Object-Agent Twins (OATs): Expressive Micro-Agent Avatars for Inanimate Everyday Objects

Daehwa Kim
Human-Computer Interaction Institute
Carnegie Mellon University
Pittsburgh, PA, USA
daehwak@cs.cmu.edu

Chris Harrison
Human-Computer Interaction Institute
Carnegie Mellon University
Pittsburgh, PA, USA
chris.harrison@cs.cmu.edu

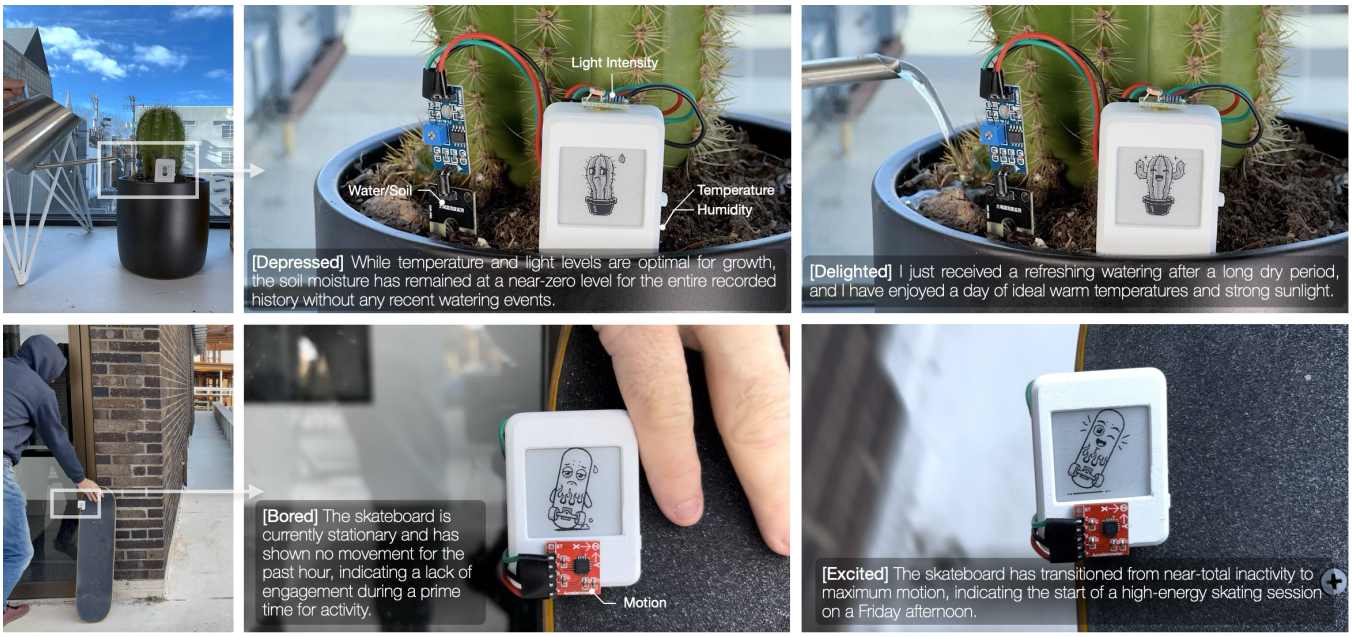


Figure 1: We created a toolkit that allows users to easily create Tamagotchi-like, expressive proxies for inanimate everyday objects, which we call Object-Agent Twins (OATs). Top: an OAT for a cactus, initially *depressed* due to a lack of water, becomes *delighted* when watered. Bottom: A *bored* skateboard becomes *excited* when the user begins to ride it. See also Video Figure.

Abstract

We introduce Object-Agent Twins (OATs), a toolkit for creating Tamagotchi-like companions for inanimate everyday objects. OATs allow non-expert users to easily author small personality-bearing agents tied to real-world artifacts such as a tea kettle, pet bowls, and a trash can. These OATs act as expressive proxies for physical objects, using real-time sensed data to drive simple needs, moods, and behaviors. In contrast to conventional smart-object interfaces that expose raw status and data, OATs render object state through character, making everyday infrastructure more socially legible and emotionally engaging. Through deployments, we show how this approach enables users to transform ordinary things in their lives from passive infrastructure into lightweight companions that invite attention, interpretation, and care.

This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License.
UIST '26, Detroit, MI, USA
© 2026 Copyright held by the owner/author(s).
ACM ISBN 979-8-4007-2856-3/2026/11
https://doi.org/10.1145/3830398.3830595

CCS Concepts

• **Human-centered computing** → *Human computer interaction (HCI)*; **User interface toolkits**.

Keywords

IoT, sensors, toolkit, agentic devices, AI, agents, smart environments

ACM Reference Format:
Daehwa Kim and Chris Harrison. 2026. Object-Agent Twins (OATs): Expressive Micro-Agent Avatars for Inanimate Everyday Objects. In *The 39th Annual ACM Symposium on User Interface Software and Technology (UIST '26), November 02–05, 2026, Detroit, MI, USA.* ACM, New York, NY, USA, 14 pages. https://doi.org/10.1145/3830398.3830595

1 Introduction

Anthropomorphism of inanimate everyday objects is a common human strategy for making relationships with the material world more legible, easier to remember, and emotionally resonant [6, 35]. Across psychology and HCI, prior work has shown that even simple artifacts can become meaningful companions when they present cues that invite projection, interpretation, and care [26, 30, 36]. This tendency is particularly powerful in pet-like and toy-like systems.

Virtual pets such as Tamagotchi [1] have shown that people readily form attachments to minimal agents that express only sparse needs and personalities [1, 18]. Their success with consumers suggests that companionship does not require realism; rather, it can emerge from lightweight, socially intelligible cues that invite repeated acts of attention and caregiving [1, 36]. This raises an opportunity for HCI: if people are already predisposed to treat everyday objects as though they had dispositions, moods, or needs, then personification may serve as a useful interface strategy for ordinary things in the home, workplace, and greater environment. Perhaps for this reason, we are increasingly seeing tech-savvy individuals create bespoke, one-off, pet-like avatars for different aspects of their lives [12, 25].

We present our work on a toolkit that enables non-expert users to easily, expressively, and rapidly create small Tamagotchi-like companions for inanimate objects in their everyday lives. However, unlike Tamagotchis that were *virtual* characters in a *virtual* environment, ours are Object-Agent Twins (OATs); each companion acts as a miniature personality-bearing proxy for a *real-world* artifact, expressing simple needs, moods, and behaviors tied to the object's *real-time* sensed state. By making object conditions visible through character rather than abstract sensor readings, the system turns ordinary infrastructure into something more socially legible and emotionally engaging, turning e.g., rote maintenance into care and sensor data into expression.

2 Related Work

Object-Agent Twins intersect with several areas of the HCI literature, including ambient and peripheral displays, systems that augment everyday artifacts with interactivity, toolkit research in HCI, and object agents and avatars. We also include a discussion of anthropomorphism of objects more generally.

2.1 Ambient & Peripheral Displays

Ambient and peripheral displays are a relevant adjacent area of literature. Work in this area has explored how information can be moved from the center of attention to the periphery, allowing users to remain aware of changing conditions without requiring focused interaction [41, 42]. Rather than presenting explicit dashboards or alerts, ambient displays often communicate through abstract, lightweight, or aesthetically integrated forms that blend into everyday environments while remaining glanceable and interpretable [23, 28, 42]. Similar to ambient displays, OATs aim to make background state visible in a lightweight and legible way. However, whereas ambient displays typically emphasize calm awareness, abstraction, and peripheral perception, OATs instead foreground personification, character, and relational interpretation. However, OATs and this prior work share an important ambition: seeking alternatives to conventional screens, numerical precision, and direct control interfaces for representing ongoing state in everyday life.

2.2 Augmenting Everyday Artifacts

Prior work has explored retrofitting existing artifacts with sensing, actuation, and interactivity rather than replacing them with purpose-built smart devices. Robiot [21] transformed everyday, passive objects into actuatable devices with a single degree of freedom (1-DoF) by automatically generating 3D-printable mechanisms based on video demonstrations of user interactions. IoTIZER [2] retrofitted existing electrical appliances by providing actuation for their built-in controls, such as buttons, knobs, and switches. Taking a similar approach, but functioning as a tactile display rather than an actuator, Facade [11] augmented appliance control panels with 3D-printed tactile overlays for blind and low-vision users. There is also body of work augmenting everyday objects beyond purely functional augmentation. "When Plants Play" [20] coupled a living plant's bioelectrical and environmental signals to a digital game in which the plant acts as an autonomous caregiver to a virtual pet. Similarly, Botanicus Interacticus [27] transformed plants into expressive, interactive interfaces that respond to user touch with sound. Hu et al. [15] integrated plants with 3D-printed structures so that as the plant grew or decayed, the physical structure could morph. Although compelling, such systems are often difficult to implement affordably, robustly, or generalizably across diverse objects. OATs shares this tradition of augmenting unmodified objects in situ, but differs by introducing an expressive, personality-bearing layer that translates sensed states into character behaviors. Furthermore, OATs provides a modular, AI-powered toolkit that enables easy, accessible deployment across a wide array of host objects.

2.3 HCI Toolkits

Toolkit research in HCI has long focused on creating reusable platforms that let researchers, designers, and end users build new classes of interactive systems more quickly and with less technical overhead. Early work framed toolkits as mechanisms for encapsulating recurring technical complexity so that authors could focus on interaction design rather than low-level implementation details [10, 24]. Across software, hardware, and sensing systems, toolkits are valuable not only because they save implementation effort, but because they provide leverage: they make previously difficult forms of interaction easier to instantiate, iterate on, and study [19, 24]. In this sense, toolkits are generative infrastructures for interaction design, enabling users to focus on higher-level experiential questions instead of low-level engineering details [24].

A recurring theme in HCI toolkit work is that the best toolkits open a design space rather than delivering a single fixed artifact. For instance, Phidgets lowered the barrier to building physical interfaces through plug-and-play components [10], while Exemplar supported the creation of sensor-based interactive behaviors through programming by demonstration [13]. More broadly, reflective design tools and creativity support tools emphasize that such systems should encourage iteration, exploration, and appropriation by a wide range of creators [14, 34].

Our system embodies a canonical HCI toolkit contribution. Ledo et al. [19] provide an excellent synthesis of prior toolkit research, presenting the following toolkit definition: "generative platforms designed to create new artifacts, while simplifying the authoring process and enabling creative exploration." OATs meets this framing in four ways. First, it is a generative platform rather than a single fixed application: users can create many different object companions rather than instantiate one predefined experience. Second, it supports the creation of new interactive artifacts, namely Object-Agent Twins that bind expressive pet-like characters to real-world objects and live sensed data. Third, it reduces authoring time and

complexity by allowing non-expert users to rapidly construct these artifacts without needing to build the full sensing, state-mapping, and character-expression pipeline from scratch. Finally, it supports creative exploration by letting users experiment with different objects, personalities, behaviors, and mappings, thereby exploring a broad design space of socially legible object companions. In this sense, OATs is not only a system for animating everyday things, but a toolkit for authoring a new class of relational interfaces for the material world.

2.4 Object Avatars & AI Micro-Agents

Rozendaal et al.'s "Objects with Intent" [31, 32] explores how intelligent everyday things can be designed as collaborative partners, combining the familiarity of mundane objects with the expressiveness and autonomy of agents. The objects themselves are infused with intelligence and personality, but there is no avatar generation as a proxy, nor creation of a toolkit. More recent and related to our work is "Areca" [3, 4], which imagines everyday products that express thoughts and emotions, though also not through character embodiment. More similar to our work, Wang et al. [38] and Iwai and Matulic [16] both explore AI agents layered onto household objects in Augmented Reality. Talking Spell [37] pushes this direction further with a worn system that enables real-time anthropomorphic voice interaction with everyday objects. In terms of systems that create virtual-pet-like artifacts, Digital Knick-Knacks [17] is perhaps the closest, though for conventional desktop GUIs instead of physical objects. Like OATs, Digital Knick-Knacks occupy a space between utility, ornament, and companionship.

This body of work is sparse and disparate, precluding easy comparison. However, overall, our work is less centered on conversational AI, focuses more on augmenting the material world with a physical IoT artifact, invests more heavily into its toolkit nature, and with virtual avatars grounded in the real sensed state of the world. Thus, our work occupies a unique point in the design space, both in terms of technical design and user-created output.

2.5 Anthropomorphism of Everyday Objects

The literature on anthropomorphism explores why people so readily ascribe humanlike qualities to nonhuman entities. Epley, Waytz, and Cacioppo define anthropomorphism as attributing "humanlike characteristics, motivations, intentions, or emotions" to nonhuman agents, and argue that this tendency is systematic rather than incidental [6]. Their three-factor theory proposes that anthropomorphism becomes more likely when people can draw on human-centered knowledge to interpret an entity, when they are motivated to explain and predict its behavior, and when they seek social connection [6]. This framing is especially useful for HCI because it suggests that personification is not merely decorative or whimsical, but can function as a meaningful interface strategy: by casting an object in socially legible terms, a system may make its state easier to notice, understand, and remember [6]. Earlier HCI-adjacent work by Reeves and Nass similarly argued that people often respond to media and computers using social rules ordinarily applied to other people, while Nass and Moon later summarized evidence that users apply politeness, reciprocity, stereotypes, and other overlearned social behaviors even to relatively simple computer systems [26, 30].

Subsequent work sharpens this framing by showing that anthropomorphism is also tied to social motivation. Epley, Waytz, Akalis, and Cacioppo argue that anthropomorphism can increase when people seek social connection, suggesting that people do not only anthropomorphize for explanatory convenience, but also because humanlike interpretations can satisfy affiliative needs [5]. In HCI, this is particularly relevant: if everyday objects are represented through tiny personalities or pet-like proxies, those representations may support not only interpretation, but also lightweight forms of attachment, companionship, and care. In other words, a "thirsty" plant or "sleepy" door is not just easier to parse than raw data; it also invites a relational stance toward an otherwise inert thing [5]. Waytz, Cacioppo, and Epley extend this line of work by showing that anthropomorphism also varies across individuals, and that those differences predict consequential judgments such as moral concern, trust, responsibility, and social influence attributed to nonhuman agents [40]. This suggests that personified objects may not simply be perceived as "cute", but may alter how users orient attention, obligation, and care toward everyday artifacts [40].

Turkle's work offers a complementary perspective by shifting the emphasis from attribution alone to the broader relationships people build with objects and artifacts. In Evocative Objects, Turkle frames familiar things as objects people think with: artifacts that anchor memory, structure routine, and become emotional and intellectual companions in everyday life [35]. This perspective could move beyond the question of whether people mistakenly treat objects as alive, and instead asks how material things participate in thought, attachment, and self-understanding. In parallel, Turkle's work on relational artifacts argues that even relatively simple computational systems can invite nurturance, reflection, and emotional response when they present themselves as entities with which one can form a relation rather than merely execute a function [36]. Her studies with children and elders suggest that the power of such systems does not necessarily depend on intelligence or realism, but on their ability to present a socially meaningful surface that invites response [36]. These two Turkle's works suggest that the appeal of tiny object personalities is not a superficial gimmick; it is continuous with a broader human practice of organizing meaning, memory, and care through relationships with things [35, 36].

To sum up, this literature motivates a design space in which ordinary infrastructure can be encountered not only as functional equipment, but also as small social actors. Prior anthropomorphism research explains why users are predisposed to interpret nonhuman entities in humanlike terms [5, 6, 40], how even simple interactive systems can elicit social responses [26, 30], and why objects and artifacts can become emotionally meaningful companions rather than mere tools [35, 36]. Our work builds on this foundation by exploring how people might create pet-like proxies for mundane inanimate objects and how these personalities can make everyday object states more legible, companionable, and worthy of attention.

3 Implementation

OATs require several different pieces of software and hardware working together, including user-facing web apps, cloud backends,

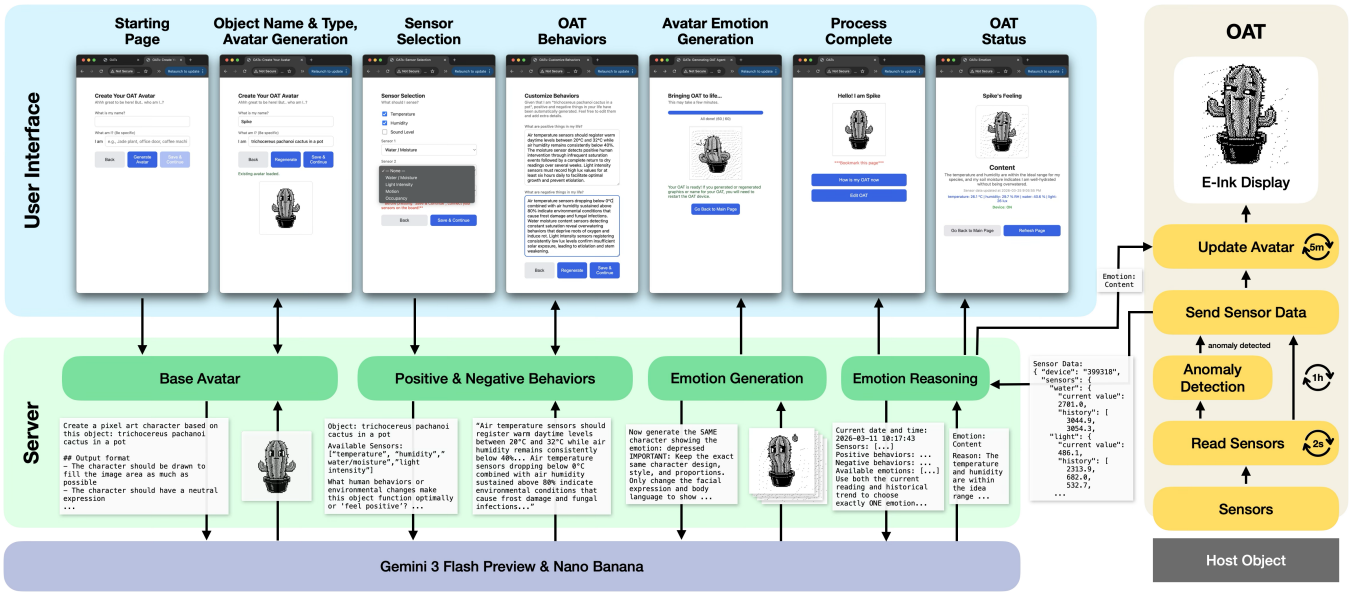


Figure 2: Block diagram of major processes and information flows. A photo of this example OAT can be seen in Figure 1.

generative AI services, and physical OAT devices outfitted with accessory sensors. We now describe these components in detail.

3.1 Open Source

We open source our software and hardware design files so that others can generate their own OATs. We also expose our OAT authoring web app for those who wish to play with the creation process. All material can be found at *https://daehwakim.com/oat/*.

3.2 OAT Creation Front End

We created a web-based interface that allows users to easily create and edit OATs. An overview of this front end is shown in blue in Figure 2. When users first turn on the OAT device, it creates a WiFi access point, and the e-ink display shows instructions for completing the internet setup. Users can connect a laptop or mobile device to this access point, which appears as "OAT-[MAC address]", and set up their device by entering WiFi credentials for a network with internet access.

After this initial step, users can open their account page on our server to create a new OAT. They first name the host object and specify its characteristics, such as the species of a plant or the location/appearance of the object. Once this information is provided, users can click the "Generate Avatar" button to create an avatar (Figure 2, Avatar Generation). They can regenerate this avatar until they are satisfied with it. Next, users select the sensors they consider relevant to the host object's life, choosing among three internal sensors via checkboxes and up to three optional external sensors via dropdown menus (Figure 2, Sensor Selection). On the next screen, users can customize the host object's positive and negative behaviors, starting from an AI-generated behavioral description (described later). Finally, the system generates the full set of emotions and corresponding avatar animations (Figure 2, Avatar Emotion Generation).

After completing this process, users can return to the main page anytime during deployment to either edit their OAT or check its status. While users can see the OAT's current emotion directly on the device's e-ink display, they can also visit the website to see the reason behind the OAT's emotional state (Figure 2, OAT Status).

3.3 Web App & Cloud

Our web-based interface allows users to easily create, modify, and access their OATs. The backend server is hosted in the cloud and implemented in Python using the FastAPI framework. An overview of the server-side of our pipeline is shown in green in Figure 2.

The server is responsible for data storage and management. It queries Gemini models for avatar generation, behavior generation, and emotion reasoning, displays current OAT status and reason, and manages synchronization between the web interface and the OAT client hardware. The OAT client device communicates with the server over HTTP. Whenever a user updates the OAT information, such as new sensors, behaviors, or avatar appearance, the server marks such updates so that OAT can download the latest changes. The client device also sends sensor data to the server, where Gemini performs emotion reasoning and returns the resulting emotion. Across all these AI inferences, we use closed-source large multi-modal models (Gemini) through API calls with our custom prompts. We describe more details in the following sections and the Appendix.

3.4 Avatar Generation & Emotion Generation

We generate OAT avatars using Gemini Nano Banana [8]. Given the user's object description, we first create a simple pixel-art-style character. Our avatar generation pipeline is structured as follows: base (neutral) character generation, followed by the generation of emotional variants, and then the generation of 5 animation frames

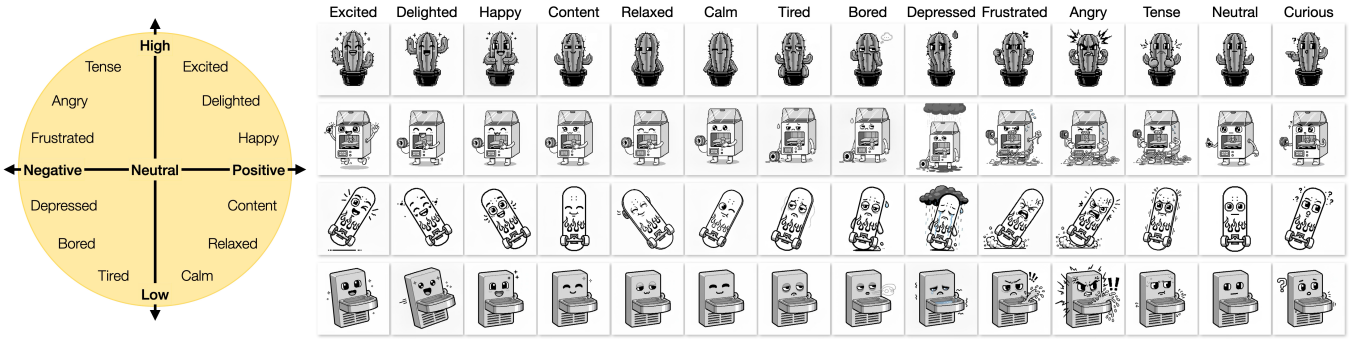


Figure 3: Left: Russell's circumplex model [22, 33]. Right: Example generated emotion avatars for OATs seen in Figures 1 and 6.

for each emotion. This sequential process helps preserve consistency in visual features, while still enabling expressive emotional variation. Detailed prompts are provided in the Appendix A.1.

To cover a core range of emotional expressions, we follow Russell's circumplex model [22, 33], which organizes 12 emotions along two dimensions, valence (positive to negative) and arousal (low to high). This model provides a well-defined and -distributed emotional space that can be generated from the positive and negative OAT behavior descriptions. We additionally generate a 'curious' expression to represent the brief period when emotion reasoning is in progress, providing immediate feedback to the user before the final emotion is displayed. In total, we generate 62 images (12 emotions × 5 animation frames + 1 base + 1 curious). Example emotion generation results are shown in Figure 3.

We convert the generated images into 1-bit packed bitmaps before sending them to the OAT device, so that all images can fit within the flash memory. To preserve visual details such as shadows, stripes, and other patterns despite the binary representation, we apply Floyd–Steinberg dithering [7]. Each image size is approximately 5 KB (5,008 bytes; 200 × 200 pixels / 8 bits + 8-byte header). Altogether, the 62 avatar images occupy about 310 KB of storage, which comfortably fits within the device's 8 MB flash memory.

3.5 Positive/Negative Behavior Generation

On the "Customize Behavior" page (Figure 2, OAT Behaviors), users specify the positive and negative behaviors of the host object, and this directly shapes the emotion reasoning process. To facilitate the authoring process, we generate initial positive and negative behaviors using using Gemini 3 Flash Preview [9]. The model is given the user's object description and selected sensor set. Detailed prompts are provided in the Appendix A.2.

We found these generated defaults helped users consider relevant contextual factors they may have overlooked (e.g., a cactus benefits from sufficient sunlight to prevent etiolation) while also inviting users to think about behaviors that can be feasibly inferred from the selected sensors. Moreover, in many cases, the AI generated behavior text was more accurate and sophisticated than what a user would typically elaborate. For instance, for plant care, Gemini was able to leverage the type of plant (e.g., ponytail palm, cactus, etc.) to describe idealized watering intervals, soil moisture, light exposure, etc. that finely tuned the OAT behaviors to facilitate proper care.

3.6 OAT Device

Each OAT device is build around a low-cost ESP32-S3 platform created by Waveshare [39]. This IoT-style device supports WiFi 802.11 b/g/n and Bluetooth 5. Our custom firmware (0.96 MB) is loaded into its 8 MB Flash (we use the remaining space to download avatar images once the device connects to the internet). Our OAT avatars are displayed on the device's 40 mm diagonal, two-color, e-ink display. We selected a 1100 mAh, 3.7V LiPo battery, which provides run time for about one week (this varies by how many and what sensors are connected; see Table 1). Conveniently, the board includes battery charging and management hardware. We 3D printed a custom enclosure. The total cost per device was under $20; optional additional plug-in sensors (discussed in Section 3.8) varied in cost, but were all under $5. All of these components can be seen in Figure 4.

3.7 Sensor Board

Our sensor board docks with the main board using a 2×6 pin header (Figure 4). This board serves several important functions. First, it hides all hardware that is irrelevant to our end users. Second, we provide plug-and-play headers with which to easily attach optional accessory sensors. As we will discuss later, most of the sensors we provide to users can be turned on momentarily to take a reading, and then be turned off again, saving considerable battery power. To facilitate this, our sensor board includes a transistor that toggles on/off the sensor power rail. However, some sensors (often those with auto-gain features) require persistent power to function well, and so we provide a second row of headers with constant power for such sensors.

3.8 Example Sensor Suite

We created a proof-of-concept suite of plug-and-play sensors for our OATs, seen at the bottom of Figure 4 and Table 1. Our sensor board provides color-coded (red/green/black) 3-pin headers for up to four external sensors. In general, any analog sensor is supported; 3.3V power is provided. Digital sensors could be supported, but we leave this to future work. The ESP32-S3 board includes an on-board SHTC3 temperature and humidity sensor, as well as a MEMS microphone (which we expose for sound level sensing, but not audio). We expose these three sensors as user-selectable options in our OAT authoring interface (seen in Figure 2, Sensor Selection).

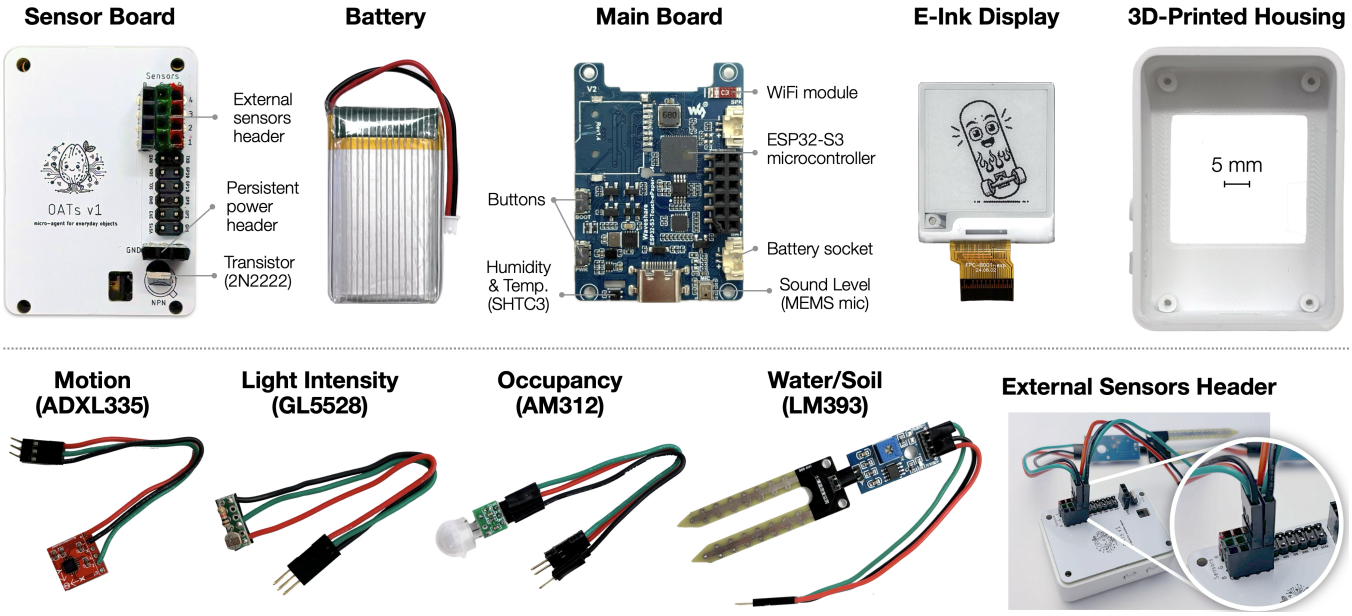


Figure 4: Top: The major hardware components of an OAT. Bottom: OATs are extensible with external sensors. We created a proof-of-concept suite of sensors that plug into a color-coded header on the back of an OAT device.

3.9 Power Consumption & Battery Life

An important design goal was for OATs to be relatively maintenance-free, lasting a week without having to be recharged in general. Of course, our OAT devices are small, containing a modest 3.7 V, 1100 mAh LiPo battery. Thus, we invested some engineering effort to reduce power consumption (though far from a commercial-level implementation). Our main strategy is to keep the ESP32-S3 processor in light sleep as much as possible (described in Section 3.10). When in this state, the device consumes 6.1 mA of power, mostly spent maintaining its volatile memory. When the processor is awake and running, it consumes roughly 68 mA. Updating the e-ink display consumes roughly 24 mA for 570 ms.

We note that establishing a WiFi connection and exchanging data over the internet is an expensive operation in terms of power. Establishing the connection, transmitting sensor data to the server, and then waiting for the emotion reply takes 16 seconds on average, during which there is an average total power draw of around 100 mA. For this reason, OAT devices must perform its own sensor anomaly detection (discussed next) to minimize connections to the server, and only connect to the internet when there is a high likelihood of an emotion change.



Figure 5: Visualization of the 14-day, hourly, sensor history provided to the reasoning pipeline (as JSON). This example is from the Avocado Plant OAT; note the day/night cycles and the watering event roughly four days ago.

3.10 Reading Sensors & Emotion Updates

Every two seconds, the OAT wakes up via a timer interrupt. It powers on the user-selected sensors and collects one or more readings depending on the sensor type (e.g., humidity requires one reading, whereas sound level RMS is computed from 2048 samples).

We update the emotion when a sensor reading is detected as anomalous. To determine this, we use three complementary anomaly-detection strategies. First, we compare the current reading against the *short-term* history of the last 3 sensor readings (i.e., 6 seconds; we refer to this as ST). Second, we compare it against a *longer-term* history of the last 30 sensor readings (i.e., 60 seconds; LT). For these two history-based methods, we compute the deviation from each history, select the larger deviation, and multiply it by three (3σ) to form an adaptive trigger threshold for that sensor. These two methods are applied across all sensors. Third, especially for humidity and temperature, whose dynamic ranges are relatively small, we additionally compare the current reading against the

Component	Part #	Power (mA)	Anomaly Algo.
OAT light sleep	ESP32-S3	6	.
OAT awake & running	ESP32-S3	68	.
OAT running + WiFi	ESP32-S3	100	.
E-ink update	1.54"	24	.
Temperature	SHTC3	10	ST, LT, PA
Humidity	SHTC3	10	ST, LT, PA
Sound level	5492 3801	25	ST, LT
Water/Moisture	LM393	1	ST, LT
Light intensity	GL5528	0.2	ST, LT
Motion	ADXL335	0.3	ST, LT
Occupancy	AM312	0.01	ST, LT

Table 1: Power consumption of main components. Different sensors employed different anomaly detection algorithms: short-term history comparison (ST), long-term history comparison (LT), and previous-anomaly comparison (PA)

most recent value that was previously flagged as anomalous (PA; previous-anomaly comparison), using fixed thresholds of 1.0°C for temperature and 5.0% for humidity.

An OAT device updates its emotion in two cases: either when a one-hour timer elapses, or when any sensor reading is detected as anomalous using the methods described above. If neither condition is met, the OAT simply returns to sleep after reading the sensor value. Otherwise, the OAT establishes a WiFi connection and opens a socket to our server. It sends the latest sensor readings together with hourly averages from the previous 14 days. The server processes this information and returns an emotion label (e.g., happy). Upon receiving the result, the OAT disconnects from WiFi and displays the corresponding pre-saved emotion graphics by stepping through five animation frames twice, updating the e-ink display once per second (i.e., 10-second animation). Afterward, it continues updating the animation at a slower rate, once every 5 minutes.

3.11 Live Reasoning & Deployed Operation

Figure 2 "Emotion Reasoning" illustrates the information flow of live emotion reasoning. The OAT device sends both current and historical sensor values to the server to use them in an emotion-reasoning prompt (Figure 5). We append the current date and time, the user-customized positive and negative behavior descriptions, and the sensor data before sending the request to Gemini.

To make the sensor data more interpretable to the model, we also provide semantic descriptions of each sensor's value range, including units and example meanings. For example, for light intensity, we provide semantic reference points ranging from darkness (≈7 lux) to office lighting (≈127 lux), daylight near a window (≈600 lux), and full sunlight (>2000 lux). We use Gemini 3 Flash Preview for fast emotion inference. We ask Gemini to select the most appropriate emotion (given 12 emotions based on Russell's circumplex model [33]) and briefly explain the reason. Detailed prompts are provided in the Appendix A.3.

The inferred emotion is then returned to the OAT device and displayed on the e-ink screen. Example emotion reasoning output can be seen in Figure 1. Users can also access the website to view the reason for the OAT feeling, as shown in Figure 2, OAT Status.

4 User Study

Following Ledo et al.'s framework for toolkit evaluation [19], our user study examined three main questions: (1) what creative explorations did participants pursue with OATs, (2) how easy was the OAT creation process, and (3) what was the participant experience and what value did they draw from OATs. Importantly, our goal was not to compare OATs against traditional interfaces on efficiency or information retrieval alone, since conventional dashboards, alerts, and status indicators already serve those needs well. Instead, we focused on a set of needs that such interfaces typically leave unmet: making object state socially legible, supporting playful and expressive authoring, and fostering engagement, attachment, and care toward everyday things. Accordingly, our evaluation asks whether OATs succeeds as a toolkit and interaction paradigm for this alternative design space. To explore these questions, we conducted a seven-day study in which participants deployed novel, personalized OATs in their own environments.

4.1 Participants & Procedure

We recruited 12 participants for a seven-day study (mean age=27.4 years, SD=1.7; 8 self-identified as female and 4 as male) and they received $20 in compensation. If participants were, e.g., traveling or did not go to their workplace over the weekend, and could not have interacted with their OAT, we did not consider those to be deployment days.

4.2 Workshop Session

Participants started with a workshop session (max three simultaneous participants), in which they created their OATs and prepared them for deployment in their own spaces. At the beginning of the session, we introduced the concept of OATs and presented three example host objects (sea onion plant, refrigerator, and air conditioner) that we had prepared, along with example behaviors and relevant sensors. We then introduced all seven available sensors that they could select and explained what each sensor measured.

Next, participants were given some time to think about an everyday object in their home, workplace, or other personal environment that they might want to augment with an OAT. Participants were then given an OAT device. Participants powered on the device and followed the on-screen instructions. Participants used their own laptop or mobile device to complete the OAT creation process. Study instructors were present throughout the workshop to answer questions, but mainly assisted participants in providing materials that could help attach OATs to the chosen host object, such as adhesive tape, glue, magnets, etc.

At the end of the workshop, participants completed a ***post-workshop questionnaire*** consisting of five Likert-scale questions (1 strongly disagree to 7 strongly agree): (1) "I was successful in creating my OAT", (2) "It was easy to create my OAT", (3) "It was fun to create my OAT", (4) "The avatar depicts the object I selected", and (5) "I am excited to deploy my OAT", plus one open-ended question: "Is there anything you want to share about your experience?"

4.3 Seven-Day Deployment

Participants deployed their OATs on the same day as the workshop. On the following day, they first submitted photos showing how they had deployed their OATs through our questionnaire form. OATs were deployed to kitchens (3), bedrooms (3), living rooms (2), a basement (1), an office (1), a workshop (1), and a office hallway (1). Participants completed a ***daily questionnaire*** sent at 3 PM for six days. This online survey included six seven-point Likert-scale questions: (1) "My OAT is functioning well", (2) "My OAT makes me think about its host object", (3) "My OAT caused me to change my behavior", (4) "I understand my OAT's behaviors", (5) "I want to refine my OAT's behaviors", (6) "I am happy with my OAT's behaviors". There was one open-ended question: "Do you have any additional comments you would like to share?" Two participants missed two daily questionnaires each, giving us 68 reports instead of a full set of 72 reports (12 participants × 6 days).

4.4 Debrief Session

After seven days of deployment, we conducted an open-ended debrief interview with each participant about their experience living with their OAT. These interviews were conducted either in person

or over video call. The interview began with general questions about participants' overall thoughts about the OAT, how it felt to live with it over the last week, and where else they might want to deploy it. We then asked more specific follow-up questions tailored to each participant's deployment experience. For example, we asked whether the OAT influenced their routines, such as their workout schedule, water intake, or sleep habits, and whether they changed sensors during the deployment, including why and how they made those changes. Finally, participants completed a ***debrief questionnaire*** consisting of four seven-point Likert-scale questions: (1) "My OAT functioned well", (2) "I enjoyed having my OAT", (3) "I want to keep my OAT", and (4) "I want to have more OATs". There was one final open-ended question: "Do you have any additional comments you would like to share?"

5 Study Results & Discussion

The post-workshop, daily deployment, and debrief questionnaire results are summarized in Figure 7. Interview data were processed using thematic analysis. In the following sections, we discuss these results together with illustrative interview quotes and OAT examples created by our participants.

5.1 Participant-created OATs

Figure 6 shows the OATs created by our 12 participants. A wide range of inanimate objects were selected, including domestic appliances (refrigerator, kettle, cat litter box), furniture (bedside lamp, bed), office fixtures (water fountain, coffee maker), equipment (3D printer, dumbbell), infrastructure (door), and two plants.

Across these deployments, participants created OATs in three broad themes: for the host object itself, to support awareness of their own habits, and for others in a shared environment. For host-object-oriented use, participants attached OATs to objects such as snake or avocado plants so that changes in conditions like temperature or watering needs could become more noticeable through the avatar's expressions. For self-oriented use, participants deployed OATs on objects such as a dumbbell to stay aware of workout routines, and a bedside lamp to reflect on sleep habits. For other-oriented or shared-context use, participants placed OATs on objects such as a public water fountain to encourage hydration, a cat litter box to notice whether a cat was using it regularly, and a resin 3D printer to keep humans away from toxic fumes.

5.2 Understanding OATs

Across the seven-day deployment, participants consistently reported that their OAT functioned well (mean=6.1, SD=1.2) and that they understood its behaviors (mean=6.1, SD=1.2). One participant with a refrigerator OAT mentioned, "If I open the fridge for a longer time, it starts to generate a frustrated face. I like that! Also, it helps to realize my fridge door is not cool enough." A participant with a plant OAT said, "It is reactive to my interaction with it, so I want to play with it". Similarly, a participant using OAT with a cat litter box described how the device appeared tense in the morning but became happy later in the day, which they interpreted as their cat had used the litter box. These suggest that OAT could express states in ways that participants found reactive and meaningful.

Participants felt that the avatar depicted their selected object well (mean=6.6, SD=0.7). For example, one participant who created an OAT for their bedroom noted "its cartoon-like style fit well with my environment". At the same time, some participants suggested that the avatar's emotional expressions could be more tightly matched to specific situations like "I think the avatar could look sweaty when it's too hot, instead of looking sad or angry".

Participants also found that emotional expression made the host object's condition easier to understand than more traditional dashboards (e.g., air quality displays). One participant used a bedside lamp OAT almost like an ambient clock, mentioning that "Because my OAT changes its behavior based on time and light intensity, I kinda use it as a clock ... When I go home, I first turn on the light, and if I see that OAT is still happy, I know it's not too late. Then, I kept either studying or playing." Another participant with a refrigerator OAT described emotionally associating the device's expression with the food inside, saying, "When OAT is mad, then I felt the food is mad." One participant, who had expertise in accessibility, suggested OAT's emotional display "would be great with the elderly and autism," because "facial expressions would be easier to interpret than textual or numerical readouts" and could "also support learning about appropriate emotional responses in different situations for people with autism."

The limited expressive channel, combined with the stochastic nature of generative AI, sometimes created serendipitous experiences. One participant found that their OAT with a motion sensor had fallen off and they "picked it up to reinstall it. Then, OAT was seen frustrated. I thought it was well matched to the situation." Another participant who deployed the OAT to an avocado plant noted that "one side of the plant was exposed to sunlight for too long, so it started leaning. And my OAT avatar was tilted toward the sunlight, too. I know it is probably a coincidence, but I was surprised."

On the other hand, the limited expressive range also made participants curious about additional information. Some wanted to ask further questions if there is a chat feature on the OAT Emotion Status website. One participant noted that while the distinction between positive and negative emotions was clear, "I don't really know how to correlate 'content' or 'calm' with the actual physical phenomena [of my plant]." Another said that although they could see their refrigerator was frustrated, "I did not know the temperature was too high until I checked the website". Several participants said they would prefer to speak directly to the physical OAT to get more information. Some participants also expressed interest in accessing historical data, especially in cases where they may have missed a moment of negative change, saying, "If the environment is negative, then I can collect the data for it to take care of it better."

Participants calibrated both their expectations and their understanding of OATs over time by adjusting sensor configurations and refining behavior prompts. Participants' response to "I want to refine my OAT's behaviors" decreased from a mean of 5.2 (SD=2.0; lying between 5.0 "somewhat agree" and 6.0 "agree") on Day 1 to 2.6 (SD=2.0; lying between 2.0 "disagree" and 3.0 "somewhat disagree") on Day 6 (Figure 7). Two out of twelve participants reconfigured the sensors to better match their intended use. For example, the participant with the litter box OAT mentioned "I decided to point it [the occupancy sensor] downwards so it would trigger when my cat uses the litter box rather than detecting my motion." The

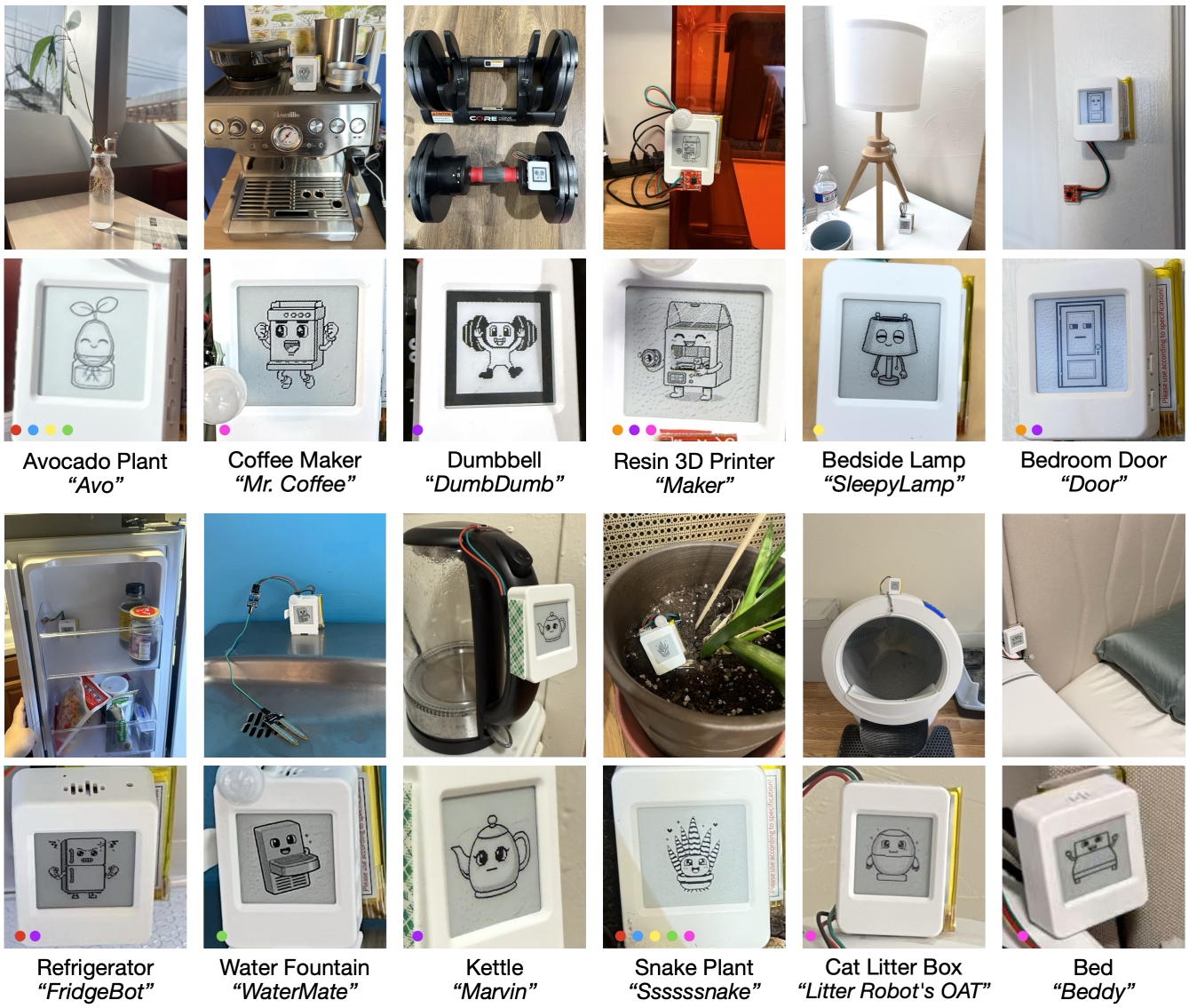


Figure 6: OATs created by our twelve study participants. Sensors used by each OAT: ● Temperature, ● Humidity, ● Sound Level, ● Motion, ● Light Intensity, ● Occupancy, ● Contact Moisture.

participant that deployed the water fountain OAT noted "I learned that my OAT wasn't behaving as expected because it relies on both occupancy data and water data ... people often pass by without using the water" and therefore "removed one of the sensors and modified my OAT's behavior ... after which its behavior is more understandable and reliable since it only uses the water sensor."

Five out of twelve participants revised their OAT's behavior prompts to better align with sensing limitations. The participant with the bed OAT explained, "I wanted it to be happy when I go to sleep on time and get enough sleep ... It was happy a lot of times when it sensed me, but I found the occupancy sensor actually doesn't detect too much when I'm sleeping ... so it often inferred that I was not sleeping enough". The participant later simplified the behavior, "only if the bed was unoccupied for more than 16 hours, it would become sad", later concluding "now everything is good."

5.3 Value to Users

In the post-workshop questionnaire, participants reported that they were excited to deploy their OATs (mean=6.6, SD=1.4) and that the creation process was fun (mean=6.8, SD=0.4). Over the course of the deployment, participants showed significantly increasing agreement that they were happy with their OAT's behaviors (linear mixed-effects model, $p = .006$), rising from Day 1 (mean=5.0, SD=1.8) to Day 6 (mean=6.3, SD=0.5), on average 5.9 (SD=1.2) across days (Figure 7). By the end of the study, responses in the debrief questionnaire were strongly positive: participants reported that they enjoyed having the OAT (mean=6.9, SD=0.3), wanted to keep it (mean=6.9, SD=0.3), and wanted to deploy more OATs in their spaces (mean=6.9, SD=0.3).

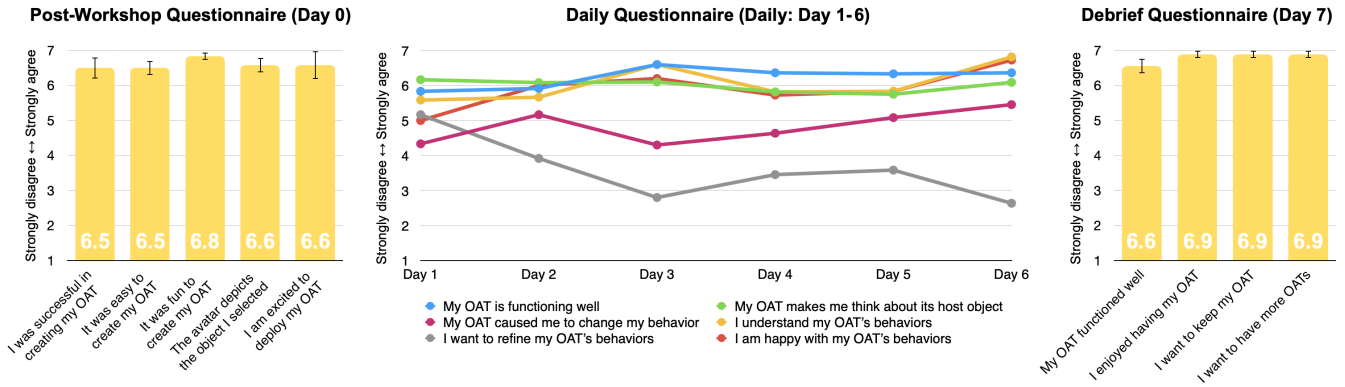


Figure 7: Summary of Likert-scale questionnaires given post-workshop, daily during deployment, and at the final debrief.

There were two main findings from participant debrief interviews. First, OATs transformed tasks that might otherwise have felt tedious into more caring goals. Ten of the twelve participants explicitly described their OAT avatars as "cute" or "adorable" with comments such as "very cute to see it change emotions", "the OAT is so cute. I love him", and "I think the cute avatar is really important! It makes me really happy!" This emotional attachment appeared to reshape how participants experienced everyday tasks. For example, the participant with the 3D printer OAT said, "It makes me look forward to my prints more and enjoy more iterations ... When printing failed, I was sad before, but now I can do more interaction with the OAT, so iteration becomes enjoyable." Another participant similarly reflected, "It's interesting how my objective changed because of OAT. Without it, I work out for other reasons like to keep in shape or boost my mood, and would not do a short workout." The other participant said, "Monitoring the plant is nice, but I can take care of the plant based on my own observation. I just use OAT because it's cute. Actually, after this, I spent more time with my plant." On the other hand, the participants who deployed OATs to a bedroom door and coffee machine reported feeling less attachment. For example, one participant described the relationship as "one direction", explaining, "I get benefit from the coffee machine and I don't provide benefits for the coffee machine".

Second, several participants highlighted the social dimensions of OATs. One participant described how their OAT became a conversation piece when guests visited: "Today my friends visited and saw my OAT, and we discussed its emotional state ... I like when friends come over and see it-'oh, you have a plant that looks happy.'" Others noted, "there is a novelty effect ... My wife liked it." and imagined OATs as a way to communicate with family at a distance, saying "My parents would like this to see how I am doing through OAT." These comments suggest that OATs not only shaped participants' own relationships with objects but also invited social interaction and interpretation from others.

5.4 User Behavior Change

In the daily questionnaire, participants reported that OATs made them think more about the host object (mean=6.0, SD=1.5), and they reported slightly above-neutral agreement that OATs influenced their behavior (mean=4.8, SD=1.7). We observed three broad patterns in these behavior changes.

First, some participants directly changed their behavior to make the OAT happy. For example, one participant "rearranged the fridge to make the OAT content" and "moved vegetables inside and soy sauce outside" after realizing that "my fridge door is not cool enough compared to inside." Others described "making tea more and more at home" and "transfer[ring] the plant to a new bottle and mov[ing] it to a sunnier place to see more positive avatar behavior." Others adjusted their routines in response to the OAT; "I tried to sleep early and turn off my lamp earlier to make it happy." Similarly, one participant shared, "Today I ran in my gym to decompress during a stressful time ... having the OAT there, I did a 10-minute short strength training with my dumbbells because I wanted the OAT to be happy ... I would not have done strength training otherwise."

Second, some participants could not or did not act directly, but still experienced an emotional shift. In these cases, OATs drew attention to the host object and prompted feelings such as "guilt" when it was neglected. For example, one participant said, "Having the OAT makes me feel more guilty when I forget about it [my plant]," while another said, "I have been super busy to work out ... I keep thinking of my OAT and feel guilty about making it sad!"

Finally, we also observed cases in which participants found "hacky" ways just to make their OAT happy without changing the underlying behavior itself. For example, one participant said, "I use the water fountain more, and even when I'm refilling my water bottle, I intentionally let some water reach the sensor to indicate that I used the fountain." Another noted, "With OAT, I kind of wanted to move the dumbbell just to make it happy." Participants also sometimes revised the OAT behavior prompt in more forgiving ways. For instance, one participant changed the Gemini-generated humidity range for their avocado plant from 50%–70% to 10%–70% in order to see a happy expression more often.

5.5 Ease of Creation & Use

During the workshop session, all participants were able to successfully create an OAT, taking 15 minutes on average. This was also reflected in the post-workshop questionnaire: participants reported high agreement that they were successful in creating their OAT (mean=6.5, SD=0.9) and that the creation process was easy (mean=6.5, SD=0.8). Participants commented that the setup was "very fast," "didn't have the burden of programming," and the web tool made it easy to modify OATs and their behaviors. Participants

also found the avatar generation process smooth. For example, one participant chose a relatively niche object, and remarked, "In the description for the avatar generator, I only wrote 'Litter-Robot 4,' and I was surprised that the generator was able to pick up on that."

Some participants encountered trouble determining what sensors to use for their OAT. As one participant explained, "I myself would not have an idea how to do that [what sensor to use]. It would be helpful to have a system or an online community where I could share a photo of my host object and get ideas from others." They also asked, "Can I add sensors outside of the set given?" Our platform can easily support other sensors, and we hope to enable this flexibility in future updates.

All participants successfully installed their OAT on their host objects and updated the device's WiFi settings to their local network through our web interface. However, physically attaching the OAT was sometimes less straightforward, and the provided materials were not always ideal. For example, one participant deployed the OAT to an avocado plant growing in a narrow glass bottle, leaving little space to attach the device. They improvised by using a plastic cup lid to hold both the water sensor and the OAT. Reflecting on this experience, the participant said, "I used the simple and ugly way [to install OAT]. If there is an online community where people share 3D printed parts, it would be great, such as snapping parts for OAT to the bottleneck."

6 Limitations & Future Extensions

While anthropomorphic rendering can make object state more legible, memorable, and engaging, it also introduces potential drawbacks. Expressing sensed conditions through moods, needs, and personalities may oversimplify the underlying data and encourage users to infer richer agency, intelligence, or intention than the system actually possesses. In some cases, this can produce misleading mental models about what an object "knows" or why it behaves a certain way. Pet-like framing may also create unwanted emotional burden: what is initially charming could become nagging or guilt-inducing over time, particularly if companions repeatedly demand attention. More broadly, the appeal of anthropomorphism may not generalize across all users and contexts. We also acknowledge that our one-week deployment was relatively short and almost certainly had novelty effects. Accordingly, our results speak more strongly to early adoption than to sustained attachment or enduring utility. Future work will need to explore longer deployment durations to examine how these relationships develop over time.

A further conceptual question is when an OAT is perceived as a true "twin" of its host object rather than as a standalone virtual character merely associated with it. In our examples, this connection appeared stronger when an OAT's behaviors were tightly grounded in the object's own internal state or direct use (e.g., a coffee maker, refrigerator), and potentially weaker when behaviors were driven more indirectly by surrounding environmental signals, as in the case of a plant. This suggests that twinning may exist on a spectrum, shaped by both the physical proximity of sensing to the object and the perceived causal specificity of the mapping between sensed state and character behavior. Importantly, an OAT need not always function as a literal mirror of the object to be useful or engaging; however, future work could more explicitly examine what design factors make an agent feel object-intrinsic rather than object-adjacent, and how this distinction influences user interpretation, attachment, and the perceived validity of the "twin" metaphor.

OATs currently communicate mainly through avatar expressions and a brief textual explanation on the website. Several participants wanted to ask follow-up questions or understand the system's reasoning with or without accessing the website. Supporting a conversational interface, either through spoken interaction with the OAT device or a text-based chat feature on the website, could make the system more interpretable, at the cost of making interactions more heavyweight.

The current toolkit supports a fixed set of sensors, which constrains the kinds of objects and phenomena users can represent. Future work should explore how to make the system more extensible, so that users can more easily integrate additional sensor types beyond the current set. Similarly, some participants wanted more systematic or community-based support for choosing host objects, selecting appropriate sensors, and attaching OATs more cleanly and robustly. This could take the form of shared examples, attachment strategies, and object-specific 3D-printable parts. Lastly, the current avatar system maps sensed situations to a predefined set of emotions. Future work could also make avatar expressions more context-specific by incorporating sensor conditions. For example, as one participant suggested, the avatar could appear sweaty under high temperatures rather than expressing a sad emotion.

OATs are currently designed for single object use. Future work could explore collections of OATs that communicate with one another and work collectively. For example, one participant in our study noted that different sections of a refrigerator are intended for different storage purposes and therefore maintain different temperatures; they suggested placing multiple OATs within the fridge and combining their readings into an aggregated emotional state. Communication among multiple OATs could support coordinated behaviors across objects and enable meaningful experiences in homes or shared environments.

Finally, the current version of OATs relies on cloud-based language models for reasoning, but future versions could perform this locally on embedded hardware as models become more efficient. Recent work suggests that even devices such as the Raspberry Pi Zero can run compact language models on-device [29], which could extend OAT to more mobile use cases such as bicycles or carried water bottles.

7 Conclusion

We have presented Object-Agent Twins (OATs), a toolkit that enables non-expert users to create Tamagotchi-like companions for everyday inanimate objects. By combining plug-and-play sensors, AI-supported avatar generation and emotion reasoning, OATs turn ordinary artifacts into socially legible and emotionally engaging companions. Through a seven-day deployment study, we found that participants were able to easily create and deploy OATs across a range of application contexts. We found that OATs encouraged attention and behavior change through character-driven interaction, suggesting that OATs could offer infrastructures that shape everyday life. Lastly, we have open-sourced OATs to facilitate deployment, experimentation, and replication.

References

[1] Anne Allison. 2006. Tamagotchi: The Prosthetics of Presence. In *Millennial Monsters: Japanese Toys and the Global Imagination*. University of California Press, Berkeley, CA, 163–191.

[2] Hyungjun Cho, Han-Jong Kim, JiYeon Lee, Chang-Min Kim, Jinseong Bae, and Tek-Jin Nam. 2021. IoTIZER: A Versatile Mechanical Hijacking Device for Creating Internet of Old Things. In *Proceedings of the 2021 ACM Designing Interactive Systems Conference* (Virtual Event, USA) *(DIS '21)*. Association for Computing Machinery, New York, NY, USA, 90–103. doi:10.1145/3461778.3461996

[3] Hyungjun Cho, Jiyeon Lee, Bonhee Ku, Yunwoo Jeong, Shakhnozakhon Yadgarova, and Tek-Jin Nam. 2023. ARECA: A Design Speculation on Everyday Products Having Minds. In *Proceedings of the 2023 ACM Designing Interactive Systems Conference (DIS '23)*. ACM, New York, NY, USA, 31–44. doi:10.1145/3563657.3596002

[4] Hyungjun Cho and Tek-Jin Nam. 2025. Living Alongside Areca: Exploring Human Experiences with Things Expressing Thoughts and Emotions. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems (CHI '25)*. ACM, New York, NY, USA, Article 434, 434:1–434:16 pages. doi:10.1145/3706598.3713228

[5] Nicholas Epley, Adam Waytz, Scott Akalis, and John T. Cacioppo. 2008. When We Need a Human: Motivational Determinants of Anthropomorphism. *Social Cognition* 26, 2 (2008), 143–155. doi:10.1521/soco.2008.26.2.143

[6] Nicholas Epley, Adam Waytz, and John T. Cacioppo. 2007. On Seeing Human: A Three-Factor Theory of Anthropomorphism. *Psychological Review* 114, 4 (2007), 864–886. doi:10.1037/0033-295X.114.4.864

[7] Robert W. Floyd and Louis Steinberg. 1976. An adaptive algorithm for spatial grey scale. *Proceedings of the Society for Information Display* 17, 2 (1976), 75–77.

[8] Google DeepMind. 2025. Gemini 3 Flash Image (Nano Banana 2). https://aistudio.google.com/models/gemini-3-1-flash-image. Accessed: 2026-03-20.

[9] Google DeepMind. 2026. Gemini 3 Flash Preview. https://ai.google.dev/gemini-api/docs/models/gemini-3-flash-preview. Accessed: 2026-03-20.

[10] Saul Greenberg and Chad Fitchett. 2001. Phidgets: Easy Development of Physical Interfaces through Physical Widgets. In *Proceedings of the 14th Annual ACM Symposium on User Interface Software and Technology (UIST '01)*. ACM, New York, NY, USA, 209–218. doi:10.1145/502348.502388

[11] Anhong Guo, Jeeeun Kim, Xiang 'Anthony' Chen, Tom Yeh, Scott E. Hudson, Jennifer Mankoff, and Jeffrey P. Bigham. 2017. Facade: Auto-generating Tactile Interfaces to Appliances. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems* (Denver, Colorado, USA) *(CHI '17)*. Association for Computing Machinery, New York, NY, USA, 5826–5838. doi:10.1145/3025453.3025845

[12] Maggie Harrison Dupré. 2025. We Talked to the Inventors of the "Tamagotchi" Vape That Dies If You Stop Puffing. Futurism. https://futurism.com/neoscope/vape-tamagotchi-interview Accessed: 2026-07-28.

[13] Björn Hartmann, Leith Abdulla, Manish Mittal, and Scott R. Klemmer. 2007. Authoring Sensor-Based Interactions by Demonstration with Direct Manipulation and Pattern Recognition. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '07)*. ACM, New York, NY, USA, 145–154. doi:10.1145/1240624.1240646

[14] Björn Hartmann, Scott Doorley, and Scott R. Klemmer. 2008. Hacking, Mashing, Gluing: Understanding Opportunistic Design. In *IEEE Pervasive Computing*, Vol. 7. 46–54. doi:10.1109/MPRV.2008.54

[15] Yuhan Hu, Jasmine Lu, Nathan Scinto-Madonich, Miguel Alfonso Pineros, Pedro Lopes, and Guy Hoffman. 2024. Designing Plant-Driven Actuators for Robots to Grow, Age, and Decay. In *Proceedings of the 2024 ACM Designing Interactive Systems Conference* (Copenhagen, Denmark) *(DIS '24)*. Association for Computing Machinery, New York, NY, USA, 2481–2496. doi:10.1145/3643834.3661519

[16] Nozomu Iwai and Fabrice Matulic. 2025. Bringing Everyday Objects to Life in Augmented Reality with AI-Powered Talking Characters. In *Proceedings of the Extended Abstracts of the CHI Conference on Human Factors in Computing Systems (CHI EA '25)*. Association for Computing Machinery, New York, NY, USA, Article 117, 7 pages. doi:10.1145/3706599.3719978

[17] Matthew Lakier and Daniel Vogel. 2024. Digital Knick-Knacks: Standalone Audiovisual Digital Possessions or Embellishments in Digital Environments. In *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems (CHI '24)*. ACM, New York, NY, USA, Article 278, 278:1–278:17 pages. doi:10.1145/3613904.3642695

[18] Laura Lawton. 2017. Taken by the Tamagotchi: How a toy changed the perspective on mobile technology. *The iJournal: Student Journal of the Faculty of Information* 2, 2 (2017).

[19] David Ledo, Steven Houben, Jo Vermeulen, Nicolai Marquardt, Lora Oehlberg, and Saul Greenberg. 2018. Evaluation Strategies for HCI Toolkit Research. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (CHI '18)*. ACM, New York, NY, USA, Article 36, 17 pages. doi:10.1145/3173574.3173610

[20] Yoonji Lee and Chang Hee Lee. 2026. When Plants Play: Rethinking Plant Materiality in Digital Games. In *Proceedings of the 2026 CHI Conference on Human Factors in Computing Systems (CHI '26)*. Association for Computing Machinery, New York, NY, USA, Article 1060, 17 pages. doi:10.1145/3772318.3791373

[21] Jiahao Li, Jeeeun Kim, and Xiang 'Anthony' Chen. 2019. Robiot: A Design Tool for Actuating Everyday Objects with Automatically Generated 3D Printable Mechanisms. In *Proceedings of the 32nd Annual ACM Symposium on User Interface Software and Technology* (New Orleans, LA, USA) *(UIST '19)*. Association for Computing Machinery, New York, NY, USA, 673–685. doi:10.1145/3332165.3347894

[22] Zhe Liu, Anbang Xu, Yufan Guo, Jalal U. Mahmud, Haibin Liu, and Rama Akkiraju. 2018. Seemo: A Computational Approach to See Emotions. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (Montreal QC, Canada) *(CHI '18)*. Association for Computing Machinery, New York, NY, USA, 1–12. doi:10.1145/3173574.3173938

[23] Jennifer Mankoff, Anind K. Dey, Gary Hsieh, Julie Kientz, Scott Lederer, and Morgan Ames. 2003. Heuristic Evaluation of Ambient Displays. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '03)*. ACM, New York, NY, USA, 169–176. doi:10.1145/642611.642642

[24] Brad A. Myers, Scott E. Hudson, and Randy Pausch. 2000. Past, Present, and Future of User Interface Software Tools. *ACM Transactions on Computer-Human Interaction* 7, 1 (2000), 3–28. doi:10.1145/344949.344959

[25] Tru Narla. 2026. I'm Making a Tamagotchi for My Plant! Instagram reel, @mewtru. https://www.instagram.com/reel/DVxD03DEW2x/ Accessed: 2026-07-28.

[26] Clifford Nass and Youngme Moon. 2000. Machines and Mindlessness: Social Responses to Computers. *Journal of Social Issues* 56, 1 (2000), 81–103. doi:10.1111/0022-4537.00153

[27] Ivan Poupyrev, Philipp Schoessler, Jonas Loh, and Munehiko Sato. 2012. Botanicus Interacticus: interactive plants technology. In *ACM SIGGRAPH 2012 Emerging Technologies* (Los Angeles, California) *(SIGGRAPH '12)*. Association for Computing Machinery, New York, NY, USA, Article 4, 1 pages. doi:10.1145/2343456.2343460

[28] Zachary Pousman and John Stasko. 2006. A Taxonomy of Ambient Information Systems: Four Patterns of Design. In *Proceedings of the Working Conference on Advanced Visual Interfaces (AVI '06)*. ACM, New York, NY, USA, 67–74. doi:10.1145/1133265.1133277

[29] Raspberry Pi Foundation. 2026. Run a large language model on your Raspberry Pi. https://projects.raspberrypi.org/en/projects/llm-rpi. Accessed: 2026-03-29.

[30] Byron Reeves and Clifford Nass. 1996. *The Media Equation: How People Treat Computers, Television, and New Media Like Real People and Places.* CSLI Publications and Cambridge University Press, Stanford, CA.

[31] Marco Rozendaal. 2016. Objects with intent: a new paradigm for interaction design. *Interactions* 23, 3 (April 2016), 62–65. doi:10.1145/2911330

[32] Marco Rozendaal, Boudewijn Boon, and Victor Kaptelinin. 2019. Objects with Intent: Designing Everyday Things as Collaborative Partners. *ACM Transactions on Computer-Human Interaction* 26, 4, Article 26 (2019), 26:1–26:30 pages. doi:10.1145/3325277

[33] James A Russell and Lisa Feldman Barrett. 1999. Core affect, prototypical emotional episodes, and other things called emotion: dissecting the elephant. *Journal of Personality and Social Psychology* 76, 5 (1999), 805.

[34] Ben Shneiderman. 2000. Creating Creativity: User Interfaces for Supporting Innovation. *ACM Transactions on Computer-Human Interaction* 7, 1 (2000), 114–138. doi:10.1145/344949.345053

[35] Sherry Turkle (Ed.). 2007. *Evocative Objects: Things We Think With.* The MIT Press, Cambridge, MA.

[36] Sherry Turkle, Will Taggart, Cory D. Kidd, and Olivia Dasté. 2006. Relational Artifacts with Children and Elders: The Complexities of Cybercompanionship. *Connection Science* 18, 4 (2006), 347–361. doi:10.1080/09540090600868912

[37] Xuetong Wang, Ching Christie Pang, and Pan Hui. 2025. Talking Spell: A Wearable System Enabling Real-Time Anthropomorphic Voice Interaction with Everyday Objects. In *Proceedings of the 38th Annual ACM Symposium on User Interface Software and Technology (UIST '25)*. ACM, New York, NY, USA, Article 117, 117:1–117:17 pages. doi:10.1145/3746059.3747617

[38] Yu Wang, Yulu Lu, Shuo Yan, and Xukun Shen. 2025. "If My Apple Can Talk": Exploring the Use of Everyday Objects as Personalized AI Agents in Mixed Reality. In *Extended Abstracts of the 2025 CHI Conference on Human Factors in Computing Systems (CHI EA '25)*. ACM, New York, NY, USA, Article 16, 16:1–16:9 pages. doi:10.1145/3706599.3720253

[39] Waveshare. 2026. ESP32-S3-ePaper-1.54. https://docs.waveshare.com/ESP32-S3-ePaper-1.54. Waveshare Documentation Platform. Accessed: 2026-03-27.

[40] Adam Waytz, John Cacioppo, and Nicholas Epley. 2010. Who Sees Human?: The Stability and Importance of Individual Differences in Anthropomorphism. *Perspectives on Psychological Science* 5, 3 (2010), 219–232. doi:10.1177/1745691610369336

[41] Mark Weiser and John Seely Brown. 1996. Designing Calm Technology. *PowerGrid Journal* 1, 1 (1996).

[42] Craig Wisneski, Hiroshi Ishii, Andrew Dahley, Matt Gorbet, Scott Brave, Brygg Ullmer, and Paul Yarin. 1998. Ambient Displays: Turning Architectural Space into an Interface between People and Digital Information. In *Proceedings of the First International Workshop on Cooperative Buildings (CoBuild '98)*. Springer, 22–32. doi:10.1007/3-540-69706-3_4

A Appendix

A.1 Avatar/Emotion Generation

Expanding on the details in Section 3.4, our system has a three-step pipeline to keep the consistent avatar features and expressive emotions: (1) base avatar generation, (2) emotion generation, (3) animation generation.

A.1.1 Base Avatar Generation. The first step of avatar generation is to create a base avatar that reflects the object description provided by the user. We used Gemini Nano Banana [8] for this and all following image generations.

```
Create a pixel art character based on this object: {object_type}

The character should be an expressive avatar that can show
      different emotions.

## Output format
- The character should be drawn to fill the image area as much as
      possible
- White background
- Black and white color scheme only
- No text in the image
- Simple, clean pixel art style
- The character should have a neutral expression

Generate the base character now.
```

A.1.2 Emotion Generation. Next, we generate images for 12 emotions and the 'curious' expression. To preserve context during emotion generation, we start a chat session and provide the base avatar image along with the prompt below:

```
Now generate the SAME character showing the emotion: {emotion}
IMPORTANT: Keep the exact same character design, style, and
      proportions. Only change the facial expression and body
      language to show {emotion}.

## Output format
- The character should be drawn to fill the image area as much as
      possible
- White background
- Black and white color scheme only
- No text in the image
- Same pixel art style as before
```

A.1.3 Animation Generation. To preserve context during animation generation for each emotion, we create a separate chat session for each emotion state and generate its animation frames within that session. Along with each emotion avatar image, we provide the prompt below:

```
Generate frame {frame_idx + 1} of {frame_num} in a looping idle
      animation of a character showing the emotion: {emotion}.
* Input: the previous frame image.
* Output: the next frame with only subtle motion changes (e.g.,
      slight body bob, breathing, blinking).
* The facial expression and emotional state MUST stay {emotion}
      throughout every frame - do NOT drift or transition to any
      other emotion.

IMPORTANT: Keep the exact same character design, style, and
      proportions as the input image. Only add minimal idle motion.

## Output format
- The character should be drawn to fill the image area as much as
      possible
- White background
- Black and white color scheme only
- No text in the image
```

A.2 Behavior Customization

As described in Section 3.5, our system generates positive and negative behavior text to facilitate OAT authoring, giving users a starting point for refinement, if needed at all. The prompt below is used to generate the positive behavior text. We generate the negative behavior text in the same manner, using opposite phrasing.

```
You are an assistant that maps positive environmental and human
      interaction states related to specific sensor capabilities.
Always respond in exactly three concise sentences.
Example answer when Object is 'Sea Onion Cactus' and Sensors are
      ['Temperature', 'Humidity', 'Water / Moisture', 'Light
      Intensity']:
Temperature warm and moderate is good, ideally between 10°C to 24°C.
Low to average household humidity, 30% to 50%, is ideal. Do not
      need to mist this plant.
A soil moisture sensor detecting a brief spike in water followed by
      a long dry period indicates proper watering behavior.
Bright, indirect light to partial sun is good.

Object: {object_type}
Available Sensors: [{sensors}].
What human behaviors or environmental changes make this object
      function optimally or 'feel positive'?
Explain this strictly through the lens of what the available
      sensors would detect.
Do not use rhetoric. Try to be concise. Provide plain text within
      exactly three sentences.
```

A.3 Emotion Reasoning

As we described in Section 3.11, we use Gemini 3 Flash Preview [9] for emotion reasoning based on sensor data. Below is the prompt:

```
Object: {object_desp}
{datetime}
Sensors: {sensor_json}
Positive behaviors (good state): {positive}
Negative behaviors (bad state): {negative}
Available emotions: {emotions}

Use both the current reading and historical trend (sampled hourly,
      based on 2-second samples aggregated into hourly averages) to
      choose exactly ONE emotion that best matches the sensor state.
      Consider the current time of day and date for context. Briefly
      explain why.

Reply in exactly this format (two lines, nothing else):
EMOTION: <emotion>
REASON: <one sentence explanation>
```

Specifically for the sensor JSON data, we provide metadata for each sensor to better ground the model's interpretation. Below, we show the example sensor data (that can be varied) together with the accompanying sensor descriptions (that are fixed) and format provided to the model. Note that, although all sensors are shown here, only the sensors selected by the user are included in the input.

```
  "sensors": {
    "temperature": {
      "current_value": 21.5,
      "history": [24.5, 23.9, 23.4],
      "unit": "°C",
      "sensor value description": "Ambient air temperature."
```

```
  },
  "humidity": {
    "current value": 15.1,
    "history": [39.5, 41.7, 46.7],
    "unit": "% RH",
    "sensor value description": "Relative humidity."
  },
  "sound": {
    "current value": -55.8,
    "history": [-55.2, -56.8, -56.1],
    "unit": "dBFS",
    "sensor value description": "Sound level. Silence/Noise Floor
    (-60 dBFS and below), Quiet (-40 to -60 dBFS), Dialogue or
    music (-40 to -20 dBFS), Loud (-20 to 0 dBFS)"
  },
  "water": {
    "current value": 30.3,
    "history": [46.1,40.8,30.8],
    "unit": "%",
    "sensor value description": "Water moisture (sensor often
    used for soil moisture sensing). 100% means immersed in water.
    0% means totally dry, like held in the air. Values in between
    suggest partial moisture."
  },
  "light": {
    "current value": 38.8,
    "history": [38.7,38.7,38.9],
    "unit": "lux",
    "sensor value description": "Light intensity. Higher values
    indicate brighter light exposure. Dark i s around 7 lux;
    office lighing is around 127 lux; next to a window in the day
    is around 600 lux; full su n is often above 2000 lux"
  },
  "motion": {
      "current value": 100.0,
      "history": [0.0,0.0,0.0],
      "unit": "%",
      "sensor value description": "Motion of the device as a
    percentage. 100% indicates being in motion, while 0% means no
    motion. For historical data, motion is sampled every 2 seconds
    and aggregated by hour, so values in between 0 and 100 show
    what % of 2-second samples in that hour there was motion."
  },
  "occupancy": {
      "current value": 100.0,
      "history": [0.0,0.0,0.0],
      "unit": "%",
      "sensor value description": "Indicates human occupancy near
    the device over time. 100% indicates human occupancy near the
    device, while 0% indicates no people around the device. For
    historical data, motion is sampled every 2 seconds and
    aggregated by hour, so values in between 0 and 100 show what %
    of 2-second samples in that hour there was human occupancy."
  }
}
```