# HumanEYEze 2024: Workshop on Eye Tracking for Multimodal Human-Centric Computing


Michael Barz
michael.barz@dfki.de
German Research Center for Artificial Intelligence (DFKI)
Saarbrücken, Germany
University of Oldenburg
Oldenburg, Germany

Roman Bednarik
roman.bednarik@uef.fi
University of Eastern Finland
Joensuu, Finland

Andreas Bulling
andreas.bulling@vis.uni-stuttgart.de
University of Stuttgart
Stuttgart, Germany

Cristina Conati
conati@cs.ubc.ca
University of British Columbia
Vancouver, British Columbia, Canada

Daniel Sonntag
daniel.sonntag@dfki.de
German Research Center for Artificial Intelligence (DFKI)
Saarbrücken, Germany
University of Oldenburg
Oldenburg, Germany



## Abstract

The HumanEYEze 2024 workshop aims to explore the role of eye tracking in developing human-centered multimodal AI systems. Over the past two decades, eye tracking has evolved from a diagnostic tool to an important input modality for real-time interactive systems, driven by advancements in hardware that have improved its affordability, availability, and performance. Initially used in specialized applications, eye tracking now significantly impacts research on gaze-based multimodal interaction. Recently, eye-based user and context modeling has emerged, utilizing eye movements to provide rich insights into user behavior and interaction contexts. The workshop aims to bring together researchers from eye tracking, multimodal human-computer interaction, and AI. It aims to enhance understanding of integrating eye tracking into multimodal human-centered computing. The expected outcomes include fostering collaborations and promoting knowledge exchange.



## CCS Concepts

• **Human-centered computing** → **User models**; **Human computer interaction (HCI)**; • **Computing methodologies** → *Artificial intelligence*; *Machine learning*;

## Keywords

Eye Tracking, Gaze, Multimodal Interaction, User Modeling, Human-centric Computing, Human-centered AI, Workshop




*ICMI '24, November 04–08, 2024, San Jose, Costa Rica*



ACM ISBN 979-8-4007-0462-8/24/11
https://doi.org/10.1145/3678957.3688384

**ACM Reference Format:**
Michael Barz, Roman Bednarik, Andreas Bulling, Cristina Conati, and Daniel Sonntag. 2024. HumanEYEze 2024: Workshop on Eye Tracking for Multimodal Human-Centric Computing. In *INTERNATIONAL CONFERENCE ON MULTIMODAL INTERACTION (ICMI '24), November 04–08, 2024, San Jose, Costa Rica.* ACM, New York, NY, USA, 2 pages. https://doi.org/10.1145/3678957.3688384


## 1 Introduction

Over the last 20 years, eye tracking has evolved from being a diagnostic tool to a powerful input modality for real-time interactive systems. This was partly driven by advances in eye tracking hardware concerning the devices' affordability, availability, performance, and form factor. Eye tracking was first used in niche applications in the '80s and '90s and then gathered significant attention through research on gaze-based interaction and gaze-supported multimodal interaction [8, 12, 15]. In the last 10-15 years, a third very promising direction has emerged: eye-based user and context modeling, i.e., seeing the eyes as an additional modality that provides rich information about user (interactive) behavior and their (interaction) context [3, 6, 10, 13, 14]. The eyes reveal information about visual activities [7], personality [9], user intents and goals [4, 5], attention [2], expertise and other cognitive abilities [1, 11], and emotions [16], to name a few. With that, eye tracking bears great potential for developing human-centered multimodal AI systems. Gaze-based multimodal user models can be used to, e.g., generate direct feedback to steer the training of AI systems or trigger explicit feedback requests (or show model explanations) if the user seems to disagree with the output of an AI system. The goal of this workshop is to bring together researchers from eye tracking, multimodal human-computer interaction, and artificial intelligence. We will welcome contributions on the following topics:

- Methods and systems to analyze everyday eye movement behavior
- Real-time vs. post-hoc analysis and modeling
- Eye tracking in human-centered AI systems

- Adaptive gaze-based and gaze-supported multimodal user interfaces
- Eye-based user modeling with limited data
- Eye tracking for multimodal user modeling
- Eye-supported multimodal activity and context recognition
- Computer vision methods for gaze estimation and multimodal behavior analysis
- Gaze sensing systems - real-world benchmarks, requirements, techniques
- Privacy-preserving eye tracking
- Repositories and datasets
- Focused reviews and meta-analyses

## 2 Expected Outcome and Impact

Our goal is to establish a unique discussion platform for researchers in the field of eye tracking with a focus on eye tracking in multimodal human-centric interfaces and AI as an enabling technology. We aim to bring together researchers and practitioners from the fields of eye tracking, multimodal interaction, machine learning, human-computer interaction, psychology, and other related fields. This shall foster collaborations among researchers in the field and enable a knowledge exchange on the role of eye tracking in multimodal interaction and AI-based systems.

## 3 Workshop Contributions

We included four papers in the workshop proceedings. The topics include studying collaborative interaction behavior based on eye tracking in mixed-reality settings, predicting driving decisions using vision transformers and gaze, modeling when users disagree with the output of generated image captions, and investigating the impact of ambient illumination change on the accuracy of head-mounted eye trackers. The accepted papers include:

- *3D Gaze Tracking for Studying Collaborative Interactions in Mixed-Reality Environments* by Eduardo Davalos, Yike Zhang, Ashwin T S, Joyce Horn Fonteles, Umesh Timalsina, and Gautam Biswas.
- *Gaze-Informed Vision Transformers: Predicting Driving Decisions Under Uncertainty* by Sharath Koorathota, Nikolas Papadopoulos, Jia Li Ma, Shruti Kumar, Xiaoxiao Sun, Arunesh Mittal, Patrick Adelman, and Paul Sajda.
- *Detecting when Users Disagree with Generated Captions* by Omair Shahzad Bhatti, Harshinee Sriram, Abdulrahman Mohamed Selim, Cristina Conati, Michael Barz, and Daniel Sonntag.
- *Investigating the Impact of Illumination Change on the Accuracy of Head-Mounted Eye Trackers: A Protocol and Initial Results* by Mohammadhossein Salari and Roman Bednarik.

## Acknowledgments

This work was funded, in part, by the European Union under grant number 101093079 (MASTER, https://www.master-xr.eu), the German Federal Ministry of Education and Research (BMBF) under grant number 01IW23002 (No-IDLE), by the Lower Saxony Ministry of Science and Culture, and the Endowed Chair of Applied Artificial Intelligence of the University of Oldenburg.

## References

[1] Oswald Barral, Sébastien Lallé, Grigorii Guz, Alireza Iranpour, and Cristina Conati. 2020. Eye-Tracking to Predict User Cognitive Abilities and Performance for User-Adaptive Narrative Visualizations. In *Proceedings of the 2020 International Conference on Multimodal Interaction (ICMI '20)*. Association for Computing Machinery, New York, NY, USA, 163–173. https://doi.org/10.1145/3382507.3418884 event-place: Virtual Event, Netherlands.

[2] Michael Barz, Sebastian Kapp, Jochen Kuhn, and Daniel Sonntag. 2021. Automatic Recognition and Augmentation of Attended Objects in Real-time using Eye Tracking and a Head-mounted Display. In *ACM Symposium on Eye Tracking Research and Applications (ETRA '21 Adjunct)*. Association for Computing Machinery, New York, NY, USA, 1–4. https://doi.org/10.1145/3450341.3458766

[3] Michael Barz and Daniel Sonntag. 2021. Automatic Visual Attention Detection for Mobile Eye Tracking Using Pre-Trained Computer Vision Models and Human Gaze. *Sensors* 21, 12 (Jan. 2021), 4143. https://doi.org/10.3390/s21124143 Number: 12 Publisher: Multidisciplinary Digital Publishing Institute.

[4] Michael Barz, Sven Stauden, and Daniel Sonntag. 2020. Visual Search Target Inference in Natural Interaction Settings with Machine Learning. In *ACM Symposium on Eye Tracking Research and Applications (ETRA '20 Full Papers)*, Andreas Bulling, Anke Huckauf, Eakta Jain, Ralph Radach, and Daniel Weiskopf (Eds.). Association for Computing Machinery, 1–8. https://doi.org/10.1145/3379155.3391314

[5] Roman Bednarik, Hana Vrzakova, and Michal Hradis. 2012. What do you want to do next: a novel approach for intent prediction in gaze-based interaction. In *Proceedings of the Symposium on Eye Tracking Research and Applications (ETRA '12)*. Association for Computing Machinery, New York, NY, USA, 83–90. https://doi.org/10.1145/2168556.2168569 event-place: Santa Barbara, California.

[6] Andreas Bulling, Daniel Roggen, and Gerhard Troester. 2011. What's in the Eyes for Context-Awareness? *IEEE Pervasive Computing* 10, 2 (April 2011), 48–57. https://doi.org/10.1109/MPRV.2010.49 Place: USA Publisher: IEEE Educational Activities Department.

[7] Andreas Bulling, Jamie A. Ward, Hans Gellersen, and Gerhard Tröster. 2009. Eye movement analysis for activity recognition. In *Proceedings of the 11th International Conference on Ubiquitous Computing (UbiComp '09)*. Association for Computing Machinery, New York, NY, USA, 41–50. https://doi.org/10.1145/1620545.1620552 event-place: Orlando, Florida, USA.

[8] Andrew T. Duchowski. 2018. Gaze-based interaction: A 30 year retrospective. *Computers & Graphics* 73 (2018), 59–69. https://doi.org/10.1016/j.cag.2018.04.002

[9] Sabrina Hoppe, Tobias Loetscher, Stephanie A. Morey, and Andreas Bulling. 2018. Eye Movements During Everyday Behavior Predict Personality Traits. *Frontiers in Human Neuroscience* 12 (2018). https://doi.org/10.3389/fnhum.2018.00105

[10] Sébastien Lallé, Dereck Toker, and Cristina Conati. 2021. Gaze-Driven Adaptive Interventions for Magazine-Style Narrative Visualizations. *IEEE Transactions on Visualization and Computer Graphics* 27, 6 (2021), 2941–2952. https://doi.org/10.1109/TVCG.2019.2958540

[11] Yan Liu, Pei-Yun Hsueh, Jennifer Lai, Mirweis Sangin, Marc-Antoine Nussli, and Pierre Dillenbourg. 2009. Who is the expert? Analyzing gaze data to predict expertise level in collaborative applications. In *2009 IEEE International Conference on Multimedia and Expo*. 898–901. https://doi.org/10.1109/ICME.2009.5202640

[12] Päivi Majaranta and Andreas Bulling. 2014. Eye Tracking and Eye-Based Human–Computer Interaction. In *Advances in Physiological Computing*, Stephen H. Fairclough and Kiel Gilleade (Eds.). Springer, London, 39–65. https://doi.org/10.1007/978-1-4471-6392-3_3

[13] Sharon Oviatt, Björn Schuller, Philip Cohen, Daniel Sonntag, Gerasimos Potamianos, and Antonio Krüger (Eds.). 2019. *The Handbook of Multimodal-Multisensor Interfaces: Language Processing, Software, Commercialization, and Emerging Directions*. Association for Computing Machinery and Morgan & Claypool. https://doi.org/10.1145/3233795

[14] Sharon. Oviatt, Björn Schuller, Philip R. Cohen, Daniel Sonntag, Gerasimos Potamianos, and Antonio Krüger (Eds.). 2017. *The Handbook of Multimodal-Multisensor Interfaces: Foundations, User Modeling, and Common Modality Combinations* (volume 1 ed.). Association for Computing Machinery and Morgan & Claypool, New York, NY, USA. https://doi.org/10.1145/3015783

[15] Pernilla Qvarfordt. 2017. Gaze-Informed Multimodal Interaction. In *The Handbook of Multimodal-Multisensor Interfaces: Foundations, User Modeling, and Common Modality Combinations - Volume 1*. Association for Computing Machinery and Morgan & Claypool, 365–402. https://doi.org/10.1145/3015783.3015794

[16] Shane D. Sims and Cristina Conati. 2020. A Neural Architecture for Detecting User Confusion in Eye-tracking Data. In *Proceedings of the 2020 International Conference on Multimodal Interaction (ICMI '20)*. Association for Computing Machinery, New York, NY, USA, 15–23. https://doi.org/10.1145/3382507.3418828 event-place: Virtual Event, Netherlands.