subin Seo, Eun Kyoung Choe

College of Information

{jzheng23, subinseo, choe}@umd.edu

## INTRODUCTION

While using smartphones, people sometimes lose track of time and spend more time than is optimal. Problematic smartphone usage (PSU) can lead to negative physical, psychological, and social consequences. Existing tools designed to mitigate PSU primarily rely on blocking or self-tracking strategies. However, blocking interventions may trigger false alarms and user backlash, while self-tracking often provides limited feedback. We propose ambient displays as a novel intervention strategy for PSU. This approach complements self-tracking by offering continuous exposure to feedback and facilitating in-situ reflection. In this study we explored the effect of the ambient display (with and without customizable attention signals) on phone usage behavior and motivation in controlling phone usage.



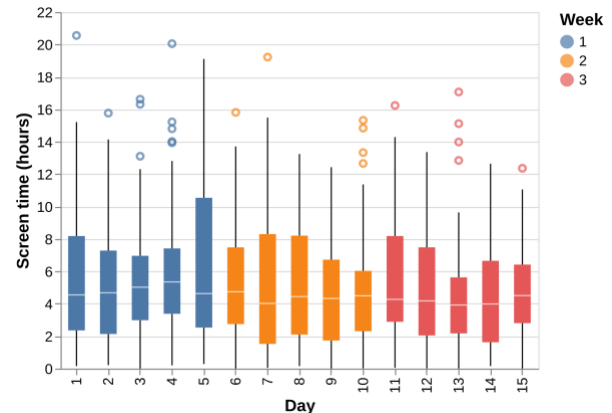
**Fig. 1.** The ambient display of the current session length (30 minutes in the demo) in the status bar.

## METHOD

We designed an Android app called TIMER. It displays the current session length (see Fig. 1) and offers customizable attention signals (i.e., animation, message, sound, and vibration) to enhance the noticeability when needed. Participants ( $N = 41$ ) used TIMER for three weeks: no intervention in Week 1, ambient display in Week 2, and ambient display with attention signals in Week 3. We measured participants' motivation in controlling phone usage before Week 1 and after each week. We interviewed 16 participants afterward.

## RESULTS

TIMER reduced participants' phone usage. The ambient display reduced daily screen time by 38 minutes (10.7% of the baseline); adding attention signals reduced it further by 32 minutes (9.0%, see Fig. 2). Session length was reduced similarly, but unlock frequency did not change.



**Fig. 2.** Day-by-day screen time changes over three weeks: no intervention (Week 1), ambient display (Week 2), and ambient display with attention signals (Week 3).

The survey data revealed an increase in competence of controlling smartphone usage. Participants appreciated TIMER for the simplicity in its design and function, the effectiveness in raising awareness and reducing phone usage, and the support for their autonomy in deciding how to act without coercion.



Try the demo on an Android phone:

## REFERENCES

- Hancock et al. (2019). A meta-analysis of flow effects and time perception. *Acta Psychologica*.
- Lee, Lee, & Choe (2023). Decorative, Evocative, and Uncanny: Reactions on Ambient-to-Disruptive Health Notifications via Plant-Mimicking Shape-Changing Interfaces. *CHI*.

# Comparing Vibrotactile and Skin-Stretch Haptic Feedback for Conveying Spatial Information of Virtual Objects to Blind VR Users

Jiasheng Li, Zining Zhang, Zeyu Yan, Yuhang Zhao, Huaishu Peng  
University of Maryland, University of Wisconsin-Madison  
jsli@umd.edu

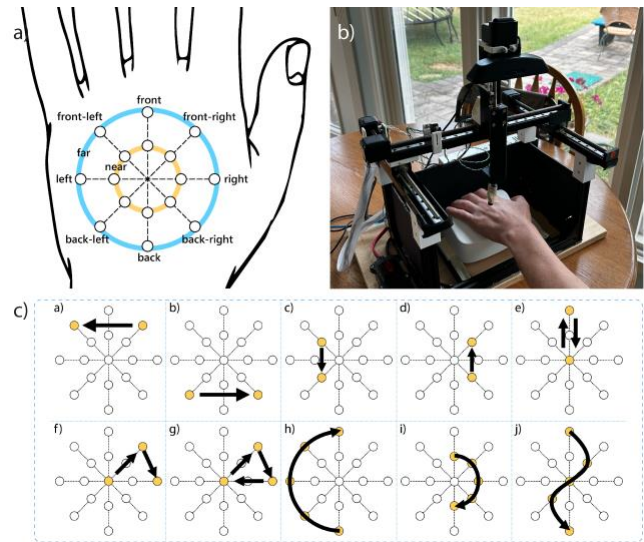
## OVERVIEW

Understanding spatial information is essential for users to perceive and navigate a 3D virtual reality (VR) environment. For sighted users, most spatial cues—such as the relative positions and directions of static objects or the motion of moving objects—are delivered primarily through visual simulations on head-mounted displays (HMDs). However, this vision-centric approach poses significant barriers to blind users, preventing them from perceiving and understanding spatial information in VR spaces.

To facilitate VR accessibility for blind users, we conducted a user study with ten blind participants to investigate the effectiveness of two types of haptic cues—vibrotactile and skin-stretch feedback—in conveying both static spatial information (e.g., the location and direction of a virtual table) and dynamic spatial information (e.g., the trajectory of a flying object in a VR game) when applied to the dorsal side of the hand. Although various haptic mechanisms and body locations are available for delivering haptic stimuli, we selected this combination based on three primary design considerations. First, the skin on the hand has relatively high sensitivity compared to many other body locations. Second, the dorsal side of the hand offers a sufficiently flat surface that is not typically involved in VR operations. Third, both vibration and skin-stretch mechanisms can be engineered with small form factors and low cost, critical considerations when designing assistive technologies.

To map the location and movement of a virtual object relative to the user's position, we designed a coordinate system on the dorsal side of the hand (Figure 1a) in representing a static virtual object's location and direction as well as ten trajectories (Figure 1c) for representing a virtual object's movement paths.

The results from the user study indicated that while blind users were capable of discerning an object's location and movement patterns through both types of haptic feedback, the skin-stretch mechanism enabled more accurate perception than vibrotactile feedback. However, vibrotactile cues—being the most frequently encountered haptic



**Figure 1** - a) coordinates and resolution on the dorsum of the hand, b) study apparatus, c) ten trajectories for rendering the spatial information of a moving object.

mechanism among blind users—provided clear sensations and were preferred by several participants.

Our findings highlight the potential of tailored haptic feedback to enhance spatial awareness for blind users in VR. We offer design insights for incorporating effective haptic solutions into VR systems and advocate for broader consideration of VR accessibility.

## PROJECT WEBSITE

The experimental device is open sourced at GitHub on <https://github.com/jsli96/handHapticforBlind>

## PAPER

Jiasheng Li, Zining Zhang, Zeyu Yan, Yuhang Zhao, and Huaishu Peng. "Comparing Vibrotactile and Skin-Stretch Haptic Feedback for Conveying Spatial Information of Virtual Objects to Blind VR Users." In 2025 IEEE Conference Virtual Reality and 3D User Interfaces (VR), pp. 686-696. IEEE, 2025.



# JetUnit: Rendering Diverse Force Feedback in Virtual Reality Using Water Jets

Zining Zhang, Jiasheng Li, Zeyu Yan, Jun Nishida, Huaishu Peng  
 Department of Computer Science, University of Maryland  
 znzhang@umd.edu

## OVERVIEW

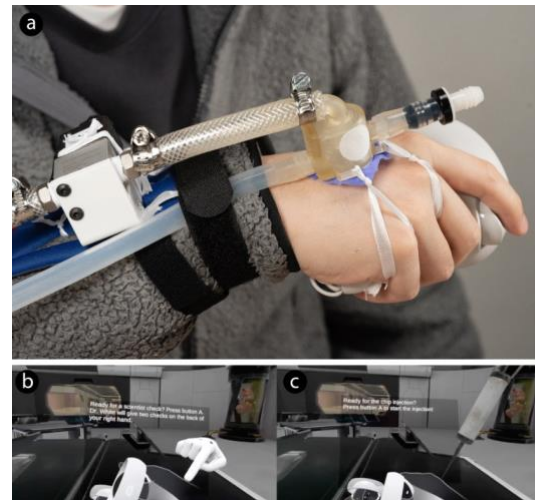
JetUnit is a wearable haptic system designed to provide realistic and diverse force feedback in virtual reality (VR) environments using water jets. Current VR haptic devices, such as pneumatic actuators, exoskeletons, or electrical muscle stimulation, often have limitations in delivering a wide range of sensations due to their physical or mechanical constraints. JetUnit uses water jets to overcome these limitations, offering a wide spectrum of intensities and frequencies force feedback without causing wetness or discomfort to the user.

The core innovation of JetUnit is its custom-designed chamber, which directs water jets onto a thin, elastic membrane that transfers force feedback directly to the user's skin. To optimize performance, several critical features were incorporated. These include a recycling pump system to efficiently drain water, a ring channel that quickly removes residual water, a protective sleeve to isolate the water jet from internal turbulence, and a balanced air pressure system. Experimental tests showed that these features enabled JetUnit to deliver strong and stable force comparable to direct water jets.

JetUnit's nozzle configuration was carefully optimized, resulting in a 1.2 mm nozzle placed 25 mm from the membrane, which provided the optimal balance of pressure (up to 442 kPa) and user comfort. A user perception study confirmed that participants found the haptic feedback realistic, comfortable, and effective.

To further validate JetUnit's effectiveness, we conducted a comprehensive VR user study involving various scenarios such as gentle taps, injections, continuous pressure, and pulsing impacts. Participants accurately perceived different tactile patterns, reporting enhanced enjoyment and a significantly improved sense of immersion due to the realistic synchronization between visual and tactile experiences.

While promising, the current JetUnit prototype faces some practical challenges, including weight, wearability, and tubing constraints. Future enhancements are aimed at addressing these challenges by implementing wireless systems and lightweight components. Additionally,



**Figure 1** – A single (a) wearable JetUnit can render interactions ranging from (b) a gentle touch to (c) a progressively accelerating injection.

integrating multi-unit configurations can be applied for broader application, and incorporating temperature control and variable contact areas could further enrich the tactile experience.

In conclusion, JetUnit greatly expands the possibilities for realistic haptic interactions with various force feedback in VR, making virtual experiences more immersive, comfortable, and engaging.

## PROJECT WEBSITE

<https://smartlab.cs.umd.edu/publication/jetunit>

## PAPERS

Zining Zhang, Jiasheng Li, Zeyu Yan, Jun Nishida, and Huaishu Peng. 2024. JetUnit: Rendering Diverse Force Feedback in Virtual Reality Using Water Jets. In Proceedings of the 37th Annual ACM Symposium on User Interface Software and Technology (UIST '24). Association for Computing Machinery, New York, NY, USA, Article 136, 1–15. <https://doi.org/10.1145/3654777.3676440>

# PCB RENEWAL: Iterative Reuse of PCB Substrates for Sustainable Electronic Making

Zeyu Yan, Advait Vartak, Jiasheng Li, Zining Zhang, Huaishu Peng  
Computer Science  
Contact: zeyuy@umd.edu

## OVERVIEW

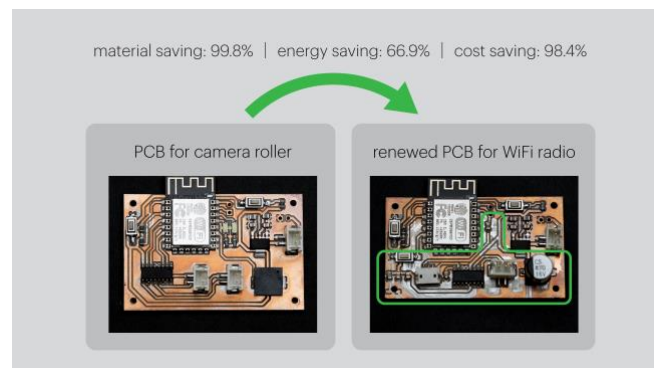
Printed circuit boards (PCBs) are critical components in nearly all electronic devices. However, the obsolescence of PCBs, from their design process to end-of-life disposal, has become an increasingly significant source of electronic waste (e-waste). PCBs are made using a subtractive fabrication method, where copper layers are permanently etched from laminated substrates, making the process *inherently irreversible*. In 2022, less than 23% of globally generated e-waste was formally collected and recycled. Even when PCBs are recycled, they are processed indiscriminately. As a result, they are rarely repaired, repurposed, or reused—even though many PCBs and their substrates remain functional.

These e-waste challenges have garnered attention in the HCI community. Recent work has called for a reimagining of end-users' roles, emphasizing their potential not only as consumers but also as active participants in PCB recycling and reuse. Additionally, researchers have advocated for the development of new processes, tools, and infrastructure to address e-waste and promote sustainable practices.

In this paper, we contribute to sustainable PCB practices by proposing a *reversible* PCB substrate fabrication process that enables the "erasure" and "reconfiguration" of copper layouts. Central to this process is the additive restoration of removed copper areas using conductive fillers, such as conductive epoxy, to renew the PCB substrate for fresh trace patterns. We call this approach PCB RENEWAL.

In the full manuscript, we introduce the workflow of PCB RENEWAL, providing a detailed examination of conductive filler materials and the key fabrication processes involved in the renewal of the commonly used PCB substrate FR-4. We validate our approach through a series of experiments that evaluate key electrical parameters. To assess the sustainability impact of PCB RENEWAL, we present a

quantitative analysis model that compares PCB RENEWAL with the fabrication of new circuits using raw FR-4. This model includes estimates of material usage, cost, time, and energy consumption.



**Figure 1** PCB RENEWAL: PCB modified to accommodate different prototypes reducing major resource consumption.

To help end-users incorporate PCB RENEWAL into their workflow, we develop an EDA software plug-in that allows end users to update a circuit design with changes visualized across iterations, evaluate the sustainability impact of specific renewed designs, and generate the fabrication profiles required for renewal. PCB RENEWAL can be applied to PCBs fabricated either in-house or through outsourcing.

## FULL MANUSCRIPT

[https://www.zeyuyan.com/s/pcbrenewal\\_comp.pdf](https://www.zeyuyan.com/s/pcbrenewal_comp.pdf)

## PAPERS

Zeyu Yan, Advait Vartak, Jiasheng Li, Zining Zhang, and Huaishu Peng. 2025. PCB Renewal: Iterative Reuse of PCB Substrates for Sustainable Electronic Making. (CHI '25). <https://doi.org/10.1145/3706598.3714276>



## Make Making Sustainable: Exploring Sustainability Practices, Challenges, and Opportunities in Making Activities

Zeyu Yan<sup>1</sup>, Mrunal Dhaygude<sup>2</sup>, Huaishu Peng<sup>1,2</sup>

<sup>1</sup>Computer Science, <sup>2</sup>College of Information

Contact: [zeyuy@umd.edu](mailto:zeyuy@umd.edu)

### OVERVIEW

Over the past decade, the traditional relationship between consumers and mass production has evolved with the rise of affordable, compact, and user-friendly desktop fabrication tools. The democratization of these tools has also reshaped the landscape of applied research. For example, personal fabrication has become an important research domain within HCI, with many research labs transitioning from computing-focused spaces to "maker" spaces equipped with personal fabrication. As the push for democratized innovation continues, so does the unintended consequence of democratized waste generation. This growing concern has recently drawn the attention of the technical HCI research community, leading to emerging concepts like "unmaking" or "sustainable making". These work urges HCI researchers to rethink the role of end-users in the broader context of sustainable computing. In addition, there is a growing body of technical research addressing sustainability in personal fabrication.

Building on this body of literature, our paper contributes to the discourse on sustainable making. However, rather than adopting a speculative approach to addressing material waste or framing unmaking as a conceptual lens for reflecting on making practices, our work is grounded in the current realities of maker culture and makerspace infrastructure in the context of sustainability. We aim to explore how today's makers perceive the sustainability impacts of their own making activities, the practices—if any—they adopt to reduce material waste, and the motivations and challenges they encounter in implementing sustainable approaches.

Our qualitative research involves semi-structured interviews with 17 makers based in the U.S. We adopt a broad definition of "maker" as anyone with sufficient experience in creating physical artifacts and potentially generating material waste. This definition allows us to engage participants with diverse backgrounds, including managers and members of grassroots and university-based makerspaces, "home" makers with personal makerspaces (e.g., in their garages), and researchers whose work centers on making. From our qualitative data, we identified four key themes at the intersection of making and sustainability: a detailed account of the types and sources of waste generated in common maker practices, the range of strategies (or absence of them) to manage this waste, the motivations behind adopting sustainable practices, such as feelings of guilt, and the challenges makers face, including the lack of adequate tools for recycling, reusing, or unmaking physical artifacts, even when they aspire to do so. Based on these findings, we organize the discussion around potential technical directions for addressing these challenges, with the goal of inspiring further research and promoting sustainable making practices within the HCI community.

### FULL MANUSCRIPT

[https://www.zeyuyan.com/s/make\\_making\\_sustainable.pdf](https://www.zeyuyan.com/s/make_making_sustainable.pdf)

### PAPERS

Zeyu Yan, Mrunal Sanjay Dhaygude, and Huaishu Peng. 2025. Make Making Sustainable: Exploring Sustainability Practices, Challenges, and Opportunities in Making Activities. (CHI '25).

<https://doi.org/10.1145/3706598.3713665>



# Challenges and Opportunities in Structuring and Facilitating an XR-focused Co-Design Program for Black and Brown Middle School Boys

Akhil Gurram, Clayton Felder, Brian Gardner,  
Elizabeth Bonsignore, Tamara Clegg, Elana Blinder  
University of Maryland | College of Information  
eblinder@umd.edu

## MOTIVATION

Tech literacy continues to become increasingly important in our modern-day society. However, historically marginalized communities often lack access to opportunities to develop these skills and to shape the development of emerging technologies. To address these inequities, we developed and facilitated a two-week summer program for Black and brown middle school boys (rising fifth through eighth graders) to learn about extended reality (XR) technologies. As facilitators, mentors, and researchers in this context, we were interested in better understanding the successes and challenges associated with the design and implementation of our summer program. Therefore, in this study we investigate the research question: *How can we create a welcoming and engaging informal learning environment for Black and brown middle school boys, who vary in age and disposition, to explore and co-design XR technologies?*

## METHODS

To answer this question, we engaged in autoethnographic focus group discussions to reflect on the challenges, successes, and opportunities associated with the implementation of our summer program from the perspectives of: (1) four of our program's near-peer mentors (local Black and brown high school and college students); and (2) our project's core research team (two PIs and one doctoral student). The core research team developed a semi-structured interview protocol to guide three online focus groups—one amongst the three core researchers, one with three program mentors facilitated by two core researchers, and one with one mentor and one core researcher.

All focus group discussions were recorded and transcribed. Then, mentors and core research team members inductively coded portions of the transcripts to develop an initial set of 20 codes, which we are currently in the process of applying to our data, in a second round of iterative inductive and deductive coding.

## PRELIMINARY FINDINGS

In our initial analyses, we observed that our team struggled with a shared desire to provide children with opportunities for self-directed learning and independence, while also enforcing constraints designed to protect their safety and support their creative project development.

Several of us expressed that children should have the freedom to choose which activities to engage with during free choice time. However, we were also concerned that the same set of children kept eagerly lining up to use the virtual reality (VR) headsets and were reluctant to relinquish their devices. This was somewhat surprising in light of frequently voiced admonishments about the “addictive” nature of VR during co-design. In addition to imposing and carefully monitoring time limits, we also soon realized the importance of preventing collisions by requiring children to sit in chairs. Finally, though we directed children to limit themselves to specific VR content, our inability to visually monitor their activity made this difficult to enforce.

## Final Project Development

Similarly, our team struggled to foster children's motivation and engagement with some of our program's more structured learning activities. However, we observed that, when given the freedom to pursue their own interests, many had difficulty getting started or settling on an idea. This often resulted in breakdowns in collaboration and a tendency to stray from the core learning goals. In addition to challenging social dynamics, many boys were preoccupied with developing novel, yet feasible, solutions for their final projects. This, in turn, pushed the limits of their digital prototyping skills and resulted in stressful time crunches to complete and share their projects by the end of the program.

## ACKNOWLEDGMENTS

This work is part of a consortium working to develop ethical guidelines for XR research with children. It was funded by an unrestricted gift from Reality Labs Research.





## Reflecting on Adult Questioning and Prompting Strategies in Co-design with Children

Arafat Bhuiyan, Jincy Njenga, Elli Papasava, Grace Rippin, Deepak Thumma,  
Elizabeth Bonsignore, Urvi Ashturkar, Nitya Jani, Elana Blinder

University of Maryland | College of Information

[ebinder@umd.edu](mailto:ebinder@umd.edu)

### MOTIVATION

As students in the Undergraduate Research Opportunity Program at the College of Information (first five authors), we have been developing our research skills by collaborating with UMD KidsTeam, a co-design team where children and adults work together as equal partners to design technology for children. Though prior work has examined the roles adults can play to foster equitable design collaborations (Yip et al., 2017), as novice researchers and co-designers, we struggled to strike a balance between not being overbearing while also guiding children who may be stuck. We wondered: “Are we asking too many questions?”; “Are we being too quiet?”; “Are we even being helpful?”. In this study, we sought to build on prior work and develop our skills and intuition as adult co-designers by asking: What questioning and prompting strategies could better support adult-child collaboration and shared creativity during co-design?

### METHODS

To answer this question, we collected data from two KidsTeam stakeholder groups: five novice adult design partners; and seven child design partners ages 8-13.

To gauge our effectiveness as adult design partners, we revisited the audio recordings and transcripts of our past co-design sessions. Each adult participant selected an excerpt and wrote a reflection about the questioning and prompting strategies we used. We then conducted a KidsTeam co-design session focused on understanding children’s perception of a positive or negative adult design partner. The session included four activities: (1) Question of the Day: children and adults each responded to the prompt, “Imagine you’re at school and your friend is stuck on something. How would you help them get started?”; (2) Line Judging: children reacted to two fictional co-design scenarios by positioning themselves along a line to indicate their desire to work with/not work with the adult design partner in each scenario; (3) Design Activity: children designed “best” and “worst” adult design partner archetypes (e.g., describing

what these partners might say, do, think, etc.); and (4) Reflection: children wrote or audio/video recorded letters offering advice to new adult partners.

To analyze these data, we are coding our written reflections, and co-design transcripts and artifacts using thematic analysis (Braun & Clarke, 2006) to identify patterns in participating adults’ and kids’ perspectives on adult prompting and questioning. At least two researchers have done a round of inductive open coding on each adult reflection and a portion of the transcripts from our KidsTeam session. We are currently developing a codebook to apply across the data, in a second round of coding.

### PRELIMINARY FINDINGS

Though we are still analyzing our data, we have observed that children generally emphasized how adults made them feel, through being *kind*, *patient*, and *not too bossy*, and asking *enough*, but *not too many questions*. Adults similarly focused on how our delivery of questions and prompts (e.g., *interrupting* or *not giving space*) could affect children’s experience and design process. Both groups emphasized listening, respect, and balanced communication as essential to being a good design partner. Our preliminary findings suggest that effective co-design partnerships rely less on specific know-how and more on emotional intelligence and interpersonal sensitivity.

### ACKNOWLEDGEMENTS

We would like to acknowledge all the children at KidsTeam for their insightful contributions to this work.

### REFERENCES

1. Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative research in psychology*, 3(2), 77-101.
2. Yip, J. C., Sobel, K., Pitt, C., Lee, K. J., Chen, S., Nasu, K., & Pina, L. R. (2017). Examining adult-child interactions in intergenerational participatory design. In *Proceedings of the 2017 CHI conference on human factors in computing systems* (pp. 5742-5754).



## Improving the Usability of an Indonesian Special Collections Repository to Enhance Library Users' Search Experiences

Antariksa Akhmadi, Stephanie Valencia Valencia  
College of Information, University of Maryland College Park  
[antarakh@umd.edu](mailto:antarakh@umd.edu), [sval@umd.edu](mailto:sval@umd.edu)

### OVERVIEW

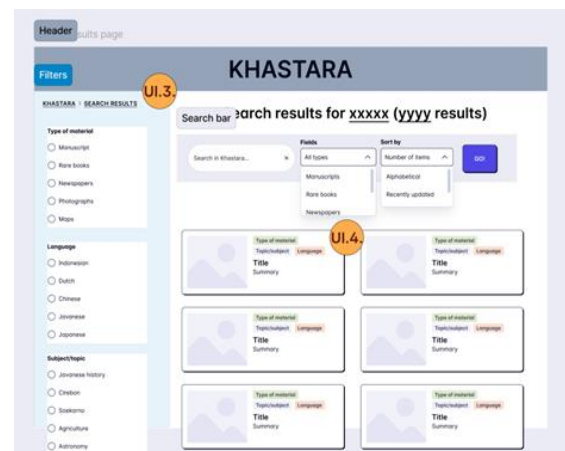
Khasanah Pustaka Nusantara (Khastara) is a repository built by the National Library of Indonesia to provide access to digitized special collections. The items of Khastara reflect the diverse ethnicities, languages, and histories of Indonesia, but it also comes with challenges in search experiences, including content discovery.

### METHODS

A mixed-methods research was conducted, incorporating methods from HCI and theoretical understanding from the information behavior (IB) model of the information search process (ISP). The quantitative element involved surveys using the Usability Scale Score (SUS) questionnaire adapted to the Indonesian language (n = 166). Meanwhile, the qualitative study consisted of contextual interviews with representative users of Khastara (n = 10), as well as librarians (n = 3). Both types of data were joined in mixed-methods data displays. To demonstrate the implications of the results, a proposed wireframe and metadata application profile for Khastara based on the research results were created as well.

### RESULTS

The aggregate SUS study score was 61.72, signifying lower than average usability perceptions in Khastara users. Interview data resulted in a list of usability problems for the Khastara platform. Common usability problems include connectivity issues, confusing navigation, a lack of search features, and missing item metadata. Users' search experiences were explained using ISP and represented in visual form through user journey mapping. Users spent much time dealing with inconsistencies between their expected results and the results they obtained from Khastara, such as in the case of different spellings, different terms being used for different topics, and other issues related to item descriptions and presentations.



### DESIGN IMPLICATIONS

The findings of this research were summarized in a set of design implications. Users need a lot of background information to navigate through the Khastara platform. In addition, they need features to save items for later use and organize their items based mainly on topic and/or date. Khastara's user interface should enable users to explore items through category displays instead of relying on user-typed queries. Metadata descriptions should be expanded with information about languages, topics/subjects, and collections, as well as additional physical descriptions, summaries/abstracts, notes, and relationships with other documents.

Further research needs to be conducted to achieve universal usability and accessibility in digital special collections and cultural heritage experiences.

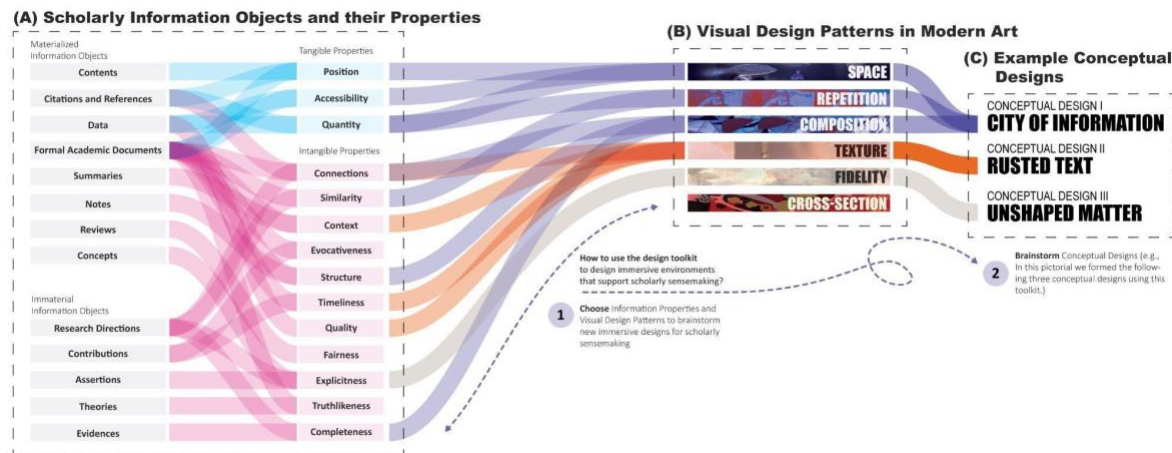
### ACKNOWLEDGEMENTS

Funding for this research was provided by the University of Maryland College of Information through the Dr. Frank G. Burke Endowed Graduate Research Award in Information Studies.



# Sense and Senseability: Exploring Future Immersive Environments for Scholarly Sensemaking

Siyi Zhu, Joel Chan  
College of Information  
zhusy@umd.edu



**Figure 1.** Design toolkit for designing senseable scholarly sensemaking support systems, consisting of (A) a set of scholarly information objects and their associated properties, and (B) a list of visual design patterns generated from the modern arts. A list of conceptual designs brainstormed from choosing properties and visual design patterns is also shown as examples of

Scholars must often make sense of vast amounts of complex and diverse scholarly information, much of which is not “senseable”: crucial information like questions, concepts, or assertions, along with key properties like truthlikeness or evocativeness, are primarily identified through effortful search or reasoning, rather than direct perception through the senses. **How might we augment scholarly sensemaking by making the full range of scholarly information more senseable?**

In this research, we explore how we might augment scholarly sensemaking by making the full range of scholarly information more senseable. First, we systematically reviewed previous empirical studies that examined the information involved in scholarly sensemaking, as well as system designs that support scholars in sensemaking tasks. And we enumerated key types of scholarly information and their properties. Then, we visited three well-known modern art museums to collect visual design patterns that can be adapted to materialize abstract scholarly information in immersive environments.

We synthesized design patterns for materializing abstract information in modern artworks, and connected them with our enumerated scholarly information and properties into a design toolkit (Fig. 1), and used it to develop three novel conceptual designs for senseable scholarly sensemaking in immersive environments. **The City of Information** is a conceptual design that allows scholars to log information in three-dimensional constructions, primarily using visual elements, space, repetition, and composition. This design enhances scholarly sensemaking by providing a set of visual constraints and affordances to aid the synthesis of information. **Rusted text** gives scholarly information a metal-like material and texture, enabling information objects to rust like real metals. The degree of rustiness of the material helps old and less visited information objects stand out visually, which is different from traditional folder structure. Finally, **Unshaped Matter** is a conceptual design that supports scholars’ sensemaking processes in ideating and refining research questions by leveraging the visual elements of Fidelity, and reflection.

# ISee: A Tool for Interactive and Semantic Exploration of Scientific Literature

Yixin Bai<sup>1</sup>, Fumeng Yang<sup>2</sup>

<sup>1</sup>College of Information, <sup>2</sup>Department of Computer Science

[yixbai01@umd.edu](mailto:yixbai01@umd.edu)

## OVERVIEW

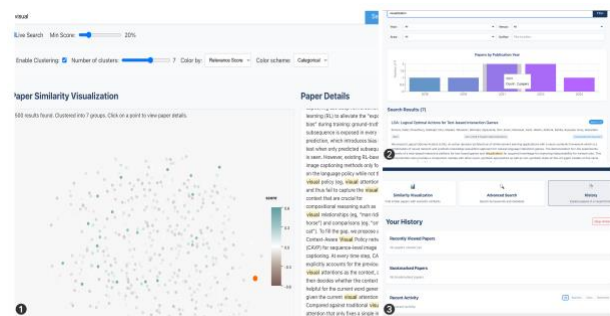
Navigating a large corpus of academic literature poses a significant challenge for researchers, particularly when they hope to understand the semantic connections between papers. Traditional methods for exploring academic literature, such as manual reviews, keyword searching, and citation analysis, prove insufficient when applied to the large research corpora. These approaches fail to provide a clear overview of the overarching theme or the semantic connections between papers<sup>[4]</sup>. Inspired by the CHI 2024 Papers Explorer<sup>[1]</sup> and Movva et al.'s work<sup>[2]</sup>, we design an interactive tool (called "ISee") that combines visualization with search and filtering features to support researchers in exploring scientific literature.

Our tool comprises three main components (see Figure 1). The primary component ① is an interactive map presenting all papers in the current corpus. We first generate semantic embeddings from the titles and abstracts using a language model, and then apply the UMAP algorithm to reduce the high dimensionality for visualization. Each paper is positioned based on semantic similarity, calculated via cosine similarity of their embedding vectors. Users can explore this map by zooming in, selecting papers, and color-coding them by relevance, cluster, year, or venue. This component uses K-means to group similar papers. Users can dynamically color-code papers based on attributes such as relevance score, cluster assignment, or publication year. Additionally, we include a robust semantic search feature that allows users to query the corpus, with keyword highlighting<sup>[3]</sup> and adjustable similarity thresholds.

The second component ② offers keyword searching across titles and abstracts, featuring result highlighting<sup>[3]</sup> and relevance sorting. The component also includes advanced filtering options (year, venue, author, research area) and a bar chart showing the number of papers in the corpus by publication year.

The third component ③ tracks user activity, featuring distinct sections for recently viewed and bookmarked

papers for easy access. It also includes a chronological timeline logging key interactions (searches, filters, paper views, bookmarks) with timestamps.



**Figure 1 – The three components of ISee:** ① a semantic paper visualization with clustering; ② advanced search with keyword highlighting and year-wise bar chart; and ③ a history dashboard to show viewed papers, bookmarks, and recent activity.

## PROJECT WEBSITE

ISee is available at [georg11a.github.io/visearch\\_explorer/](https://georg11a.github.io/visearch_explorer/).

## REFERENCE

1. John Alexis Guerra Gómez. *CHI 2024 Papers Explorer*. <https://johnnguerra.co/viz/chi2024Papers/> (2024).
2. Rajiv Movva, Sidhika Balachandar, Kenny Peng, Gabriel Agostini, Nikhil Garg, Emma Pierson. *Topics, Authors, and Institutions in Large Language Model Research: Trends from 17K arXiv Papers*. arXiv:2307.10700 [cs.DL]. (2023).
3. Hui Zhang. *Beyond Query-Oriented Highlighting: Investigating the Effect of Snippet Text Highlighting in Search User Behavior*. Computational Intelligence and Neuroscience, Article ID 7836969. (2018).
4. Daniel A. Keim. *Information Visualization and Visual Data Mining*. IEEE Transactions on Visualization and Computer Graphics, 8(1), 1–8, 2002.



## Infographics: What Do We Mean and What Do They Show?

Xiaoyu Liu, Diego Lopez\*, Shalin Muzumdar\*, Vraj Patel\*, Vishnu Sreekanth\*,  
Fumeng Yang, Zhicheng Liu

Department of Computer Science, University of Maryland

\* These authors contributed equally

Contact: dasqxl@umd.edu

### OVERVIEW

Infographics, or information graphics, is a prevalent form of information visualization that emerges in research, education and popular media [1]. However, there has not been a clear or uniform definition of infographics, nor is there a systematic analysis of the design space of infographics. In existing works, the term “infographics” has been used to refer to a wide variety of visual contents, sometimes even in contradicting ways. Previous attempts at a design space analysis were also often partial and focused only on certain aspects or subsets of infographics [2,3,4]. A clearer definition of infographics and a more comprehensive analysis of its design space will allow researchers and visualization designers to produce more effective infographics and better tailor them to the appropriate domains and audiences.

Our study investigates the following research questions:

- What do people mean when they use the term infographics?
- How do people perceive infographics in relation to other related concepts (e.g., charts, data visualization, and data comics)?
- In what ways do people design and use infographics in their work?

### METHODS

Our study consists of two phases: a literature review and an online expert survey.

For the literature review, we start with a keyword search in major visualization conferences and journals (IEEE, ACM, EuroVis and Graphics Interface) for publications that mention the term “infographics” in their title and/or abstract. We then manually extract any statement and figure within each publication that, with respect to infographics, either gives a definition (*Meaning*), provides examples (*Example*), or compares and contrasts them against other types of data visualization (*Comparison*). Finally, we perform a two-round inductive coding to highlight emerging themes and conflicts in these statements.

For the expert survey, we will distribute an online questionnaire through curated public mailing lists, Slack workspaces and social media platforms which we believe provide outreach to domain professionals in science

research, business and education. The survey will gather people’s opinions on the meaning, boundary and application of infographics to complement the findings of our literature review.

### PRELIMINARY RESULTS

In the literature review phase, we extracted 514 statements and 113 infographics-related figures from 111 unique publications. Based on this corpus, our inductive coding yielded a total of 104 codes, each originating from one of three statement types (*Meaning*, *Example*, *Comparison*) and falling into one of six code categories (*Design*, *Goal*, *Content*, *Application*, *Authoring*, *Comparison*). Two codes may be marked as *Similar*, *Related*, or *Contrary* to each other according to their arguments. Some example codes from each category are given below.

- Design: “*Infographics contain visual embellishments superimposed over more basic visualizations.*” (#2)
- Goal: “*Infographics have a certain high-level narrative goal.*” (#19)
- Content: “*Infographics require readers to think actively about it to get the intended message.*” (#30)
- Application: “*Infographics are suitable for data communication.*” (#100)
- Authoring: “*Infographics authoring is a type of graphics design.*” (#99)
- Comparison: “*Infographics are a different type of data visualization from charts and data comics.*” (#88)

For the next steps, we will conduct further analysis of the coding results and administer the online expert survey.

### REFERENCES

- X. Lan, *et al.*, “Smile or Scowl? Looking at Infographic Design Through the Affective Lens,” in *TVCG* (2021). doi: 10.1109/TVCG.2021.3074582.
- W. Cui, *et al.*, “Text-to-Viz: Automatic Generation of Infographics from Proportion-Related Natural Language Statements,” in *TVCG* (2019). doi: 10.1109/TVCG.2019.2934785.
- N. Diakopoulos, *et al.*, “Playable Data: Characterizing the Design Space of Game-y Infographics,” in *CHI* (2011). doi: 10.1145/1978942.1979193.
- Y. Wang, *et al.*, “Animated Presentation of Static Infographics with InfoMotion,” in *Computer Graphics Forum* (2021). doi: 10.1111/cgf.14325.

## Rolling into the Future: Navigating Career Pathways with a GenAI-Powered Pool World Model

Ziyi Wang, Zijian Ding  
College of Information, University of Maryland  
Contact: [zoewang@umd.edu](mailto:zoewang@umd.edu)



**Figure 1** Interface of pool world model for career event simulation: When the user pockets a random event ball, the system displays the corresponding random event and related decision-making opportunities. When the player pockets a big event ball, the right panel generates a new big event based on previously occurred events, along with five new random events.

Amid rapid advancements in generative AI, traditional professions—such as software development and customer service—are undergoing partial automation of routine tasks. This shift necessitates more dynamic and frequent career planning and pivoting, especially for individuals at the early stages of their professional journeys.

To address this challenge, we leverage the generative capabilities of large language models (LLMs) to simulate career trajectories within a "pool" world model. Framed as an interactive pool game, the system lowers the learning curve and promotes user engagement. Each ball represents a career milestone or decision point, and pocketing a ball

signifies the occurrence of an event that shapes career development. This metaphorical interface transforms abstract career planning concepts into tangible interactions. However, LLM-generated simulations often inherit survivorship bias from their pre-training data, which tend to overrepresent successful career development cases found online. To mitigate this, we fine-tune GPT-4.1 using real-world examples of both career successes and failures, collected through in-person interviews, to ensure a more balanced and realistic simulation. Beyond serving as a platform for exploring career dynamics, this system can also function as a testbed for studying new modes of interaction between humans and generative systems.





# Human-Computer Interaction Lab

## 42<sup>nd</sup> Annual Symposium



## Read, Watch, Play: From Pedagogy to Playwise

Ananya Kaipa, Samarth A, Samita Belliganood, Sanjay Balaji, Yabing Yang, & the RWP Faculty Team  
College of Information

[carowp@umd.edu](mailto:carowp@umd.edu), [akaipa@umd.edu](mailto:akaipa@umd.edu), [sashwath@umd.edu](mailto:sashwath@umd.edu), [samita@umd.edu](mailto:samita@umd.edu), [sanjayb7@umd.edu](mailto:sanjayb7@umd.edu),  
[yyang116@umd.edu](mailto:yyang116@umd.edu)

### OVERVIEW

The Read, Watch, Play (RWP) approach was developed, implemented, and iterated by Dr. Williams-Pierce over several years, who used the pedagogy to teach a variety of courses, ranging from: fully asynchronous, fully synchronous, blended; middle school mathematics teaching, user-centered design, game design, fundamentals of HCI; undergraduate, graduate, doctoral levels; and so on. One key drawback with using this approach is that online teaching platforms are commonly designed for more traditional teaching approaches, with features that directly conflict with RWP principles of failure and feedback. In addition, when supporting other faculty in using the RWP approach, Dr. Williams-Pierce began wondering: why do we have *learning management systems* that don't have *teaching management* designed in as well? In other words, just as online systems can be designed to support good learning, they can be designed to support good teaching!

As a result, Dr. Williams-Pierce gathered a cross-disciplinary, cross-campus team that applied for and received a 3-year grant to design, prototype, and test an RWP-based Teaching and Learning Management System (TLMS). We have spent the entire 2024-2025 academic year conducting research, designing, prototyping, and user-testing the RWP TLMS, now called *Playwise*.

### TLMS DESIGN GOAL

First, we have all taken at least one class with Dr. Williams-Pierce using the RWP pedagogy, so we were already familiar with the experience from the graduate student side. She additionally shared some files (video, text, slides, and an essay [1]) that described the grant and the pedagogy, but focused more on the big picture of designing a course syllabus using RWP, and essentially told us: design a mobile-first TLMS that feels playful, ignores grades as much as possible, and is simple and intuitive to use. She also recommended treating students as colleagues in a workplace, so they can see almost everything the instructor sees, and

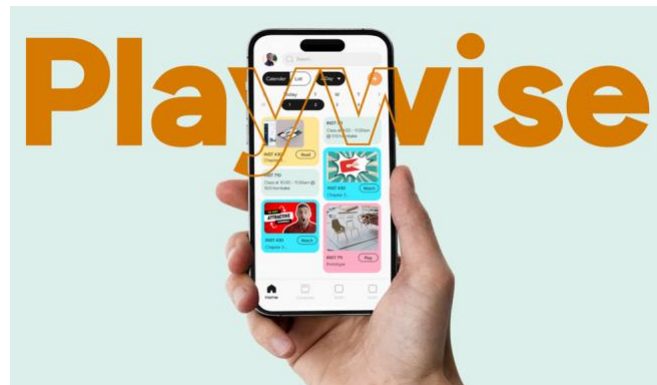


Figure 1 – A mockup of the TLMS.

share their own ideas and expertise in ways that are difficult to do in current teaching platforms.

Based on this background and guidance, we conducted interviews with students, faculty, and teaching assistants; created initial designs, tested them with students and faculty; iterated, tested again, and are now compiling our high fidelity prototypes and a design system which can be used for the project going forward. At the Symposium, we look forward to sharing details about our design process and final Playwise prototype.

### Acknowledgements

Authors alphabetical by first name. Funding for this work was provided by a UMD's Teaching & Learning Transformation Center's Teaching Innovation grant. Many thanks to the iConsultancy for their support with organizing capstones! The full RWP Faculty Team includes: Caro Williams-Pierce, Anat Szendro Sevilla, Emily Dacquist, Andrew Fellows, Katherine Iszak, Pamela Duffy, Kendall Williams, Roger Eastman, Jason Aston, and Daniel Greene.

### PAPERS

1. Williams-Pierce, C. (2021, October). *Loud Gleeful Learning*. Sparks of Innovations: Stories from the HCIL. <https://medium.com/p/86f603f27497>



## The Ethics of Engagement: When Gamification of Learning Works *Too Well*

Uhjin Sim, Yabing Yang, Ting-Wei Hsu, Dimple Chandnani Dolwani, Saie Ganeshkumar Wable,  
Ruibin Zheng, Caro Williams-Pierce

University of Maryland College of Information

[uhjsim@umd.edu](mailto:uhjsim@umd.edu), [yyang116@umd.edu](mailto:yyang116@umd.edu), [twhsu99@umd.edu](mailto:twhsu99@umd.edu), [dimple@umd.edu](mailto:dimple@umd.edu), [saiew23@umd.edu](mailto:saiew23@umd.edu),  
[rzheng13@umd.edu](mailto:rzheng13@umd.edu), [carowp@umd.edu](mailto:carowp@umd.edu)

### OVERVIEW

This dangerously interesting idea emerged from a class activity in INST730: *Games as Emergent Experiences*, where we explored gambling mechanics in game design. Our two groups independently devised satirical takes on gamifying education. We've merged these into a single pedagogical model that uses gambling-inspired strategies to boost student engagement, which raises sharp questions about the ethics of such engagement. Gamification is often widely celebrated, but what happens when it's poorly executed? Or worse—executed *too well*? Could a hyper-effective gamification approach distort the ethical and moral compass of the education system? We propose a hypothetical design study to investigate these tensions. While gamification is framed as a positive motivator, what happens when it starts to mirror the exploitative, competitive behaviors that learners—as “humans”—are already wired to navigate? And how does our current grading system reinforce elements of gambling, risk, and manipulation? We aim to spark critical reflection on the role of competition, incentives, and systemic pressures in higher education under the guise of increased engagement, and question whether gamification is truly a panacea for learning design.

### Background

Gamification—that is, introducing game-like structures, mechanics, and rewards in otherwise non-game-based environments and interactions—has a complex history. Educators use gamification strategies to increase engagement with the content and reward specific behaviors considered to be the hallmark of a ‘good student.’ These strategies include well-accepted approaches that many would be surprised to learn are a form of gamification—such as grading student work, which rewards them with points, or allowing students to choose the topic of their paper to increase their intrinsic motivation and improve their writing (which is then rewarded with points). In other areas, gamification is increasingly commonly used by companies

to increase monetized engagement and spending, such as loyalty programs, wheels the user ‘spins’ in an online storefront, or microtransaction environments posing as ‘casual games.’ In both cases, different design strategies for gaming and gambling are used to increase engagement and elicit specific behaviors (whether learning or spending). In other words, gamification is a design tool that walks a fine line between *supporting* and *manipulating* human beings. And it can go very, very wrong.

### Our Excellent, Incredible, Very Good Idea

We propose a restructuring of formal learning environments to take full advantage of the power of gamification. After all, if increasing engagement is increasing learning, then we should increase engagement in every way possible, right? So we propose a tool that empowers students to bet on other students’ learning-adjacent behavior. As sports gambling is considered to increase the engagement of fans with their teams, we anticipate that increasing engagement within the classroom community will support the development of a powerful peer network.

First, this tool can be used to bet on peers’ grades. If you place a bet on Yabing submitting her paper on time, you will work to support her in her learning and writing efforts. If she receives a cut of the winnings after submitting her paper on time, then she’ll be rewarded for being a good student too! Second, this tool can be used to place wagers on each other’s classroom attendance, specifically betting on whether a classmate will show up and/or arrive on time. To incentivize punctuality, students who arrive on time could be financially rewarded, boosting overall attendance, while also introducing opportunities for strategic interference that tap into human nature to subtly influence outcomes. In other words, this will be hilarious.

### Acknowledgements

Many thanks to our INST730 cohort for their valuable discussions and contributions throughout this semester!



## The Shape of Trust in Formal Education

Caro Williams-Pierce, Uhjin Sim, & Hayley Park

College of Information  
carowp@umd.edu

### OVERVIEW

Much of the first author's research has to do with the productive nature of play for learning, which she started by studying how play and learning are intertwined in well-designed video games, particularly focusing on mathematics video games and youth. Now, however, she is examining what role *trust* has in play-based undergraduate and graduate classrooms. Trust seems particularly important in play-based formal contexts, because such experiences are high stakes (you must do well in this course to complete your degree), don't feel as voluntary as playing a video game (often, you don't even get to choose your instructor), and because good play in classrooms involves *public risk-taking* [1]. In other words, the students need to trust that if they take a risk and fail, they will be appreciated for their risk-taking and creativity, not punished for their failure.

### Our Trust Model

The Formal Education Trust Model (Figure 1) is adapted from a model of trust variables for everyday life [2]. Weiss and colleagues combined various lines of research on trust to identify key potential contributors to the experience of trust. We then adapted that model to a particular context: formal education. For example, in the top of the original model, Weiss and colleagues included an *Interindividual Differences/Personality* part of trust: that is, when a trustor walks into a new situation, they bring with them their own *dispositional propensity to trust*. This dispositional propensity has a cross-culturally validated general trust measure that is widely accepted in the fields of philosophy and psychology of trust, but is focused on trust in everyday social interactions. In the different context of formal education, we posited that students' trust would also be influenced by their prior educational history, and – unlike their stable dispositional propensity to trust – their educational history would change over the semester, as they continued taking classes. We similarly adapted the more general trust model to the specific context of formal education, which we will describe more in the presentation.

### Methodology & Analysis

We collected data weekly in two sections of a first-semester HCIM class, taught by the first author. Each week had 1-3 surveys, and each survey targeted one part of a green bubble in Figure 1. Each survey except the general trust measure

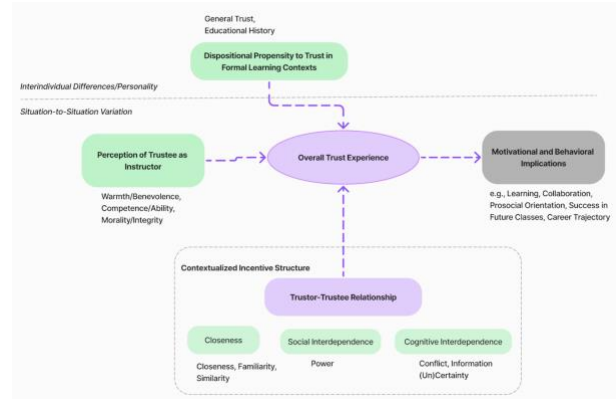


Figure 1 – Formal Education Trust Model.

was administered multiple times over the course of the semester, in order to identify how the different components of trust ebbed and flowed over time and increased familiarity with the socio-educational expectations of the course and program. The data was then transcribed into digital form, and we are in the process of our preliminary analysis, which we plan to report during the presentation.

### Next Steps

Our analysis is focusing on the trends across the semester, both by survey type (e.g., “Perception of Trustee as Instructor” across the semester), as well as by clustering students by their dispositional propensity to trust to see how initial differences manifest differently over time. Future work will connect our findings on trust more directly with learning.

### PAPERS

- Williams-Pierce, C., Deep, S., Ashwathanarayana, S., Gannamaneni, R., Horne, D., Bloodworth, A., Yasuda, S., & Ellis, A.B. (in press). Implementing playful mathematics tasks in calculus sections. In S. Cook, B. Katz, & D. Moore-Russo (Eds.), *Proceedings of the 27th Annual Conference on Research in Undergraduate Mathematics Education*. Alexandria, VA: SIGMAA.
- Weiss, A., Michels, C., Burgmer, P., Mussweiler, T., Ockenfels, A., & Hofmann, W. (2021). Trust in everyday life. *Journal of Personality and Social Psychology*, 121(1), 95–114. <https://doi.org/10.1037/>

## HCIL Honors and Awards

On the following pages, we celebrate some of the awards and honors that HCIL faculty and students have received in the last year.

Promotion & Tenure	
Joel Chan	Received tenure in 2025 (INFO)
Naeemul Hassan	Received tenure in 2025 (INFO & Journalism)
Tamara Clegg	Promoted to full professor in 2025 (INFO)

Awards for research papers	
Nathan Reitinger	Honorable Mention, IAPP's Privacy Papers for Policymakers Student Paper Award: "Artificial Intelligence is like a Perpetual Stew"
Team of authors, including Michelle Mazurek	IAPP's Privacy Papers for Policymakers Student Paper Award: "Data Subjects' Reactions to Exercising Their Right of Access"
Team of authors, including Zhicheng Liu and Chen Chen	IEEE VIS Best Paper Honorable Mention: "Manipulable Semantic Components: A Computational Representation of Data Visualization Scenes"
Team of authors, including Junhyung Moon and Eun Kyoung Choe	Best Paper Award at ACM CHI: "FluidTrack: Investigating Child-Parent Collaborative Tracking for Pediatric Voiding Dysfunction Management"
Team of authors, including Amelia Short, Ruipu Hu, Eun Kyoung Choe, I Hernisa Kacorri, Shannon Jette, Beth Barnett, and Amanda Lazar	Best Paper Award at ACM CHI: "Tracking and its Potential for Older Adults with Memory Concerns"
Team of authors, including Stephanie Valencia	Best Paper Honorable Mention at ACM CHI: "Why So Serious? Exploring Timely Humorous Comments in AAC Through AI-Powered Interfaces"
Team of authors, including Huaishu Peng and Zeyu Yan	Best Paper Honorable Mention at ACM CHI: "PCB Renewal: Iterative Reuse of PCB Substrates for Sustainable Electronic Making"

Other Awards & Achievements	
Jessica Vitak	General Chair, ACM CSCW 2025
Michelle Mazurek	Organizer, 6th Capital-Area Colloquium on Trustworthy and Usable Security/Privacy (CACTUS/P)
Susannah Paletz	Organizer, 2025 INGRoup Midyear Conference
Eun Kyoung Choe	UW iSchool Distinguished Alumni Award
Jonathan Lazar	Accessibility Initiatives Award, International Association of Accessibility Professionals

## HCIL Student Graduation

Finally, we celebrate the many HCIL students who have graduated over the last year. Students are an integral part of the lab's success, and these students have worked with faculty on a range of important research projects in recent years. Below is a partial list. We offer congratulations and good luck to all students graduating who have been a part of the HCIL during their studies!

Student Name	Advisor	Grad Date	Degree
Pramod Chundury	Jonathan Lazar / Niklas Elmqvist	Spring 2025	PhD (INFO)
Celia Chen	Jennifer Golbeck	Spring 2025	PhD (INFO)
Rie Kamikubo	Hernissa Kacorri	Spring 2025	PhD (INFO)
Siggie Gold	Wayne Lutters	Summer 2024	PhD (INFO)
Md Naimul Hoque	Kari Kraus	Fall 2024	PhD (INFO)
Yuhang Zhou	Wei Ai	Spring 2025	PhD (INFO)
Mohammad Ali	Naeemul Hassan	Spring 2025	PhD (INFO)
Chen Chen	Zhicheng Liu	Fall 2024	PhD (CS)
Hannah Bako	Zhicheng Liu	Spring 2025	PhD (CS)
Yuexi (Tracy) Chen	Zhicheng Liu	Spring 2025	PhD (CS)
Lei Mao	Stephanie Valencia	Spring 2025	HCIM
Kriti Singh	Alex Leitch	Spring 2025	HCIM
Xinyi Hu	mols sauter	Spring 2025	HCIM
Zainab Shamim	Hernisa Kacorri	Spring 2025	HCIM

[illegible]

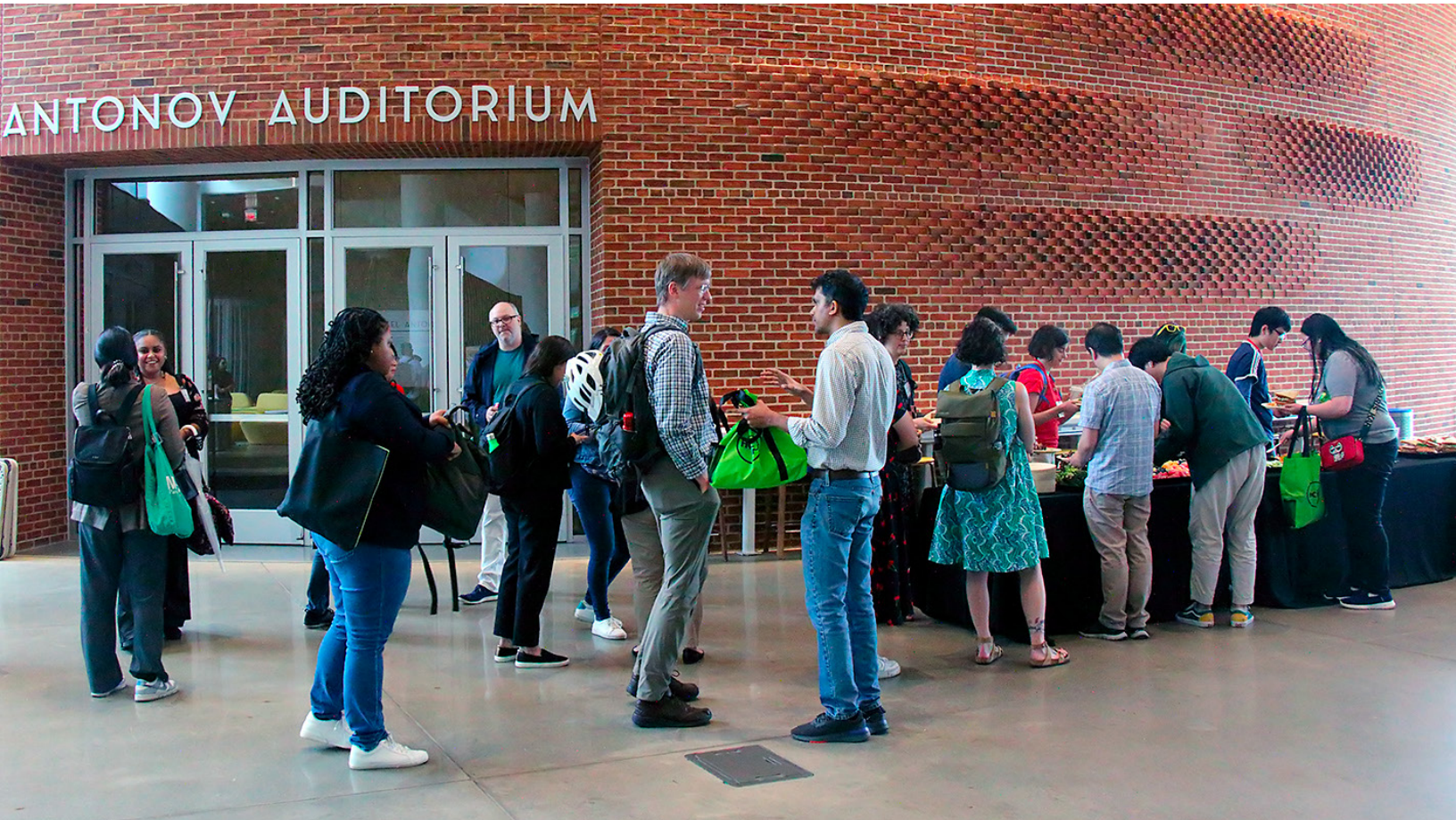
[illegible]





UNIVERSITY OF  
MARYLAND

**hcil.umd.edu**



**Human-Computer Interaction Lab**  
4130 Campus Drive,  
Hornbake Building, South Wing, Room 2105  
College Park, MD 20742  
[hcil-info@umd.edu](mailto:hcil-info@umd.edu)



COLLEGE OF  
INFORMATION



**UMIACS**

University of Maryland Institute for Advanced Computer Studies