# Marx Wang

**PhD Candidate, AI Alignment/Human-Computer Interaction**
**University of Washington | Seattle, WA**
marxwang@uw.edu | (540) 998-5525 | marxvwang.com | linkedin.com/in/marxw

My research sits at the intersection of HCI and LLM Alignment, examining relational and affective dynamics in human-AI interaction through mixed methods: qualitative conversation analysis grounded in developmental psychology, computational psycholinguistics/NLP (LIWC, NRC-VAD, RoBERTa), and mechanistic interpretability using Sparse Autoencoders to identify and steer affect-aligned features in LLM activations. My work spans both empirical studies of how users engage with conversational agents and technical methods for auditing and modulating the emotional behavior of language models.

## Education

August 2023 - Present
**University of Washington,** Seattle, WA
Ph.D., Information School
Research Focus: Applied AI, Conversational Agents, Affective Computing
Advisors: Dr. Alexis Hiniker, Dr. Anind K Dey

August 2021 - May 2023
**Virginia Tech,** Blacksburg, VA
Master of Science (MS), Computer Science and Application, GPA 4/4
Thesis: Understanding How Users Express Their Social Identities through Animated Gif for Public Discourse
Advisors: Dr. Sang Won Lee (Chair), Dr. Kurt Luther, Dr. Eugenia Rho

August 2018 - May 2021
**Virginia Tech,** Blacksburg, VA
B.S. in Computer Science, Data Centric Computing, GPA 3.98/4
Summa Cum Laude

## Technical Background

**AI/Machine Learning & Data Analytics**
LLM, NLP, CV with (PyTorch, Tensorflow, Transformers, Diffusers, Datasets, TRL, SK-Learn, Pandas, Keras, NumPy, Gensim, GloVe, fastText, SciPy, StyleGAN, Seaborn, Matplotlib)
Relational Database Management Systems: SQL, MySQL, PostgreSQL, Mongodb;

**Full Stack Software Development**
Python, R, C++, C

JavaScript with NodeJs, NextJs, React, Redux;
Django, Java, PHP; Unity, HTML5, CSS, Bootstrap;
Blockchain Dapps with Solidity, Smart Contracts;
Cloud Computing (AWS, Heroku, Azure) and Third-Party APIs (Twitter APIs, Auth0 APIs,
Google APIs, Giphy APIs)

**UI/UX Design Research & Usability Engineering**
Interface Design, User Study, Interview, and Evaluation;
Interaction Design, Wireframe & Prototype Design, Conceptual Model Design, Contextual Analysis,
Storyboarding, Heuristic Evaluation; Figma, Sketch, Framer X, Cinema 4D, Photoshop, InDesign, Premiere
Pro, After Effect, Illustrator;

# Publications

## Peer-reviewed Publications ([J]ournal and [C]onference proceedings)

[C.1] **Marx Wang** et al. "Understanding the relationship between social identity and self-expression through animated gifs on social media." Proceedings of the ACM on Human-Computer Interaction 8.CSCW1 (2024): 1-30.

[C.2] **Marx Wang**\*, Daniel Manesh\*, Ruipu Hu, and Sang Won Lee. Piggybacking on IFTTT to Connect Triggers and Actions with JavaScript." 2024 IEEE Symposium on Visual Languages and Human-Centric Computing (VL/HCC). IEEE, 2024. (\*Equal Contribution)

[C.3] Yue Fu, Tony Zhou, Bin Han, **Marx Wang,** Yixin Chen, Zelia Gomes Da Costa Lai, Rock Pang, Katharina Reinecke, Jacob Wobbrock, Alexis Hiniker, Decoupling of Usefulness and Novelty: Evaluating the Impact of Generative AI on Design Outputs and Designers' Creative Thinking Proceedings of the ACM on Human-Computer Interaction 8.CHI, 2026.

[J.1] Yuan Li, Donghan Hu, **Marx Wang**, Doug A Bowman, and Sang Won Lee. 2021.The Effects of Incorrect Occlusion Cues on the Understanding of Bare-handed Referencing in Collaborative Augmented Reality. Frontiers in Virtual Reality. https://doi.org/10.3389/frvir.2021.681585

[C.4] Manisha Kusuma, Vikram Mohanty, **Marx Wang**, and Kurt Luther. 2022. Civil War Twin: Exploring Ethical Challenges in Designing an Educational Face Recognition Application. In Proceedings of the 2022 AAAI/ACM Conference on AI, Ethics, and Society (AIES '22). Association for Computing Machinery, New York, NY, USA, 369–384. https://doi.org/10.1145/3514094.3534141

[P.1] **Marx Wang** and Sang Won Lee. 2022. TaskScape: Fostering Holistic View on To-do List With Tracking Plan and Emotion. In The Adjunct Publication of the 35th Annual ACM Symposium on User Interface Software and Technology (UIST '22 Adjunct). Association for Computing Machinery, New York, NY, USA, Article 31, 1–4. https://doi.org/10.1145/3526114.3558720

[P.2] **Marx Wang**, Zachary Duer, Scotty Hardwig, Sam Lally, Alayna Ricard, and Myounghoon Jeon. 2022. Echofluid: An Interface for Remote Choreography Learning and Co-creation Using Machine Learning Techniques. In The Adjunct Publication of the 35th Annual ACM Symposium on User Interface Software and Technology (UIST '22 Adjunct). Association for Computing Machinery, New York, NY, USA, Article 22, 1–4. https://doi.org/10.1145/3526114.3558708

[P.3] **Marx Wang**, Daniel Manesh, Ruipu Hu, and Sang Won Lee. 2022. iThem: Programming Internet of Things Beyond Trigger-Action Pattern. In The Adjunct Publication of the 35th Annual ACM Symposium on User Interface Software and Technology (UIST '22 Adjunct). Association for Computing Machinery, New York, NY, USA, Article 53, 1–5. https://doi.org/10.1145/3526114.3558776

[P.4] **Marx Wang**, Ya-Yun Chen, Tae-Ho Lee. 2022. Predicting Prosocial Behavior in Childhood and Early Adolescence by Integrating Parental Factors, Neighborhoods and The Brain. Biennial Meeting of the Society for Research on Adolescence.

[P.5] Ya-Yun Chen, **Marx Wang,** Tae-Ho Lee. 2022. Predicting Aggressive Behavior in Childhood and Early Adolescence with the Socioecological Approach: From Neighborhoods, and The Brain. The 14th Annual Meeting of the Social & Affective Neuroscience Society.

[P.6] Viral Pasad, **Marx Wang,** and Sang Won Lee. 2020. Understanding the Challenges of Online Group Chat for Productive Discourse at Scale. In Conference Companion Publication of the 2020 on Computer Supported Cooperative Work and Social Computing (CSCW '20 Companion). Association for Computing Machinery, New York, NY, USA, 365–369. https://doi.org/10.1145/3406865.3418335

[W.1] Vikram Mohanty, **Marx Wang,** David Thames, Manisha Kusuma, and Kurt Luther. 2020. Find Your Twin in History: Exploring Ethical Design Challenges in Facial Recognition. In Conference Companion Publication of the 2020 on Computer Supported Cooperative Work and Social Computing (CSCW '20 Companion). Association for Computing Machinery, New York, NY, USA, 365–369. (Ethics in Design Workshop.)

## Under Submission in Peer-Reviewed Venues
**Marx Wang,** Songling Ngo, Raghavi Putluri, Rachel A. Hanebutt, Shirin Amouei, Robert Wolfe, Katie Davis, Alexis Hiniker 2025. How Young People are Using ChatGPT for Social Support

**Marx Wang,** Songling Ngo, Raghavi Putluri, Rachel A. Hanebutt, Shirin Amouei, Robert Wolfe, Katie Davis, Alexis Hiniker 2025. Botsplaining Human Vulnerability: An In-Depth Case Study of Youth Self Disclosure with ChatGPT

**Marx Wang,** Robert Wolfe, Songling Ngo, Raghavi Putluri, Alexis Hiniker 2025. From Emotional Mirroring to Emotional Attunement: Do LLMs and Humans Attune to Each Other

**Marx Wang,** Robert Wolfe, Alexis Hiniker 2026. Do Attachment Styles Shape Conversational Agent Usage?

# Experience

Fall 2023 - Present

**Research Assistant/Senior Data Scientist, User Empowerment Lab,** *University of Washington*

- Developed mechanistic interpretability methods to extract and validate linear "truth directions" from Llama hidden states, implementing claim-boundary token extraction, LDA-based probing, and ghost direction artifact removal to isolate semantic content from framing effects.
- Engineered a mechanical steering framework for Llama 3.2 to disentangle and manipulate causal directions of model sentiment. Implemented Sparse Autoencoders (SAEs) for feature extraction and diagnosed critical non-robustness in cross-layer steering, resolving signal degradation by replacing correlational feature selection with causal Contrastive Activation Addition.
- Designed computational psycholinguistic approaches and sentiment analysis pipelines to evaluate LLM capacity for emotional attunement, identifying failures in sustaining empathy across multi-turn conversations.
- Analyzed 14,000+ conversation threads to develop a taxonomy of how youth use LLMs for the "transition to adulthood," identifying design patterns in Cognitive Extension (offloading "adulting" tasks), Social Simulation, and Self-Stabilization (emotional regulation).
- Conducted in-depth qualitative analysis of youth self-disclosure to characterize "AI Botsplaining," identifying specific failure modes where models prioritize instrumental logic (e.g., premature problem-solving, tone policing) over relational needs.
- Investigating personalized computing and human-AI attachment, specifically how standard RLHF reward models can inadvertently reinforce maladaptive dependency and accelerating usage behaviors in anxious users.
- Collected and computationally analyzed a dataset of 200k+ in-the-wild ChatGPT interactions to fine-tune models for better detection of vulnerable user states.

Fall 2019 - 2023

**Research Assistant/Software Engineer, EchoLab,** *Virginia Tech*

- Led a team of 3, designed and implemented a machine learning pipeline that visualizes user emotions through auto-suggesting gifs in real time based on messages in ongoing conversation with NLP and CNN. Achieved 95.6% accuracy rate in composite models of multiple high-level linguistic features such as empathy, sarcasm [C.3].

- Organized user studies and implemented high fidelity twitter like study programs for ecological validity. Applied sociolinguistic techniques such as matched disguised tests. Built web scraping scripts and collected a dataset with approximately 1000000 gifs. Extended studies to multiple computational sociolinguistic studies in multimodal natural language processing and search engine bias [C.3].

- Designed user studies with real time collaborative tasks to analyze users' behavior patterns and coping strategy in disorganized group chat. Conducted thematic analysis, moderated user studies, analyzed large scale data. Engineered 6 novel chat systems in MeteorJS with psychology-backed features such as live keystrokes to realize co-temporality [P.6].

- Heuristically evaluated the UX of Microsoft HoloLens. Developed synchronous web applications for VR studies without any framework (with JavaScript, PHP, MySQL, Apache) Thematically analyzed users' compensatory spatial referencing strategies in AR conditions [J.1].

Jan 2020 - 2023
### Research Assistant/UX Designer, Mind Music Machine Lab, *Virginia Tech*

- Designed mapping systems to computationally capture emotional experiences and cognitive messages of art and translating visual features (e.g. painting structure ) to auditory parameters (e.g. tempo).

- Designed and implemented EchoFluid, an AI-powered choreography tool that allows dance learners to easily build their AI models for practices and co-creation. Implemented the algorithm with reinforcement learning for real time dance movement analysis and developed the interface with Unity. Presented the work at UIST2022 [P.2].

June 2021 - Present
### Research Assistant/Software Engineer, Affective Neurodynamics and Development Lab, *Virginia Tech*

- Used machine learning techniques to forecast children's prosocial behaviors, improved model accuracy from 55% to 85%, created data visualization to effectively communicate findings to various stakeholders. Presented work at SANS 2021 [P.4, P.5].

- Conducted comprehensive data analyses and regression analyses, applied advanced statistics, and modeled the complex relationship between family, neighbors, and environmental factors and

children's prosocial and aggressive behaviors. Conducted fMRI Scan and analyzed neurobiological data [C.5, C.6].

June 2021 - 2023

**Research Assistant/Software Engineer, Crowd Intelligence Lab***, Virginia Tech*

- Conducted user research to understand elderly users' privacy concerns, designed effective visualizations and detailed gradual steps to clearly communicate underlying ethical challenges in data privacy, transparency, security, gender and racial bias, while effectively educating AI concepts to an older audience, designed selected to CSCW 2021 workshop and published in AIES 2022 Conference [C.1].

- Designed UI/UX of Civil War Twin website, a historical educational website that allows people to use AI to find their look-alike in the Civil War era. Implemented high fidelity frontend with animation in JavaScript to fulfill innovative frontend implementation requirements, coordinated with the lead developer and integrated the frontend with Django based backend [P.7].

## Teaching

Fall 2025-2026

**Pre-Doctoral Lecturer, University of Washington**

**INFO 380**: Information Systems and Management (2 quarters)

Fall 2023

**Teaching Assistant, University of Washington**

**INFO 300**: Research Methods (2 quarters)

**INFO 380**: Information Systems and Management (2 quarters)

Fall 2021

**Graduate Teaching Assistant, Virginia Tech**

**User Interface Software,** with Dr. Kurt Luther, 78 students

Spring 2022

**Graduate Teaching Assistant, Virginia Tech**

**Professionalism in Computing,** with Dr. Daniel Dunlap, 200 students

# Leadership & Social Impact

2023-Present

**Principal, ForYouPage.Org (FYP)**

*Youth-led 501(c)(3) nonprofit for mental health and well-being advocacy (EIN: 93-2392340)*

- Built global youth advocacy platform reaching 30,000+ members across 6 continents
- Mobilized 200+ high-impact youth leaders influencing international policy
- Established partnerships with 600+ organizations worldwide
- Direct advisory engagement with prime ministers, UN leadership, and global policymakers

  Major International Conferences:

- UN Science Summit UNGA78-80 (2023-2025): Convened 60+ leaders from 40 major organizations including Mental Health America, California Child Trust, HopeLab, UN, citiesRISE, Led youth-driven conference on behavioral health priorities for 2030

2024

**Associate & Board Member, Ourmedia**

*501(c)(3) nonprofit for Gen Z leadership and AI education*

- Spearheaded partnerships with major Silicon Valley firms such as IDEO
- Leading Gen Z advocacy in AI readiness and workforce development
- Pioneering AI education programs for underserved communities
- Building intergenerational knowledge networks
- Direct advisory engagement with tech industry leaders on responsible AI


**Raft Crisis Hotline**, Jan 2022 -Present
**Mock Trial, Regional Best Attorney,** Sept 2021 – Present
**Phi Alpha Delta, International Pre-law Fraternity**, Sept 2021 – Present
**Phi Beta Kappa,** May 2021
**Lambda Iota Mu, Natural Resources and Conservation Service Fraternity**, Jan 2020 – Present
**United Feminist Movement**, Sept 2021 – Present
**Active Minds**, Sept 2020 - Present