Wearing AI-Powered Gadgets: Exploring Privacy, Ethics, and Design Challenges of First-Person AI Devices

Rintaro Chujo
chujo@hc.ic.i.u-tokyo.ac.jp
The University of Tokyo
Tokyo, Japan

Ayumi Takagi
Hosei University
Tokyo, Japan

Shun Naoi
Hosei University
Tokyo, Japan

Reika Nakamura
Hosei University
Tokyo, Japan

Yuchi Yahagi
The University of Tokyo
Tokyo, Japan

Katsufumi Matsui
The University of Tokyo
Tokyo, Japan

Young ah Seong
Hosei University
Tokyo, Japan

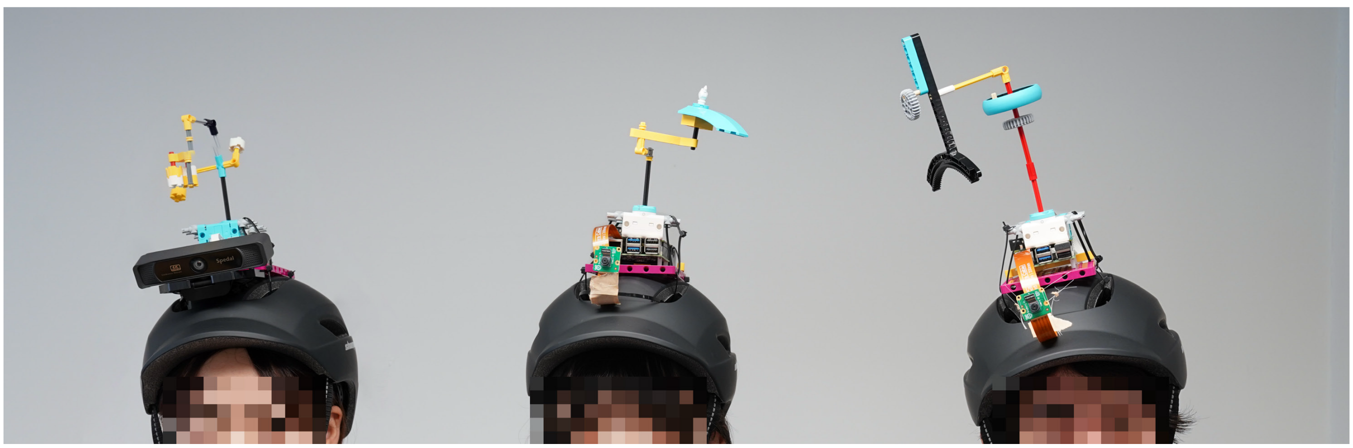

Figure 1: People wearing copters

Abstract

Advances in generative AI enable seamless real-time interactions through visual and auditory inputs, exemplified by technologies such as GPT-4o. Yet, first-person AI introduces significant privacy and ethical challenges. We present "Copter," a speculative head-mounted gadget without practical functionality, to probe everyday human-AI interactions. Through a two-week autobiographical field study with three designers, themes related to human-AI symbiosis emerged. Building on Copter's insight, especially on "projection for the moving of things", a designer developed "MuseBuddy," a diving companion AI gadget. To expand this exploration, we propose a toolkit for diverse designers to prototype similar AI gadgets, facilitating nuanced exploration of the social, ethical, and privacy implications of first-person AI, contributing critical embodied insights to Research through Design.

CCS Concepts

• **Human-centered computing** → **Human computer interaction (HCI)**.

Keywords

Material Speculation, Autonomously Operating Object, Symbiosis, Human-Object Interaction

ACM Reference Format:
Rintaro Chujo, Ayumi Takagi, Shun Naoi, Reika Nakamura, Yuchi Yahagi, Katsufumi Matsui, and Young ah Seong. 2018. Wearing AI-Powered Gadgets: Exploring Privacy, Ethics, and Design Challenges of First-Person AI Devices. In *Proceedings of Make sure to enter the correct conference title from your rights confirmation emai (Conference acronym 'XX)*. ACM, New York, NY, USA, 5 pages. https://doi.org/XXXXXXX.XXXXXXX

Conference acronym 'XX, June 03–05, 2018, Woodstock, NY

ACM ISBN 978-1-4503-XXXX-X/18/06
https://doi.org/XXXXXXX.XXXXXXX

1 Introduction

The remarkable advancement of artificial intelligence technologies, including generative AI, has made it possible for computers to interact with the surrounding world at a natural speed comparable to that of humans. For example, the demo of OpenAI's generative AI "GPT-4o" showcased its ability to engage in conversations while using cameras and microphones to understand the surrounding situation in real-time and at a natural pace. This technology enables computers to process not only text but also visual and auditory information.

One of the applications of generative AI technology is the idea of using first-person perspective videos as input for AI to complement and support human perception. For example, the "Humane Ai Pin" is a small laser projector that can be clipped onto clothing, functioning as a virtual assistant and communication device that responds to user inquiries using a proprietary Large Language Model (LMM) based on OpenAI's GPT-4. Additionally, Meta's "Ray-Ban Meta Glasses" are equipped with a camera that can capture first-person videos, and it has been announced that users will be able to interact with AI based on these first-person videos.

On the other hand, the use of AI with first-person perspective videos presents various design challenges, including concerns about privacy and ethics. For example, Bhardwaj et al. conducted a micro-longitudinal diary study involving users who have and have not used camera-equipped glasses, reporting that the design of current glasses-type devices capable of capturing first-person perspective videos is insufficient for protecting privacy, and they discussed that design space. Additionally, socially, in response to the announcement of Ray-Ban Meta, concerns about privacy have been raised by the International Bar Association, indicating that such concerns are spreading.

This study reports on the record of wearing an AI-powered gadget called "Copter," which operates based on first-person support footage, over a period of two weeks in daily life. The Copter is an AI agent that operates using first-person video, similar to "Ray-Ban Meta Glasses," but it does not possess any functions that are useful to humans; it merely acts as an artifact where motors operate according to AI instructions. This practice is similar to material speculation in the sense that it explores how technology functions in daily life and what challenges exist, using counterfactuals. On the other hand, what this study explores is not the world of if...then statements, but rather the challenges associated with AI that uses first-person perspective footage, which is already quietly spreading in the world.

In this study, along with reports on "Copter," we will also introduce "MuseBuddy," an AI-powered gadget developed by one of the designers who participated in the Copter practice, based on their personality as a diver and operating on first-person underwater footage. Additionally, we propose a toolkit aimed at allowing other designers to experience similar experiences in their daily lives. Through this research, we hope that designers from various backgrounds will implement AI gadgets using first-person footage in their daily lives with the toolkit, and through long-term use, gain new insights regarding AI that utilizes first-person footage.

2 Copter: Speculating Human-Machine Symbiosis through Everyday Life with a Autonomous Gadget on the Head

In this section, we introduce the project we worked on called "Copter." The Copter is a head-mounted gadget that autonomously operates while being worn on a person's head and moving with them. It consists of a camera for capturing first-person footage, a motor used as output from AI, and a computer that converts the video and audio input from the camera into motor movements. In an experiment, three individuals, including the authors of this paper, wore the Copter while going about their daily lives for two weeks, documenting their experiences in the form of a diary. After the two-week practice, the authors held a workshop to share insights and interesting episodes gained from living with the Copter, based on the diaries each participant recorded. Subsequently, they devised a new design for "an object that autonomously perceives the world and acts alongside us, yet has no functions for humans," which was expressed through illustrations. This chapter shares the insights gained from the practice of the Copter. It is a summary of the content presented at a research meeting held in Japan.

2.1 Implementation

A Raspberry Pi 5 (8GB) is used as the computer, and the Raspberry Pi Build HAT is connected as an add-on board. The motor employs the LEGO® Education Large Angular Motor, and a commercially available webcam is used to connect to the computer. The movable parts of the copter connected to the motor are designed to allow the user to customize the design. Power is supplied from a mobile battery, and all components are secured to a helmet, allowing it to be worn on a person's head. The video and audio data captured by the camera are sent to the multimodal generative AI model "Gemini," developed by Google, through a Python program running on the Raspberry Pi, which generates the motor's operation patterns. The mounted motor operates according to the generated patterns.

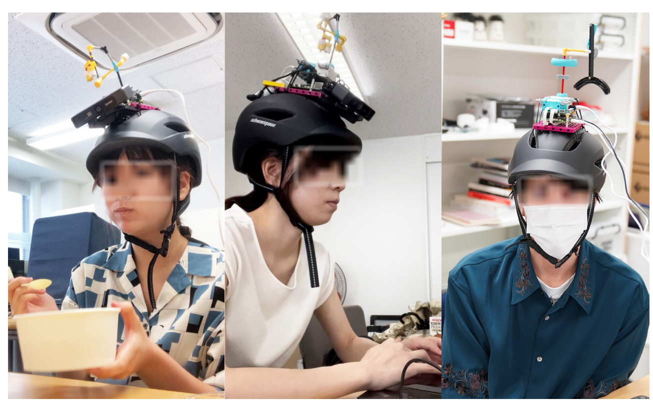

Figure 2: The Experience of Daily Life with Copters

2.2 Evaluation

To capture the interaction between the designed gadget "Copter" and humans, three authors (the first to the third authors, referred to as A1, A2, and A3) regularly wore the Copter for two weeks while going about their daily lives. The participants documented their experiences on the days they wore the Copter in a diary format using Slack's chat function, accumulating these records. After completing the two weeks of living with the Copter, the participants gathered to conduct a workshop on the coexistence of objects and humans based on the recorded diaries. Through this process, they explored how the autonomous device Copter affects human daily life and environment, and what new interactions arise from this coexistence.

2.3 Analysis

Using the transcription data from the reflection workshop (total time: approximately 270 minutes), inductive coding was conducted. Specifically, after performing open coding on a portion of the transcribed data, focused coding was carried out to explore the relationship between humans and autonomous agents and to discuss insights regarding their future state. Subsequently, the emerging codes were summarized and abstracted. The analysis of the data was performed on QDA software. Additionally, interviews and quotes were translated into English by the authors.

2.4 Result

As a result of analyzing the data from life with the copter and the workshop, the following five main themes emerged: 1) Consideration for Privacy, 2) Unexpected Intimacy, 3) projection for the moving of things, and 4) Sense of Presence toward things. Below, We will briefly describe each of these.

Consideration for Privacy. A2 reflected on a situation where he tried to wear the copter but didn't, saying, "The living room is relatively clean, but my own room is a bit messy, so I didn't want to wear the copter. I thought I wouldn't want to wear it because it has a camera. Especially since I'm not showing that footage to anyone." A1 added, "I don't want to show things like my line messages. So, there's a feeling of what is okay to show and what I feel uncomfortable showing, like a sense of unease has started to grow." It can be inferred that the awareness of being seen by the copter has led to a feeling of not wanting to reveal private aspects. A1 focuses on the awareness of being "watched" by the copter and incorporates a feature that allows users to physically block the camera when they do not want to be seen.

Unexpected Intimacy. I had come to see the copter as an intimate presence close to me. A3 felt at that moment that "it's like a pet, and I think it recognizes the copter as a presence nearby," and "when watching the sunset together, it felt like I had gained another friend." A1, while eating breakfast, was talking to itself, saying things like "this is kind of delicious."

Projection for the movement of things. There was a tendency to seek meaning in actions generated by AI. For example, A1 reflects on a situation where they wore a copter and did their makeup, saying, "I was talking while doing it like this, and when the copter moves really smoothly, it feels nice, and I feel like I'm being praised for it. If it doesn't move much, it feels a bit off, and it's really selfish of me, but that's how I felt."

Sense of Presence toward things. The presence of AI has also been reported to change depending on the situation and actions. For example, A2 reflected on the situation where A2 wore a copter and watched a drama on TV, saying, "Sometimes, when the screen goes dark in the drama, the copter reflects and I really reaffirmed myself wearing the copter." It has also been reported that by using the copter long-term, it becomes less extraordinary and discoveries decrease. For example, in the A1 household, the copter was quickly accepted, and "no one says anything even if I wear it at home, and for my family, it's probably already familiar." "When I'm with my family, I feel like it's okay not to wear it. (Omitted) It's already blended into our lives, the copter."

3 MuseBuddy: Proposal for a wearable device that encourages underwater exploration for divers

After practical experience with the copter, I would like to introduce "MuseBuddy," a product designed by one of the participants. MuseBuddy is a device that divers can take with them during underwater exploration. It uses AI to react in real-time to changes in the underwater environment based on footage captured by a camera, providing vibration feedback to the diver to encourage new insights. This concept emerged from the discovery during the copter practice that humans tend to engage in "projection for the moving of things," highlighting the potential for AI to contribute to the Art of Noticing. The concept of AI generating motion based on camera footage and conveying it to humans is the same as with the copter, but the designer of this product, who is also an author of this paper, applied this concept to enhance awareness of the underwater environment during diving, drawing from their extensive diving experience. This chapter summarizes the content presented at a research meeting held in Japan.

3.1 Implementation

"MuseBuddy" is configured to store a Raspberry Pi 5 (8GB), a Raspberry Pi Camera Module, a motor driver, a vibration motor, a mobile battery, and a 1kg weight for fishing gear inside a waterproof casing. The waterproof casing is designed to withstand a maximum water depth of 140m and can be directly attached to the harness of a buoyancy control device (BCD) worn during diving using a carabiner and Velcro. The mounted camera captures video at 15-second intervals, and using an object detection model running on the Raspberry Pi 5, it recognizes fish and divers. If a target is detected, it vibrates for 20 seconds with a 25footnote

Based on the YOLO11n model, it was trained on a custom dataset. The dataset consists of approximately 20,000 images, which are a combination of parts from the Deepfish, Ozfish, and V-DDC datasets.

3.2 Utilization and Discussion

Three divers, including the author, experienced diving while wearing MuseBuddy. After the experience, they mentioned, "I wanted to get a little closer to any fish and see what kind of reactions they would have," and "I think I was able to take an interest in all the fish and enjoy diving in the ocean" (P1). Another said, "By wearing that device, I felt like I became more sensitive along with the machine,

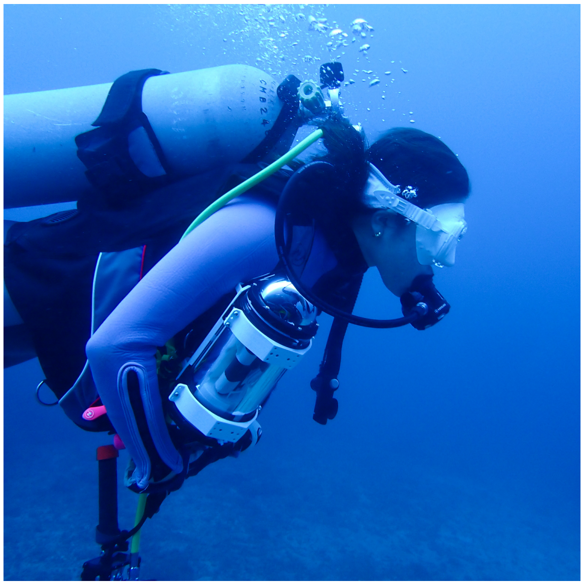

Figure 3: A scene of diving with MuseBuddy attached

probably becoming more conscious and looking around more, being aware of changes in the environment and whether there were fish" (P2). One diver noted, "I think my diving style has changed a bit" (P2).

3.3 Discussion

The intention of the design, along with the presentation of vibrations by AI using a first-person camera, suggested that participants' awareness of their surroundings during underwater exploration was enhanced. On the other hand, unlike with drones, there were no comments from participants regarding privacy concerns. In the context of everyday life targeted by drones, first-person footage is more likely to infringe on privacy, whereas in the extraordinary setting of diving, the benefits of using first-person footage were emphasized instead.

4 Proposal: A toolkit for converting first-person footage into motion using AI

This paper proposes a toolkit that enables the easy creation of devices that convert first-person video into motion using AI. As reported so far, in the current context where AI utilizing first-person video is becoming widely adopted, this practice incorporates the interaction of AI generating motion from first-person video to convey information to humans within various human activities. This allows designers to recognize the challenges of design, including privacy and ethics, and to create new designs based on that awareness. The previously mentioned MuseBuddy is an example of this, implemented by one of the participants in the copter practice, who drew inspiration from their experience as a diver and the personality associated with it to conceptualize and implement the use of first-person video and AI. Similarly, if it becomes easier to create devices that convert first-person video into motion using AI, it may allow individuals to design new devices based on various personalities, leading to the discovery of new Privacy, Ethics, and Design Challenges of First-Person AI Devices, and encouraging the design of new devices based on those findings.

The idea for the toolkit originated from one of the authors sketching a concept for a tail-shaped device equipped with AI that would move based on images after completing the copter practice 4. This designer focused on the ambiguous movements of the copter, which made it difficult to understand what it was thinking, and wanted to incorporate the easily imaginable movements of a cat's tail, which, while ambiguous, suggest what it might be thinking. Additionally, after observing the copter's movements reflected in a mirror, the designer thought, "It would be more interesting if the movements were visible," leading to the idea of a shoulder-mounted design. If AI interactions using first-person camera images can be easily implemented by anyone and used long-term in real life, the insights gained from this practice could potentially expand significantly.

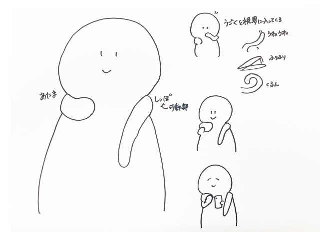

Figure 4: A tail-shaped device designed to wrap around the neck.

Designers will be able to easily incorporate AI-powered gadgets into daily life, just like the practices of Copter and MuseBuddy, by integrating this toolkit into various items used in everyday life. The toolkit is being considered for design around the microcontroller module "M5Stack," which is equipped with the ESP32 CPU that has wireless communication capabilities via Wi-Fi and Bluetooth, along with a camera module for capturing first-person video, motor drivers, and actuators. The software is planned to be made openly available through GitHub. By combining the shape of the gadget with code generation through generative AI in a GUI, we are making it possible to easily implement gadgets like the ones we created this time, even without skills or knowledge in programming.

5 Conclusion

We believe that various AI-powered gadgets implemented using the toolkit proposed in this research, utilized by designers from diverse backgrounds over an extended period, could contribute to building a better relationship between AI and humans, as new privacy, ethics, and design challenges of first-person AI devices emerge. At the workshop, we will demonstrate Computer and MuseBuddy, and share the trajectory of our practice in more detail while providing participants with the toolkit. Participants are expected to implement AI gadgets using first-person video with the toolkit and gain new insights regarding AI in first-person video through long-term use. We look forward to our practice and toolkit contributing to research through design and becoming a topic of discussion.