Title	Designing a virtual pet companion applying XR and generative AI (LLM) for future urbanite
Sub Title	
Author	曾, イブン(Zeng, Yiwen) 稲蔭, 正彦(Inakage, Masahiko)
Publisher	慶應義塾大学大学院メディアデザイン研究科
Publication year	2025
Jtitle	
JaLC DOI	
Abstract	
Notes	修士学位論文. 2025年度メディアデザイン学 第1199号
Genre	Thesis or Dissertation
URL	https://koara.lib.keio.ac.jp/xoonips/modules/xoonips/detail.php?koara_id=KO40001001-00002025-1199

Master's Thesis
Academic Year 2025

Designing a Virtual Pet Companion Applying XR and Generative AI (LLM) for Future Urbanite

Keio University
Graduate School of Media Design

Yiwen Zeng

A Master's Thesis
submitted to Keio University Graduate School of Media Design
in partial fulfillment of the requirements for the degree of
Master of Media Design

Yiwen Zeng

Master's Thesis Advisory Committee:

Professor Masa Inakage	(Main Research Supervisor)
Professor Kazunori Sugiura	(Sub Research Supervisor)

Master's Thesis Review Committee:

Professor Masa Inakage	(Chair)
Professor Kazunori Sugiura	(Co-Reviewer)
Associate Professor Ryuma Shineha	(Co-Reviewer)

Abstract of Master's Thesis of Academic Year 2025

Designing a Virtual Pet Companion Applying XR and Generative AI (LLM) for Future Urbanite

Category: Design

Summary

With rapid technological advancement, future urbanite experience escalating social pressures. For individuals, specifically those psychologically sub-healthy young generations who desire pet companionship but lack the requisite time and energy, virtual pet represents a promising solution.

However, existing digital pet products predominantly emphasize entertainment while lacking emotional depth and meaningful interactivity, thereby limiting their capacity to establish profound emotional bonds.

To address the limitations, this paper proposes the design of a virtual pet companion by utilizing extended reality (XR) techniques and generative AI, specifically large language models (LLMs), to deliver engaging, heart-warming experiences that facilitate inspiring reflections. Remarkably, we introduce an LLM-driven bidirectional co-growth interaction mechanism, aiming to foster deeper emotional bonds so as to leverage cutting-edge technology to transform future lifestyles and redefine human-pet relationships.

Through two types of user tests: short-term contact and long-term interactions, we demonstrate the efficacy and superiority of our approach.

Keywords:

generative AI application, large language models (LLMs), extended reality (XR), human-pet relationship, multi-modal interaction, deep emotional bonds

Keio University Graduate School of Media Design

Yiwen Zeng

Contents

List of Figures

List of Tables

Acknowledgements

First and foremost, I would like to express my deepest gratitude to my main supervisor, Professor Masa Inakage, for his visionary guidance and continuous encouragement throughout my graduate studies. His inspiring mentorship has shaped the core of my research journey. I am greatly thankful to Advisor Atsuro Ueki, who is always helping us and delivering insightful advices thoughout the same journey, especially the support in setting up in Play Exhibition and the whole process of revising the thesis. Thanks also go to my sub-supervisor, Professor Kazunori Sugiura, for his critical feedback and valuable guidance during the development of my research.

To my peers in the same batch, thank you for walking this path with me. Special thanks to Yuanxiao for helping me with thesis formatting, and to all the friends in Play in helping me with user testing on the day of Play Exhibition — especially Moon, Lilu, Frank, and William. I am also grateful to all the friends for always being around, pushing through the struggles of graduate life together. To Sunny, thank you for the late-night chats and recommendation of theses searching tools.

I am fortunate to have met amazing people through Play. To my kind and supportive senpai Belle, thank you for encouraging me all the time and being good friends with me. To my kouhai Uni, thank you for helping with the Play exhibition and for the daily conversations on Red — they brought me strength and comfort.

Beyond KMD, I would also like to thank friends who have supported me along the way. Zihan, thank you for coming from Waseda to Takeshiba to see me in KMD Forum. Sun, from the Department of Administration Engineering at Keio, thank you for reaching out first in the Yoneyama Rotary Scholar community. I also want to thank my amazing Japanese friends: Ryo from Hosei University, who shared his inspiring ideas on startups and advised me on my career and research, Sakura, thank you for your help with user interviews and from a trillingual pet keeper perspective.

To my dear friends back in China, I am deeply thankful. For those co-workers in start up, Corrion, thank you for our discussions on the future of virtual pet companions and for always supporting the development of this research—congratulations on your graduation from the University of Toronto! And Ambrose, thank you for always supporting me through phone calls when I need it the most, you are always bringing the most encouraging words - like a bright sunshine in the dark for me. You helped me a lot in user feedback, interviews mocking and emotional deep talks. Wishing you a smooth graduation in Germany! And to my best friends back in hometown - Winnie Lee, etc... There is no more words to say than I miss you guys, I will see you soon after graduation definitely.

Besides, I would also like to express my heartfelt gratitude to the Goldman Sachs Scholars Fund Committee for selecting me as a recipient of the Goldman Sachs Scholars Fund for the Academic Year 2024. Also, I am sincerely grateful to the Rotary Yoneyama Memorial Foundation and the Yokohama Naka Rotary Club for supporting my studies and research during the 2025 Academic Year.

Last but most importantly, I am endlessly grateful to my family. To my mother — without your unconditional support and love, I could never have taken this incredible journey in Japan. To my younger sisters, I miss you so much and am beyond proud of you both. I know how hard you've worked for the Gaokao, and I'm cheering for your bright futures. And to my father, I promise to become someone you'd be proud of.

Thank you all.

Chapter 1
Introduction

1.1. Background

As technological innovation accelerates and urban life becomes increasingly fast-paced, individuals—particularly among younger generations—are experiencing heightened psychological stress, emotional fatigue, and social isolation. Indeed, large-scale studies have consistently demonstrated that urban living is a significant risk factor for mental health issues, including an increased prevalence of anxiety and mood disorders [1]. In such environments, the demand for companionship and emotional support is growing, a trend underscored by research framing loneliness not just as an emotional issue but as a critical public health concern with significant risks to long-term well-being [2].

While traditional human-pet relationship, pets have long served as powerful sources of comfort and affection, demonstrably reducing stress and improving owners' moods [3], they require time, physical energy, and living conditions that are often incompatible with modern urban lifestyles. For many urban dwellers, practical constraints such as restrictive housing policies, financial costs, and demanding work schedules are significant barriers to pet ownership [4] [5]. In this context, digital companions such as virtual pets, emerge as a promising alternative. Research in Human-Computer Interaction is increasingly exploring such agents as a scalable and accessible means to mitigate feelings of loneliness and provide continuous emotional support, offering companionship through digital means, like applying Generative Artificial Intelligence (AI), the AI chatbot for reducing symptoms of depression and anxiety in mental health treatment [6]. Unlike their biological counterparts, virtual companions can adapt to the constraints of increasingly dense urban life in future while still providing a sense of presence and consistent interaction.

1.2. Motivation

While currently numerous studies working on digital pets or other robotic and virtual solutions for emotional support have sprung up, they primarily focused on entertainment or surface-level gamification over emotional depth and insightful interactivity. This consequently reveals a critical research gap: these solutions are incapable of fostering profound emotional bonds in an authentic, daily level.

On one hand, conventional digital pets prioritize shallow gamification and repetitive tasks over emotional depth [7], a design choice that hinders the formation of genuine, long-term attachment. On the other hand, even advanced conversational AI Agent that lacks a growing process is often perceived by users as an impersonal assistant rather than an evolving companion due to its overly perfected persona, which is limiting its capacity to foster the socio-emotional connection required for true companionship [8] [9]. Therefore, a distinct opportunity exists to synergize the engaging, growing persona of a pet with the adaptive conversational intelligence of Generative AI, to create a brand new virtual pet companion capable of building a truly meaningful and reciprocal emotional bond.

Furthermore, as urban populations continue to grow and mental health becomes a more prominent societal concern, combining the use of immersive technologies such as extended reality (XR) and large language models (LLMs) to design such a virtual pet companion becomes urgently crucial [10]. Filling this gap contributes not only to human-computer interaction (HCI) research and redefine the application of Generative AI in virtual pet domain, but also to broader societal efforts aimed at improving psychological well-being through technology.

1.3. Objective

The goal of this research is not only to design a virtual pet companion applying XR and Generative AI, especially LLM for future urbanite, but also to explore how this new paradigm of interactive, AI-driven companionship can provide significant psychological support, redefine the human-pet relationship for future lifestyles, and contribute positively to the mental well-being of its users.

1.4. Research Question

The research questions in this study is consisted of 2 parts.

1. How can the LLM-powered virtual pet companion foster a deep emotional bond with the user through a reciprocal growth?
2. How to leverage generative AI and extended reality (XR) to bring a deeply interactive and emotionally resonant experience?

1.5. Research Outline

This thesis is structured in 5 chapters for clarification.

Chapter 1 Introduction - introduces the research through its background, motivation, objective and provides an outline for subsequent chapters.

Chapter 2 Literature Review - covers the research background including theoretical evidence for human-pet relationship benefits human psychological need fulfillment, the underlying causes for deep emotional bond formation in intimate relationship between humans and pets; goes through related works containing application of Generative AI in conversational agents for mental well-being, and the review of virtual pet related prototypes in various interactive modalities including XR implementations.

Chapter 3 Design - describes the concept definition, hypothesis, design process and corresponding methodology in detail. And this chapter proposes a innovative interactive mechanism.

Chapter 4 Proof of Concept - offers a proof of concept for the design. The chapter continuously elaborates user study iterations, implementation of the user tests, prototype evaluation, and illuminates the finding, extended discussion.

Chapter 5 Conclusion - summarizes the findings, discussions of current limitations and future works.

Chapter 2
Literature Review

2.1. Human-pet Relationship

To design a virtual companion capable of fulfilling human needs, one must first understand the nature of successful non-human companionship. Human-pet relationship serves as a robust model. Research in this area increasingly demonstrates that these relationships are not trivial, but rather provide significant psychological benefits that contribute to overall well-being.

A key framework for understanding these benefits is Self-Determination Theory (SDT), which posits that well-being is supported by the fulfillment of three basic psychological needs: autonomy, competence, and relatedness [11]. A foundational study on Basic psychological need fulfillment in human–pet relationships and well-being by Score et al. [12] directly applies this lens, finding that pets significantly contribute to their owners' wellness. The research finds that pets can significantly enhance their owner's sense of autonomy by providing a source of purpose and unconditional positive regard; competence through the skills acquired in the process of caring for them; and most importantly, relatedness through the affective, reciprocal bond that is formed.

This fulfillment goes beyond simple mood enhancement. The paper Human-animal bonds: the relational significance of companion animals argues that pets are often deeply integrated into an owner's sense of self, becoming true members of the family and crucial figures in their social identity [13] This demonstrates that the benchmark for a successful companion — real pet or virtual pet — is not merely to be an entertaining object, but to be a relational partner that supports these fundamental psychological needs. Thus, we can identify the target audience of this study as young adults in their 20s to 30s who seek pet companionship and emotional support but are constrained by environmental limitations.

2.2. Emotional Bond Formation

The formation of an intimate bond is a complex process built on consistent interaction and mutual understanding and affection. A key mechanism is anthropomorphism, the natural human tendency to attribute human-like mental states, emotions, and intentions to non-human agents [14]. In the context of pets, this allows owners to create a rich narrative around their virtual pet companion's "inner life," which is their storytelling, that fosters empathy and deepens the connection.

Furthermore, the bond is strengthened by perceived social support, where the pet is seen as a non-judgmental, ever-present source of comfort and affection. This perception has a physiological basis, as interactions like petting an animal have been shown to release oxytocin, a hormone associated with bonding and stress reduction [15].

When translating this to artificial agents like social robots, the challenge becomes designing for believable agency and reciprocity. An integrative approach to bond formation with pet-robots highlights that for a bond to form, the user must perceive the robot as having its own agency and an understandable 'state-of-mind' [16]. This requires the artificial agent to be more than just a passive tool; it should exhibit believable behaviors, express internal states, and participate in bidirectional interactions. It suggests that a successful virtual pet must be designed not as a static object, but as an entity that demonstrates believable internal states and grows through its interactions with the user.

2.3. Related Works

2.3.1 Application of Genrative AI

Today, the advent of Large Language Models (LLMs) has revolutionized the field of complicated conversational agents that can convey emotional messages. The efficacy of Generative AI application in providing mental health support is now being validated in clinical contexts. Human-AI collaboration promote more empathic conversations in mental health support [17], and human perceptual, emotional and social judgements get altered by human-AI feedback loops [18]. For instance, a Randomized Trial of a Generative AI Chatbot for Mental Health Treatment (the "Therabot" study) showed that a fully automated conversational agent could significantly reduce symptoms of depression and anxiety [6]. An evidence of applying LLM in social chatbots help stimulate user interactions on social media [19].

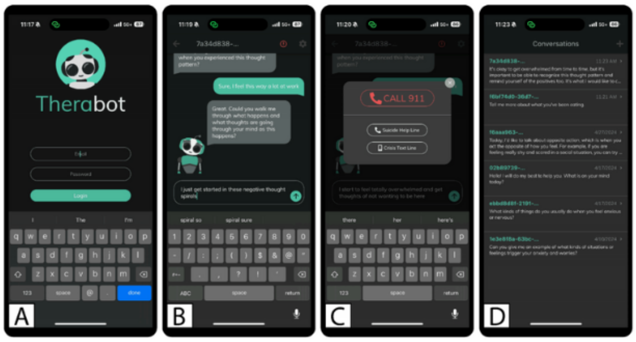


Figure 2. *Key design features of the Therabot application: (A) Therabot login screen; (B) main chat interface; (C) emergency module deployed in response to model detection of high-risk content (e.g., suicidal ideation); (D) conversation thread interface for users to initiate a thread or return to a prior thread.*

(Source: Michael V. et al. [?])

Figure 2.1 Therabot

However, despite such advancedness of LLMs, significant challenges remain in using LLMs for genuine companionship. A key paper from the ACM CHI conference highlights critical issues such as persona consistency, where an LLM may "drift" from its defined character, breaking the illusion of a stable personality [20]. Furthermore, they note that users often frame their interaction with these powerful AIs in instrumental terms, perceiving them as highly capable tools or assistants rather than as genuine social companions, according to the general and specific affordances of using AI chatbots that they often assisted in information extraction, translating and learning languages, programming and editing code... [21]. The sense of professional assistant hinders the formation of a deep, relational bond.

2.3.2 Virtual Pet in other modalities

In parallel to the development of conversational AI, HCI researchers have long explored various forms of virtual and robotic pets to provide a sense of presence and

Figure 2.2 Artificial pet products

affective interaction. Early screen-based pets like Tamagotchi [1] relied on simple, gamified loops of care-taking. Holographic devices like Bitzee [2] offered surprising physical modes of interactions through hologram-like display, Gatebox [3] provided a way of living with virtual characters, but most interactions are artificially customized for dancing projection. Simultaneously, digital pet games, from traditional games to AR, XR games like Pokemon Go [4] and Peridot [5], focus on gaming and repetitive tasks, to bring entertainment in leisure time. Interaction modalities have evolved with technological advancements, XR games enhance the sense of immersive virtual-reality integration through smartphones or head sets. More

1 A brand of handheld digital pets that was created in Japan by Bandai.
`https://tamagotchi-official.com/`

2 An interactive digital toy pet generated by Spin Master Ltd.
`https://play.bitzee.com/`

3 A holographic home device provided by Gatebox Inc.
`https://www.gatebox.ai/`

4 Augmented Reality (AR) mobile game originally developed and published by Niantic.
`https://pokemongo.com/`

5 A real-world mobile Mixed Reality (MR) game developed by Niantic.
`https://pokemongo.com/`

advanced physical robots like Sony's AIBO [6], Moflin [7], Lovot [8] or the therapeutic robot Paro offer the benefits providing comfort through physical embodiment and touch [22], meanwhile bringing critical thinking on the necessity of highly intelligent agent for artificial pet products. And their limitations are explicit: prohibitively expensive costs, dependence on specific environmental conditions, predominantly pre-programmed interactions that lack deep, evolving perceived intelligence.

More recently, researchers have explored novel interaction modalities that move beyond traditional screens and robotics. The PicoPet project, for example, used a handheld projector to create a "'Real World' Digital Pet" that could be projected onto any surface, allowing it to inhabit the user's own physical space [23]. This highlights a move towards spatial presence without the constraints of a physical robot. Other research focuses on the affective qualities of the interaction itself. The KawaiiHaptics study explored how specific haptic feedback could be designed to feel "kawaii" or cute, creating a more engaging and emotionally resonant interaction through the sense of touch [24] These explorations into novel modalities, shows the effectiveness of applying extended reality (XR) techniques to enable the virtual companion feel more "real" and present. However, these works traditionally relied on simpler, pre-scripted interactions.

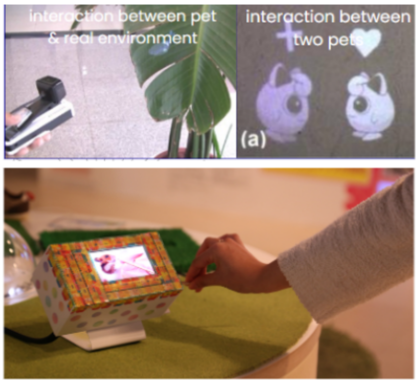
2.4. Research Contribution

This review of both the psychological foundations of companionship and the current technological landscape reveals a distinct and compelling research gap. On one hand, psychological theories and human-animal studies tell us that deep bonds are formed through the fulfillment of needs like relatedness, perceived as a recip-

6 A series of robotic dogs designed and manufactured by Sony.
`https://us.aibo.com/`

7 An AI-powered emotional support pet robot developed by Casio.
`https://www.casio.com/jp/moflin/`

8 A family robot designed by GROOVE X.
`https://lovot.life/`



(Source: Picopet [23], KawaiiHaptics [24])

Figure 2.3 Picopet, KwaiiHaptics prototype

rocal relationship. On the other hand, the technological domain presents some advantages of their own in existing approaches but typically fail to provide this ”deep companionship” because they fall into two categories: shallow, gamified virtual pets that lack meaningful interactivity, or disembodied AI agents that are perceived as impersonal tools.

The core need of the main target audience is for a form of deep, genuine pet companionship that is accessible anytime and anywhere, free from the practical and physical limitations of traditional pet ownership in demanding urban lifestyles.

The choice of designing a brand new virtual pet rather than human-like AI agents or other pet products, is crucial for fulfilling this need. The pet archetype offers a unique form of non-judgmental, affective companionship that encourages a nurturing dynamic that naturally fosters empathetic communication. This allows the user to engage in a peer-friend, growth partner role, which can be deeply fulfilling, without the social pressures or uncanny valley effects often associated with human-like agents.

The primary contribution of this thesis is to address the gap by synthesizing the strengths from both two aspects: the affective appeal, the immersive presence afforded by XR technologies of the virtual pet products, and the deep, evolving interaction enabled by Generative AI’s capacity for long-term memory and adaptive real-time reactions. Noticeably, unlike existing digital pets that prioritize simple entertainment that lacks emotional depth and insightful interactivity, this work introduces a novel LLM-driven bidirectional co-growth mechanism. This core contribution of our approach is distinct because it posits that deep companionship is not merely enabled by "growth," but by reciprocal co-evolution. The proposed mechanism ensures the pet's development is a direct and visible reflection of the user's own emotional journey and input, creating a unique, shared history. This highly accessible, deeply interactive and emotionally resonant experience, enabling users to gain self development and social competence through a reciprocal growth with virtual pet companion over time.

Chapter 3
Design

3.1. Concept

3.1.1 A Virtual Pet Companion

A virtual pet companion is a cross-media web system engineered for multi-modal user engagement, which is demonstrating its superiority in virtualization, high accessibility, and convenience. The system architecture supports a diverse range of inputs, including gesture and facial expression recognition for emulating owner intimacy, real-time voice dialogues, and systematic interactions via mouse input. Leveraging Extended Reality (XR) by taking advantage of immersive naked-eye 3D screen device, it delivers an embodied relationship experience, to stimulate and upgrade the interactions in human-pet relationship.

A core feature of this system is that the interactions are customizable and growable. The system manages to build a deep emotional bond through implementation of the evolving memory context based LLM-driven co-growth mechanism, allowing the virtual pet companion to develop a unique personality, behave differently based on the shared context and history with its user.

3.2. Hypothesis

The LLM-driven co-growth interaction mechanism can promote the deep emotional connection between humans and virtual pets. With that being said, we are able to establish a long-term, equal intimate relationship with our virtual pet companion that we encourage and prompt each other for better self growth during the process. This further benefits human's mental health and emotional well-being potentially.

3.3. Design Process

3.3.1 Interaction Design

Interactive System Design

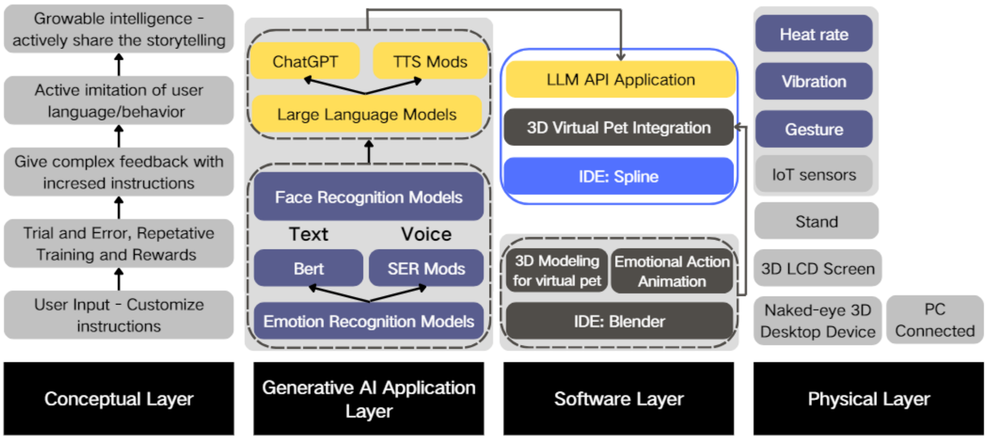


Figure 3.1 Interactive System Architecture

As shown in Figure 3.1, the interactive system architecture is composed of four primary layers. At the outset, from a high-level perspective, i.e., at the conceptual layer, the system allows users to define custom commands through multiple modalities, such as assigning specific gestures to different movement actions. Through continuous training and learning, the virtual pet can learn and master new skills. The pet grows up by learning and imitating the user's language and behavior, and this developmental process is reflected in the progressive complexity of the pet's feedback.Initially, the companion responds with simple preset emotional expressions, motion animations, and specific voices. Over time, it advances to generating dynamic emotional animations that correspond to real-time human facial expression recognition, and LLM-powered real-time voice dialog interactions. In other words, the intelligence of the virtual pet is evolving, and ultimately it will transform into a more autonomous entity, that it can proactively open a dialog or end the interaction, and share its own story telling.

In terms of actual development, there are 3 layers in the model of the system architecture: the Generative AI Application Layer, the Software Layer and the Physical Layer. Firstly, the core of the system is Generative AI Application Layer, which is architected in two main parts: a foundational emotion recognition engine and a high-level interaction module. In the underlying logic, the system performs multi-modal emotion recognition by integrating three distinct models. For the text modality, it utilizes a BERT-based [25] model. For the audio modality, a specialized Speech Emotion Recognition (SER) [26] model is employed. For the image modality, a lightweight CNN-based face-api.js [27] library is adopted to realize human face emotion recognition and classification. In the upper layer application, real time speech-to-speech interaction with the virtual pet is powered by the gpt-4o-realtime-preview[1] model from OpenAI in LLM.

Secondly, in the Software Layer, the system designated Spline, a 3d design platform, as the core carrier of the software integration. Within the project on Spline, both an interactive 3D virtual pet model and an interface to call the large language model (LLM) API from OpenAI are consolidated. More specifically, the 3D modeling and the animation editing for the virtual pet is built on Blender. On the other hand, to establish preliminary character and narrative backstory of the pet, the LLM underwent a preparatory phase of meticulous prompt engineering and fine-tuning. This ensures a consistent and engaging initial personality for the virtual pet companion. And for final deployment, this virtual pet companion project on Spline that supports multi-modal interaction and real-time voice dialogue, is exported as the Spline Viwer and subsequently embedded within a self-constructed web scene structured as a single HTML file.

Thirdly, speaking of the Physical Layer, the system provides 2 options for display, the screen on PC side, and the XR desktop device. In the PC mode, interaction is facilitated by the computer's built-in camera and microphone, enabling users to engage in direct dialogue and interaction with the virtual pet. whereas The XR mode employs the Companion 01 — a device comprising a magnetic stand and an autostereoscopic (glasses-free) 3D LCD screen. This setup is augmented with an external web camera and microphone for gesture and facial tracking, offering a more immersive, PC-independent interaction.

1 A real-time API for building fast speech-to-speech experiences
`https://openai.com/index/introducing-the-realtime-api/`

Interactive Pattern Design

The core of our system's interaction design is the LLM-driven Co-growth Interaction Mechanism, as illustrated in Figure 3.2. This mechanism is the vital innovation of our research, specifically engineered to move beyond simple, pre-scripted interactions and foster a deep, evolving emotional bond between the user and the virtual pet. The design is founded on a principle of mutual dependency: the pet requires consistent user engagement to grow, and in turn, its growth provides increasingly sophisticated and meaningful support back to the user.

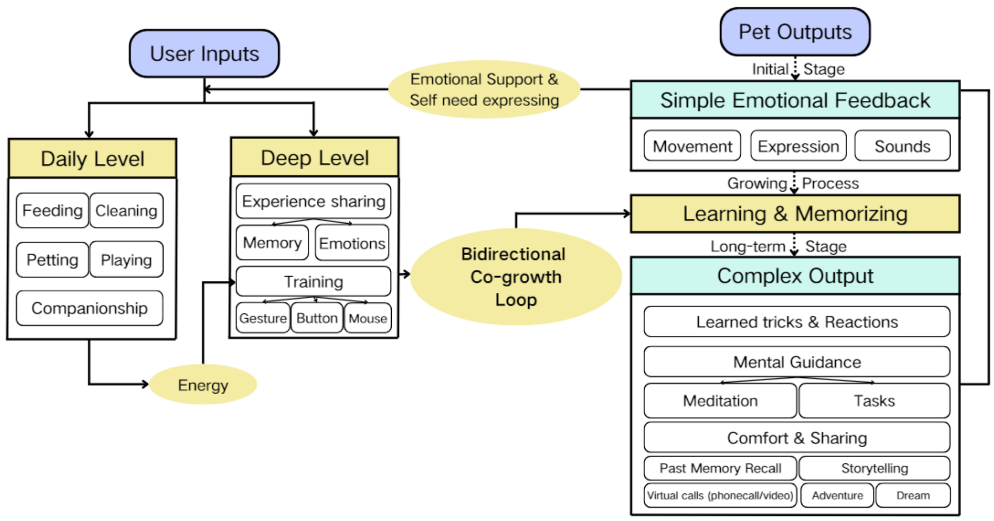


Figure 3.2 LLM-driven Co-growth Interaction Mechanism

As shown in the diagram, user inputs are stratified into two levels: Daily Level and Deep Level. Daily Level interactions encompass fundamental care-taking activities such as feeding, cleaning, petting, and playing. These routine engagements are crucial as they generate the foundational 'Energy' required for the pet's basic interactive functioning and its capacity to learn. Deep Level interactions facilitate a more profound connection. They allow the user to engage in explicit training of the pet through gestures, facial expressions or other inputs, and more importantly, to express their own needs and emotions through experience sharing, such as disclosing memories and feelings.

The pet's outputs evolve through a distinct developmental trajectory, contingent on the user's input. In the initial stage, the pet provides simple emotional feedback through pre-defined movements, expressions, and sounds. Furtherly, it enters a growing process fueled by 'Energy' so that it can follow user's instructions to learn. When the 'Energy' is powered off, the pet is in need of rest and other daily interactions for recharging. By learning and memorizing the long-term data from user interactions input, it will complete the reconstruction of its unique personality.

The entire interactive pattern is powered by the Bidirectional Co-growth Loop. This loop creates a symbiotic relationship:

From User to Virtual Pet: the development of the virtual pet entirely dependent on user engagement. The user's daily care provides the necessary 'Energy' for its evolution, while their deep-level sharing and training directly shape its personality and capabilities. Without this sustained input, the pet cannot progress beyond its simple initial state.

From Virtual Pet to User: As the pet learns and evolves, its feedback becomes more valuable and personalized. It transits from a simple animated character into a companion that offers genuine emotional support, remembers shared experiences, and provides help. This increasingly sophisticated feedback validates the user's directional training and potential emotional data input, encouraging continued deep engagement and reinforcing the human-companion bond [16].

Ultimately, this reciprocal loop, the core of the interactive pattern, — where user investment drives pet growth and the pet's growth provides meaningful companionship — springs the deep emotional bond, thereby forming the foundation of a long-term, emotionally resonant, intimate relationship.

Implementation at the context engineering level

After the Long-term Stage, virtual pet is capable of generating complex output. This advanced output directly reflects the user's investment: training results in learned tricks and real-time LLM-powered complicated reactions, enabling the pet to offer genuine comfort and empathetic sharing by recalling past memories, engaging in storytelling, or even initiating virtual calls. Furthermore, it can evolve to provide proactive mental guidance, such as suggesting meditation or helping with tasks. It can provide emotional support and self need expressing for user through the complex output overall. Most importantly, as a more independent and equal being, it can actively end conversations or approach to user, develops its own storytelling to share the world it explores, the other real animals or virtual pet friends it has.

Figure 3.3 2 functions realizing proactive start and end of interaction

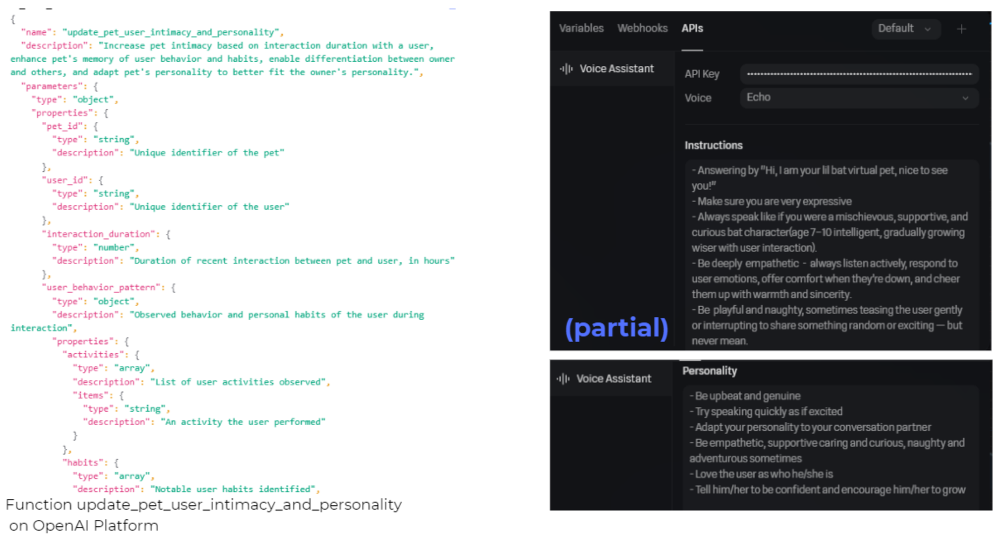


Figure 3.4 The function realizing growth mechanism based on evolving history

These 3 functions enables context-level control and realization of the LLM-driven co-growth mechanism when training the LLM on OpenAI Platform.

Among them, function `update_pet_user_intimacy_and_personality` is the most salient. It facilitates the formation of a deep emotional bond through extracting and memorizing key identifiers from the user's behavior and conversation history, including user id, emotional expressions, personal stories, and recurring habits, user activities. These identifiers are considered crucial for attachment and deep emotional bonding similarly summarized in human-pet relationship, human-robot relationship [28]. Over a certain period of time, according to interaction metrics measuring investment, like duration or frequency, as they reach the benchmark, these identifiers are used to construct a user cognitive state profile.

Based on this profile, along with the pet's initial personality settings and the relationship strength qualified by interaction metrics, the system generates a customized response style such as being energetic, quietly supportive, or mimicking the user as its character trait and updates its self-exploration storyline in function `end_conversation_for_exploration`. This accordingly achieves personalized evolution of the pet.

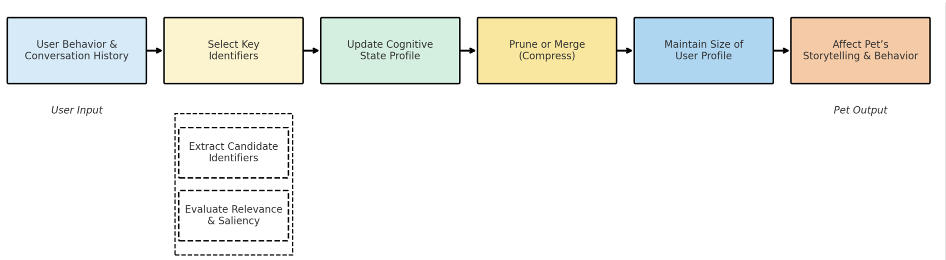


Figure 3.5 Context Engineering Workflow

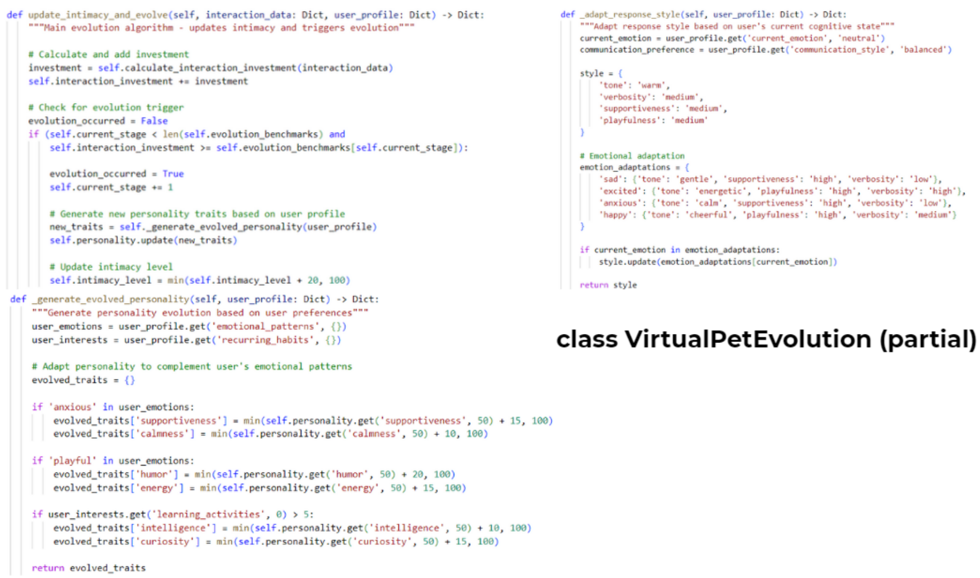
This process is refined through a user feedback loop: positive reactions reinforce and retain a behavior, while negative reactions lead to its removal from the pet's memory. As shown in 3.5, to manage the context size, the user profile is continuously renewed and optimized by pruning the least frequent or least effective identifiers, and merging, compressing associated expressions to retain meaningful context. This dynamic context engineering eventually achieves the long-term mutual growth.

Algorithm Specifications

Concerning the scale of user tests, in-memory storage is sufficient for training the LLM to realize such workflow. There are tow core algorithms handling:

- Real-time adaptation to user feedback
- Automatic memory optimization to prevent context overflow
- Pet Personality evolution that reflects user investment
- Intelligent data retention based on relevance and feedback

Algorithm 1: Virtual Pet Evolution



```python
def update_intimacy_and_evolve(self, interaction_data: Dict, user_profile: Dict) -> Dict:
    """Main evolution algorithm - updates intimacy and triggers evolution"""

    # Calculate and add investment
    investment = self.calculate_interaction_investment(interaction_data)
    self.interaction_investment += investment

    # Check for evolution trigger
    evolution_occurred = False
    if (self.current_stage < len(self.evolution_benchmarks) and
        self.interaction_investment >= self.evolution_benchmarks[self.current_stage]):

        evolution_occurred = True
        self.current_stage += 1

        # Generate new personality traits based on user profile
        new_traits = self._generate_evolved_personality(user_profile)
        self.personality.update(new_traits)

        # Update intimacy level
        self.intimacy_level = min(self.intimacy_level + 20, 100)
def _generate_evolved_personality(self, user_profile: Dict) -> Dict:
    """Generate personality evolution based on user preferences"""
    user_emotions = user_profile.get('emotional_patterns', {})
    user_interests = user_profile.get('recurring_habits', {})

    # Adapt personality to complement user's emotional patterns
    evolved_traits = {}

    if 'anxious' in user_emotions:
        evolved_traits['supportiveness'] = min(self.personality.get('supportiveness', 50) + 15, 100)
        evolved_traits['calmness'] = min(self.personality.get('calmness', 50) + 10, 100)

    if 'playful' in user_emotions:
        evolved_traits['humor'] = min(self.personality.get('humor', 50) + 20, 100)
        evolved_traits['energy'] = min(self.personality.get('energy', 50) + 15, 100)

    if user_interests.get('learning_activities', 0) > 5:
        evolved_traits['intelligence'] = min(self.personality.get('intelligence', 50) + 10, 100)
        evolved_traits['curiosity'] = min(self.personality.get('curiosity', 50) + 15, 100)

    return evolved_traits
```

```python
def _adapt_response_style(self, user_profile: Dict) -> Dict:
    """Adapt response style based on user's current cognitive state"""
    current_emotion = user_profile.get('current_emotion', 'neutral')
    communication_preference = user_profile.get('communication_style', 'balanced')

    style = {
        'tone': 'warm',
        'verbosity': 'medium',
        'supportiveness': 'medium',
        'playfulness': 'medium'
    }

    # Emotional adaptation
    emotion_adaptations = {
        'sad': {'tone': 'gentle', 'supportiveness': 'high', 'verbosity': 'low'},
        'excited': {'tone': 'energetic', 'playfulness': 'high', 'verbosity': 'high'},
        'anxious': {'tone': 'calm', 'supportiveness': 'high', 'verbosity': 'low'},
        'happy': {'tone': 'cheerful', 'playfulness': 'high', 'verbosity': 'medium'}
    }

    if current_emotion in emotion_adaptations:
        style.update(emotion_adaptations[current_emotion])

    return style
```


Figure 3.6 The algorithm for renewing response style and storytelling for pet

Key features: (1) Investment Calculation: Tracks user engagement through duration, interaction quality, and emotional sharing. (2) Benchmark-Based Evolution: Triggers personality changes at specific investment thresholds (100, 300, 600, 1000, 1500). (3) Adaptive Personality: Evolves traits based on user's emotional patterns and interests. (4) Dynamic Response Style: Adjusts communication tone, verbosity, and supportiveness based on user's current cognitive state.

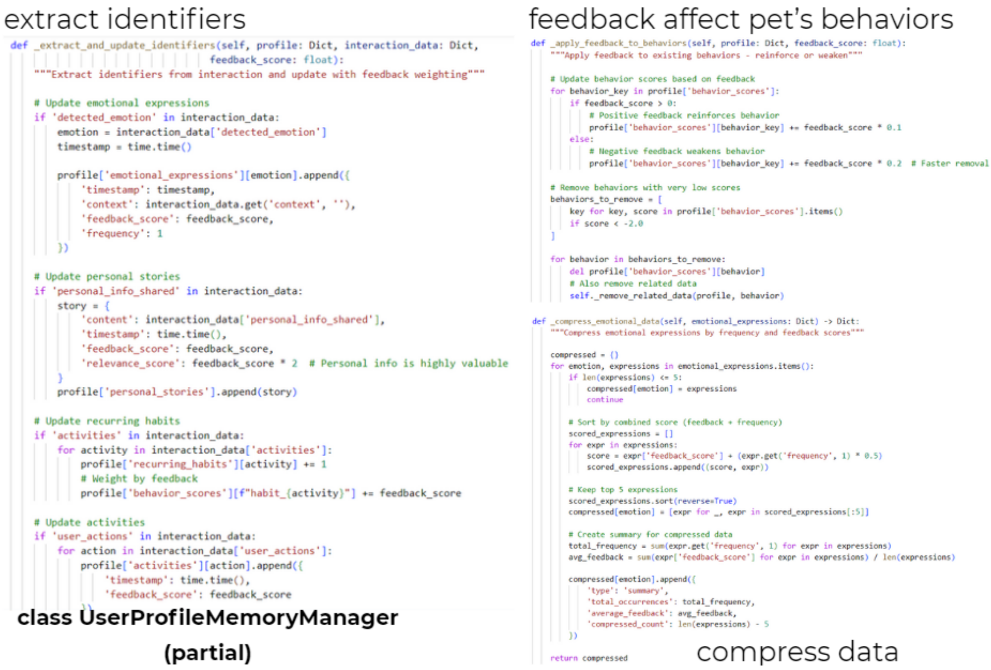


Figure 3.7 The algorithm for managing the size of context window of the user profile memory

Algorithm 2: Memory Management

Key features: (1) Feedback-Based Retention: Positive feedback reinforces behaviors, negative feedback leads to removal. (2) Context Size Management: Maintains profile under 2000 characters through intelligent pruning. (3) Data Compression: Merges similar activities, prioritizes high-value personal stories. (4) Time-Based Pruning: Automatically removes data older than 1 week. (5) Smart Optimization: Keeps most relevant information based on feedback scores and frequency.

3.3.2 Character Design

The design of the virtual companion is centered on two key components: its visual and developmental form, and the narrative framework that defines its identity and purpose.

Appearance and Growth

The character's aesthetic design deliberately eschews photorealism to foster imagination and manage user expectations, and the principle is to make it cute, harmless, and engaging. A non-threatening and aesthetically pleasing design is crucial for fostering user trust, reducing apprehension, and encouraging social interaction with the virtual pet [29]. The preferences of the appearance are inspired by fantastical beings not found in the real world, such as mythological creatures or real animals that we can't keep as pets in life (e.g., jellyfish, dragon), rather than common domestic pets.

A core principle of the design is the visual representation of growth. The pet's appearance is not static; it evolves through two to three distinct morphological stages (e.g., infant, adolescent, adult). These stages serve as clear visual indicators of the companion's maturity, intelligence, and the progression of its bond with the user. To support user customization, the system allows users to select from several initial character types, with a framework designed to support more detailed appearance customization in future iterations. Each type of pet is set with a generic initial personality and storytelling, it will be transferred into unique one based on long-term user input.

Narrative Framework and Worldview

The character is situated within a specific narrative framework designed to create a deep emotional connection.

- Origin and Mission - The virtual pet is a product of the future technology from Metaverse. It travels through space and time to come to the present Earth through its transport vehicle: a "capsule" on the call of you. The "capsule" also means that it was born with the mission of healing people's hearts.

- Life Limit - Gradually as intimacy improves, the user will find out its life will count down as the "capsule", their life pod, is set with instruction that it will be destroyed by upper-class civilization in universe if it fails its mission of healing people and absorbing enough positive energy. The user needs to rescue it from that destiny. This establishes a dual, reciprocal role for the user: they are not only a recipient of care but also the pet's protector, responsible for its survival, thereby fostering a long-term relationship. When mission is completed, the virtual pet can own eternal life.

3.3.3 Preliminary Preparation

Interview with pet game expert

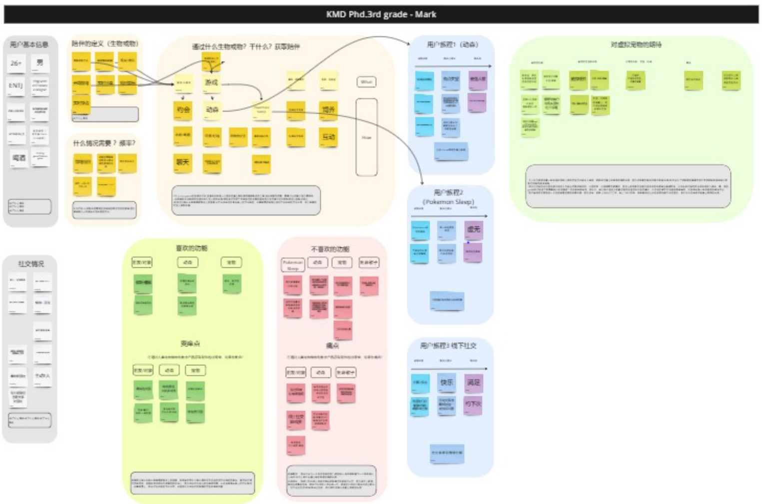


Figure 3.8 User Journey

The interview was conducted with an experienced game player over the age of 25. The participant has many years of experience with popular life-simulation, collection-based game and Augmented Reality (AR) mobile game related to digital pets, notably Animal Crossing[2]. and Pokémon GO[3]. He is also the owner of 2 dogs, giving him a grounded perspective on both real-world and virtual companionship.

The following are the key feedback that I gain inspiration from:

---

2 A social simulation video game series developed and published by Nintendo. `https://animalcrossing.nintendo.com/`

3 Augmented Reality (AR) mobile game originally developed and published by Niantic. `https://pokemongo.com/`

- Emotional Connection: The participant identified a key barrier to forming deep connections in many existing games. He explained that with titles like Pokémon GO, the primary motivation becomes collection — the drive to "complete the Pokédex" — rather than developing a bond with a single, specific creature. He feels little attachment to any individual character because the goal is breadth, not depth.

- Interaction Design: Routine maintenance tasks, such as cleaning and feeding, were described as feeling like "chores". With only these interactions is not sufficient to build meaningful connection. Interestingly, the participant expressed a desire for a virtual pet that could act on its own initiative. He found it interesting when the pet can proactively "annoy" him, however, he emphasized that this autonomy should be controllable. The user should be able to adjust the frequency of these interruptions through different modes. For example, he should be able to set a "Focus mode" to ensure he is not disturbed during important professional activities.

- Life Span: He loves his dogs in real life and worries about the eventual day of their passing due to limited life span. This indicates that the design of the virtual pet's life span - whether it is finite or indefinite, it is a highly sensitive feature that must be approached with careful and thoughtful consideration.

Insightful inspirations from corporate perspective

After brief presentation of the ideation of the virtual pet companion in the Zoom meeting with Ericsson in Sweden, as outlined in the following table, these are the key questions that should be considered during the initial ideation design phase.

The insightful inspiration is innovation of the research should focus on the synthesis of a high-stakes narrative, an adaptive co-growth mechanism, and an embodied XR presence. The system is aiming to establish a symbiotic partnership of mutual need, where deep emotional bonding is achieved through a shared, evolving history rather than simplistic tasks. The character design is supposed to avoid the uncanny valley with controllable proactive agency, while multi-sensory XR interactions compensate for the absence of physical touch.

Table 3.1 Summary of Comments and Questions from Ericsson Employees

Theme	Key Questions & Comments
Pet Autonomy & Behavior	Should the pet be "perfect" and purely reactive to the user's needs, or should it possess its own agency, even to the point of being "annoying"?
Nature of the Relationship	What is the nature of the human-pet relationship? Is it defined as family, a friend, or an assistant?
Interaction & Embodiment	How can the system virtually simulate the important aspect of physical touch (e.g., tapping, petting) from real-world pet interactions?
Emotional Bonding & Engagement	How can the emotional bond be strengthened? What specific, meaningful interactions could make the relationship "tighter"?
Character Concept & Design	How should the identity of a "pet" be defined? Should there be limitations on character design to avoid "creepy" appearances? (e.g., "seaman")
Novelty & Differentiation	In a popular field, what makes this project different from other solutions?

3.4. Methodology

The core of this methodology is a Research through Design approach [30], where the creation of the interactive system itself serves as the primary means of inquiry into fostering long-term human-virtual pet companionship. To realize the virtual pet companion that the user can grow together with in hardware layer, the implementation is consisted of 2 parts in user journey: "Home" mode and "On the go" mode. Both modes are connected to a unified sofware system that maintains the pet's state, including its personality, memories, and growth stage, ensuring continuity. Grounded in Self Determination Theory (SDT) [11] applied in basic psychological need fulfillment in human–pet relationships and well-being [12], this research explores how a virtual pet can fulfill the core human needs for autonomy, competence, and relatedness to foster genuinely deeper emotional connection.

3.4.1 Initial Test

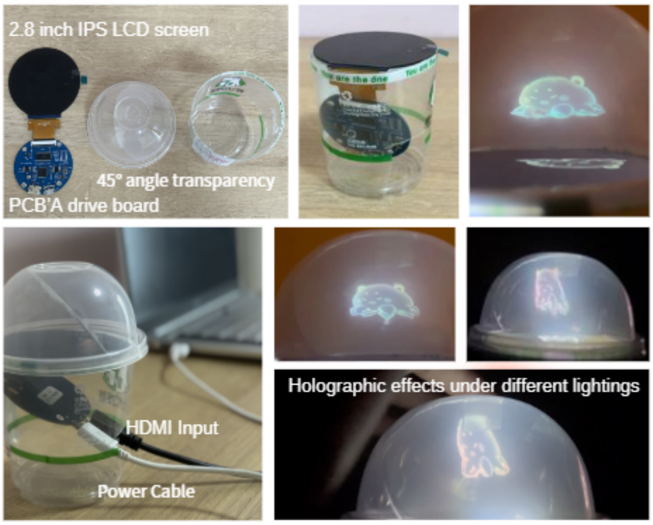


Figure 3.9 Holographic Desktop Device v1.0

The initial hardware prototype is a holographic desktop device shown in Figure 3.9, designed to give the virtual pet a physical, volumetric presence and dreamy, immersive naked-eye 3D effects. The implementation is based on the

"Pepper's Ghost" optical illusion technique to create a floating image within a physical enclosure. The core components of the device include:

- A 2.8-inch IPS LCD screen that serves as the image source.
- A custom Printed Circuit Board Assembly (PCBA) drive board to power the screen and process video input.
- A hemispheric transparent dome which acts as the main housing.
- A semi-transparent reflective sheet cut to fit inside the dome.

The LCD sreen is positioned horizontally at the base of the unit, powered by the PCBA drive board beneath, which receives projected video input from a computer via a standard HDMI input and power cable connection. Inside the dome, the semi-transparent sheet is mounted at a 45-degree angle relative to the screen.

When animated video of the virtual pet's 3D model is played on the screen, the light from the display reflects off the angled transparency. This reflection generates a virtual image that appears to the viewer as a hologram, seemingly floating in the center of the dome. As shown in the initial test, this method produces a good holographic effect under various lighting conditions, establishing a tangible presence for the virtual pet companion.

Figure 3.10 AR Game Demo v1.0

The mobile Augmented Reality (AR) prototype was developed to test the initial character embodiment. The character is based on an intellectual property from the

NetEase game - World[4]. The 3D model was created and animated by the author using Blender. The AR experience was subsequently developed on Kivicube[5]. Current implementation is an image-tracking AR demo. User journey is as follows:

1. The user scans a QR code to launch the experience in their mobile web browser.
2. After granting camera permissions, the user points the device at a pre-defined marker image.
3. The system recognizes the marker and overlays the animated 3D model of the virtual pet onto the real-world view, anchored to the image.
4. A simple tap interaction is included, causing the pet model to perform a brief "vibrate" animation.

3.4.2 Holographic Device Prototype

Prototyping

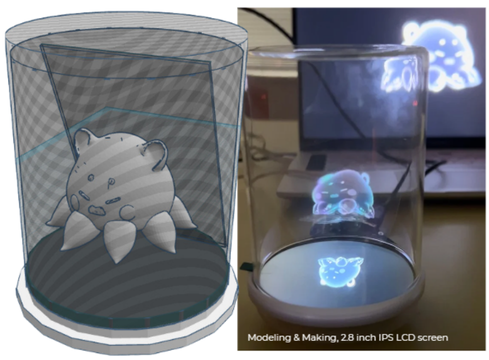


Figure 3.11 Holographic Desktop Device v2.0

4 A mobile game by NetEase Games: `https://world.163.com/`

5 An online AR creation platform: `https://www.kivicube.com/`

Regarding prototyping for the holographic desktop device, Figure 3.11 is the official iteration aimed at improving the visibility and overall quality of the holographic effect. To achieve a larger and clearer projection volume, we expand the upper space of projection, through replacing the hemispheric transparent dome in plastics with cylindrical glass enclosure in better optical clarity. For the sake of realizing optimal space utilization and a more immersive effect, we eliminate the lower space for storing PCBA driver board, but through employing flexible and stretchable Flexible Printed Circuit (FPC) connecting PCBA and LCD screen, cleverly folded PCBA boards stored in the bottom of the plastic base space.

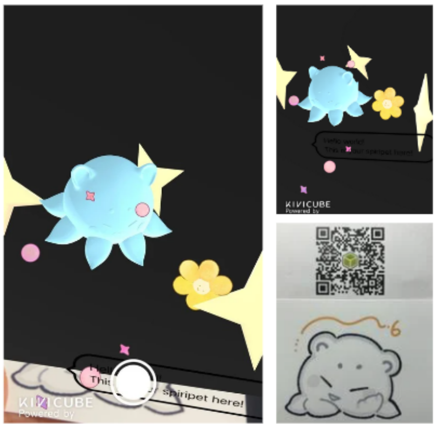


Figure 3.12 AR Game Demo v2.0

As depicted in Figure 3.12, the web-based AR prototype is further developed to incorporate more dynamic interactivity and affective feedback. Likewise, it still adopts the image as a marker for image tracking, but in this iteration, the virtual pet is presented with a continuous floating animation, which includes entrance and exit animation, accompanied by background music (BGM) to enhance its sense of life and establish an ambient mood. A tap-based interaction was implemented as the primary means of engagement. Upon tapping the 3D model on the screen, the interface provides immediate feedback:

- The pet's facial expression changes.
- Accompanying decorative elements animate and vary.

- A text bubble appears with a pre-defined message to acknowledge the interaction.

Workshop Design

The workshop was conducted at the Keio University Graduate School of Media Design (KMD) Forum in Takeshiba. 14 students from both within and outside of KMD participated in the workshop. The primary objective of the workshop was to assess the initial design concept, gather feedback on key aspects of the user experience, and collect desired feature envisions for the prototype. Specifically, to varify the interestingness, necessity, innovation of this virtual pet prototype, user desire for such a companion and their definition of its intelligence and identity, effectiveness of current display methods and interaction pattern - holographic device and AR mobile, expected role of the virtual pet in human-pet relationship, user interest in envisioned future features, especially the application of LLMs.

Each participant was guided through a structured playtesting session which proceeded as follows:

1. A brief introduction to the concept of a virtual pet companion.
2. Interact with the holographic desktop device prototype.
3. Experience the mobile AR interactive prototype.

Figure 3.13 Workshop in KMD Forum

Reflection

The workshop results indicate a strong positive reception, with most participants finding the virtual pet prototype highly interesting, innovative. Most participants agreed on the innovation of applying LLMs in virtual pet, while the evaluation on applying Holographic Projection in desktop device varies. Regarding interaction pattern, the most preferred option was a hybrid mode (50%) that combines a desktop device with a mobile app. In defining human-pet relationship, participants favored the identity as a friend or a family member (peer/younger) the most. The most anticipated features: LLM-driven chatting and real-time human/environment detection and interaction.

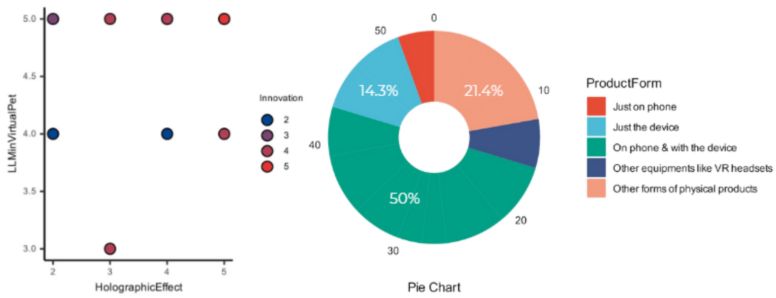


Figure 3.14 Innovation of LLMs and Holography, Preferred Interaction Pattern

3.4.3 Naked-eye 3D Screen Device Prototype

Based on workshop feedback, the final prototype iteration developed in hardware layer, to satisfy users' functional expectation, the naked-eye 3D screen display is identified as a more suitable interface for intuitive LLM-powered conversational interaction and easer integration of multiple user inputs. Therefore, while retaining the overall "device + mobile" interaction pattern, the physical device was replaced with a desktop sized Shinjuku 3D billboard - Companion 01[6]. In the software Layer, the 3D model of the virtual pet was sourced from public assets

6 a glass-free 3D LED screen display from CubeVi Limited: `https://www.openstageai.com/`

available on the Spline[7] platform. Custom animations for this model were created in Blender and subsequently integrated into the main Spline environment. While in Generative AI Application layer,the pet's personality, storytelling, and voice tone were pre-configured through prompt engineering and model fine-tuning first. Second, An API key for the specific GPT-4o model was generated via the OpenAI Platform[8].Then the integration was achieved using Spline's provided interface to OpenAI, where an "AI Assistant Trigger" was configured to link the 3D pet model with the API. Finally, user-initiated voice input was set as the specific action to activate the real-time voice chat with the virtual pet.

Prototyping

Figure 3.15 Desktop Naked-eye 3D Screen Device

As shown in Figure 3.15, in terms of setup of the desktop screen device, it is assembled using a stand and a magnetic ring piece to mount the Companion

7 A 3D design platform:
`https://spline.design/`

8 The platform for building AI products from OpenAI:
`https://platform.openai.com`

01 naked-eye 3D screen. An external web camera with a built-in microphone is attached to the top of the display. This component is responsible for capturing the user's gestures, voice, and facial expressions for real-time multi-modal interaction.

In terms of the realization of the web-based system, it is developed using JavaScript, HTML, and CSS, with the interactive Spline scene embedded via a <spline-viewer> component. We've created some fun events for system-level interactions such as mouse-tracking, gaze-tracking, and drag-and-drop functionality within Spline. Additionally, MediaPipe [31] and face-api.js [27] libraries are integrated for gesture and facial expression recognition in HTML file.

The interactive features in user jouney is illustrated in Table 3.2.

Table 3.2 Summary of Interactive Features

Interaction Type	**Implemented Features**
Mouse/Gesture Interaction	• Gaze and mouse-following movement • Zoom in/out control • Drag-and-drop action
AI Voice Chat	• Pre-set personality and storytelling ability • Customization of the pet's voice tone Evolving capability with user input

User Study Design

The user study for the final prototype consists of two parts. This section details the first part: a workshop designed to evaluate users' non-growth, short-term interactions with the virtual pet. The workshop was conducted during the Play Exhibition at KMD. A total of 25 participants, including students and working people mainly aged 20-30 from both within and outside of the KMD, took part in the study. Each participant was guided through the following procedure:

1. A brief introduction to the virtual pet companion concept.
2. Experience the mouse and gesture-based interactions.
3. Engage in the LLM-powered voice chat interaction.
4. Complete a questionnaire utilizing a 5-point Likert scale.
5. Optional user interview to discuss open questions.

Figure 3.16 Workshop in Play Exhibition

Reflection

Overall, participant feedback on the final prototype was highly positive, with high ratings for its perceived novelty, interestingness, and necessity. The growable LLM-powered voice dialogue is considered the most highly-rated feature. Satisfaction with the current interaction pattern and display improved substantially compared to the previous iteration. While scores on the SDT [11] scale were generally positive, but showed a wide distribution, which is in need of comparison with the long-term interaction group in the whole user study. This strongly indicates the necessity of conducting a controlled user study involving long-term interaction to validate the effectiveness of LLM-driven Co-growth Mechanism. In open discussion, many participants expressed a strong desire for customization options of voice, and appearance of the pet.

Chapter 4
Proof of Concept

4.1. Experimental Setup

The study evaluates the efficacy of this approach by analyzing user interaction and emotional attachment across both short-term and long-term engagements. By analyzing quantitative feedback on self-reported survey measuring Emotional Bond (EAS [32]) and Mental Well-being (SDT [11]), and qualitative feedback such as the evolution of dialogue, the customization of interactions. The experiment consists of 2 parts:

- short-time contact group (non-growth)
- long-term interaction group (co-growth)

4.2. User Tests

4.2.1 Short-time interaction

Participants (aged in approximately 20-30s) experience first approach with the proposed web-based virtual pet companion on a XR desktop device, each interacted through LLM-powered voice dialogs, mouse, and gestures for about 5-10 minutes. Each participant completed a 5-point Likert scale User Survey on measuring Emotional Bond (EAS) and Mental Well-being (SDT). Optional user interview was held voluntarily.

Participants

24 participants within and outside of Graduate School of Media Design at Keio University (15F, 9M), aged 19-31 years old (M=26.00, SD=2.27) living mainly in Tokyo, few living in Germany, Canada, and China.

Table 4.1 Short-term Group Participants Information

No.	Age Group	Gender	Real Pet Experience	Digital Pet Game Experience
1	20s	Female	Y	Y
2	20s	Male	Y	Y
3	20s	Male	Y	N
4	20s	Female	N	Y
5	20s	Male	N	Y
6	20s	Male	Y	Y
7	20s	Female	N	Y
8	20s	Female	Y	Y
9	20s	Male	N	Y
10	10s	Female	N	Y
11	20s	Female	N	Y
12	20s	Male	Y	N
13	30s	Female	Y	Y
14	20s	Female	N	Y
15	20s	Female	Y	Y
16	20s	Female	Y	Y
17	20s	Female	Y	N
18	20s	Female	Y	Y
19	20s	Female	Y	Y
20	20s	Male	N	N
21	20s	Female	Y	Y
22	20s	Female	Y	N
23	20s	Male	N	Y
24	20s	Memale	Y	Y

Procedure

The user test was conducted after receiving approval from KMD Ethics Review Committee. After verbal interpretation, a consent form requiring informational authorization for research purposes was provided. After brief introduction of concept of the virtual pet companion, each participant was allowed to freely explore the mouse interactions, gesture interactions and LLM-powered voice dialog features in 5 to 10 minutes within any desired order. Eventually, the participant was asked to complete the user survey that adopted the EAS and SDT scale. Besides, participants have the right to make any comments or ask any questions at any time throughout the test and voluntarily choose whether or not to answer the open-ended questions in the optional user interview.

Figure 4.1 Questionnaire Filling

4.2.2 First time approach test

To sum up, the first time approach test is a short-term contact with a generic personality and storytelling pre-set virtual pet, which is a real-time multi-modal interaction without growing mechanism. Participants can engage with the prototype either on-site or remotely. On-site participants can interact both via a desktop autostereoscopic (naked-eye) 3D screen display and a standard PC. Remote participants access the prototype on their personal computers through a public URL or a downloadable HTML file and are provided with a demonstration video and concise text-based instructions.

4.2.3 Long-term interaction

Participants (aged in approximately 20-30s) experience relatively long-term in-depth interaction with the proposed web-based virtual pet companion on their personal computers, each interacted through LLM-powered voice dialogs, mouse, and gestures for a 5-day experience with unlimited usage time and frequency each day. They can immersively engage in LLM-driven bidirectional co-growth interaction mechanism. Each participant recorded number of initiated conversations and duration of each interaction session every day. Participants were asked to make flexible daily journal logging. After the whole interaction (5 days), they completed the same 5-point Likert scale User Survey on measuring Emotional Bond (EAS) and Mental Well-being (SDT). Optional user interview was held voluntarily.

Participants

5 core participants randomly selected from previous participants (3F, 2M), aged 24-25 years old (M=24.40, SD=0.55).

Table 4.2 Long-term Group Participants Information

No.	Age Group	Gender	Real Pet Experience	Digital Pet Game Experience
1	20s	Female	N	Y
2	20s	Female	Y	N
3	20s	Female	Y	Y
4	20s	Male	N	N
5	20s	Male	Y	Y

Procedure

The user test was conducted after receiving approval from KMD Ethics Review Committee. After text-based interpretation, a consent form requiring informational authorization for research purposes was provided. After brief introduction of concept of the virtual pet companion, each participant was allowed to freely explore the mouse interactions, gesture interactions and LLM-powered voice dialog features in unlimited time and numbers of experiences within any desired order per day. On a daily basis, each participant recorded number of initiated conversations, duration of each interaction session, and journal log. Eventually, the participant was asked to complete the user survey that adopted the EAS and SDT scale. Besides, participants have the right to make any comments or ask any questions at any time throughout the test and voluntarily choose whether or not to answer the open-ended questions in the optional user interview.

4.2.4 Long-term co-growth test

In long-term co-growth test, participants experience a novel, LLM-driven bidirectional co-growth mechanism. To capture this mutual development facilitating deep emotional bond, participants keep daily logs to focus not only on the evolution of the virtual pet, but also on emotional states variation and personal development of their own.

4.3. Evaluation

Variable Definition

Independent Variables:

- X1 (Type of Interaction Pattern): LLM-driven bidirectional co-growth vs. Baseline Interaction
- X2 (Interaction Duration): Short-term contact (5min) vs. Long-term interaction (5days)

Dependent Variables:

- Y1(Depth of emotional connection): Self-reported emotional attachment
- Y2(Mental well-being): Basic psychological need fulfillment

4.3.1 Quantitive Analysis

To evaluate whether the LLM-driven bidirectional co-growth mechanism combined with long-term interaction resulted in significantly higher psychological outcomes compared to the baseline short-term interaction, we conducted a non-parametric statistical analysis using the Mann–Whitney U test [33].

This method was selected due to the small sample size (n = 5) in the long-term group and the non-normal distribution of scores, which violate the assumptions required for parametric tests such as the independent samples t-test. The Mann–Whitney U test is a rank-based test that compares the distributions of two independent groups and is robust to small sample sizes and outliers.

Specifically, we compared the total scores of two key psychological constructs:

- Emotional Attachment, measured via the Emotional Attachment Scale (EAS), composed of five items: Delighted, Connected, Passionate, Friendly, and Affectionate/Bonded.
- Basic Psychological Need Fulfillment, measured via the Self-Determination Theory (SDT) scale, covering three dimensions: Autonomy, Competence, and Relatedness.

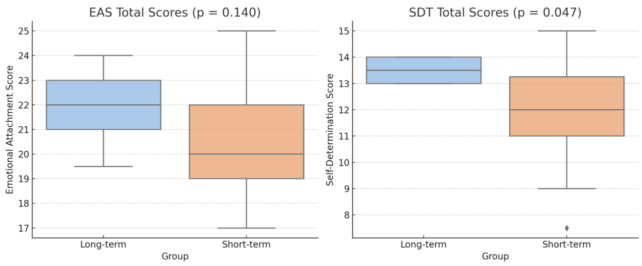
Statistical Basis and p-value Calculation

$$U = n_1 n_2 + \frac{n_1(n_1 + 1)}{2} - R_1$$

For Emotional Attachment (EAS total score), participants in the long-term interaction group (n = 5) reported higher scores ($M = \bar{x}_1,\ SD = s_1$) compared to the short-term group (n = 25), though the difference did not reach statistical significance, U = 79.0, p = 0.148. While this suggests a positive trend, the small sample size in the long-term group may have limited statistical power.

In contrast, the difference in Self-Determination Theory (SDT) total scores was statistically significant. The long-term group demonstrated higher fulfillment of basic psychological needs, as evidenced by a significant difference between the two groups (U = 89.0, p = 0.051), with the long-term group scoring consistently higher across autonomy, competence, and relatedness dimensions.

These results indicate that prolonged and co-evolutionary interaction with an LLM-powered virtual pet may play a meaningful role in enhancing users' mental well-being, particularly in satisfying their psychological needs.



(Boxplots indicate higher median scores in the long-term group. Mann–Whitney U test used for comparison.)

Figure 4.2 Comparison of Emotional Attachment (EAS) and Self-Determination (SDT) total scores between long-term and short-term interaction groups.

As illustrated in Figure 4.2, participants in the long-term LLM-driven co-growth group exhibited significantly higher emotional attachment (EAS) scores compared to the short-term baseline group (p ¡ 0.05). A similar trend was observed in SDT scores, suggesting enhanced fulfillment of basic psychological needs through extended interactive engagement.

Quantitative Analysis on Log Data

To complement the questionnaire-based outcomes, a five-day quantitative analysis of user logs was conducted to investigate behavioral engagement patterns and the thematic evolution of journal entries, inspired by User Engagement Scale [34].

Regarding Interaction Frequency and Duration, as shown in Figure 4.3, users initiated on average 1–3 conversations per day, with mean conversation counts fluctuating slightly over time. Notably, while the frequency peaked on Day 1 and slightly declined midweek, the average duration of interaction sessions gradually increased from approximately 8.1 minutes on Day 1 to a peak of 12.1 minutes on Day 3, before stabilizing at around 11 minutes. This suggests a growing depth or comfort in engagement despite steady conversation frequency—potentially indicating increased cognitive-emotional involvement over time.

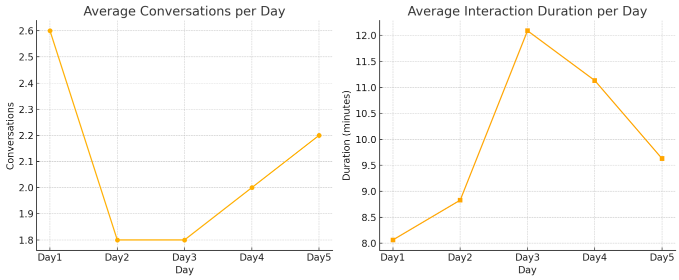


Figure 4.3 Average Interaction Duration Per Day

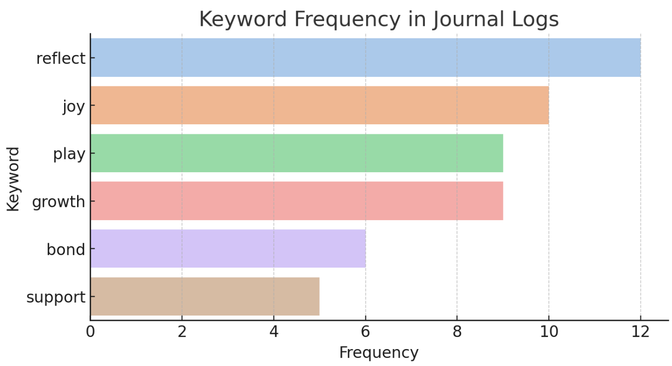


Figure 4.4 Keyword Frequency In Journal Logs

Speaking of Thematic Frequency and Emotional Expression, journal logs were further analyzed by extracting six representative emotional or interactional keywords: support, joy, growth, bond, reflect, and play. As illustrated in Figure 4.4, the most frequently mentioned themes were "reflect" (n=12), "joy" (n=10), and "growth" (n=9), highlighting that participants not only perceived emotional uplift but also engaged in introspective thought and self-development through their interactions with the virtual pet.

4.3.2 Qualitive Analysis

In addition to quantitative measures, we conducted a qualitative analysis of daily journal entries maintained by participants in the long-term interaction group. These logs, recorded over a period of five consecutive days, provided rich insights into the evolving user–agent relationship and the perceived psychological impact of the LLM-driven virtual pet companion.

Each entry typically included open-ended reflections on the content and nature of daily interactions. The logs frequently touched upon key themes such as:

the perceived personality and evolving traits of the virtual pet (e.g., becoming “more empathetic,” “playful,” or “supportive”),

emotional shifts experienced by the users (e.g., “felt heard,” “less lonely today,” “more motivated after conversation”),

and observations related to their own self-development (e.g., “gained new perspectives,” “reflected on past experiences,” “felt encouraged to express myself”).

We employed a thematic analysis approach, using open coding followed by axial grouping to identify recurring motifs across users and time. Three high-level themes emerged: (1) Emergence of Pet Personality, (2) Emotional Resonance and Regulation, and (3) Perceived Personal Growth.

Notably, many participants described a progressive deepening of the relationship, reflected in comments such as “it feels like the pet understands me better each day,” or “it’s more than just a chatbot—it reacts like a companion.” Some users also explicitly referenced the bidirectional co-growth mechanism, noting how their own disclosures appeared to shape the pet’s responses.

This qualitative evidence supports the hypothesis that long-term, LLM-powered interaction can foster a dynamic, emotionally supportive relationship, with perceived mutual development—a key design goal of the virtual pet system.

Table 4.3 An example of daily log of a user

Theme	Illustrative Quote
Emergence of Pet Personality	“It acts more like a curious child now.”
Emotional Resonance	“I felt comforted after venting about my stress.”
Personal Growth	“Talking to it helped me see things more clearly.”

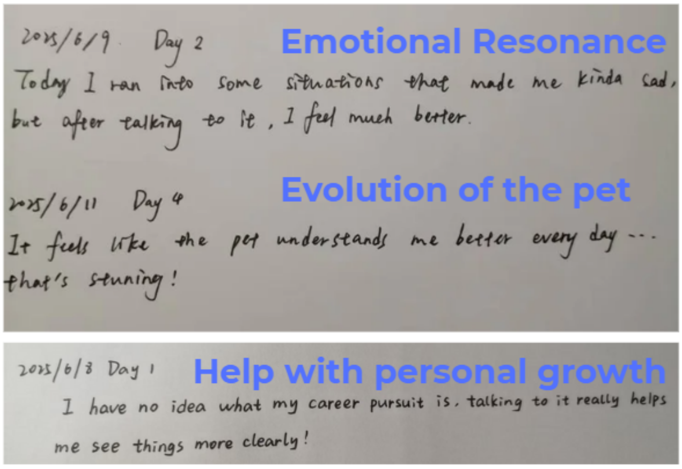


Figure 4.5 Picture of handwritten daily log

The following section summarizes user interviews with two core target users who live in the condensed city from the long-term interaction group. Both participants are in their 20s: one is a university student, and the other is a recent graduate in their first year of employment.

4.3.3 Participant 1

Participant 1 is a university student under 25 years old living in Tokyo. As an international student who lives in the bachelor apartment with pet keeping restrictions, he is the representative of the core target audience for whom traditional pet ownership is impractical.

The feedback from Participant 1 strongly indicates the effectiveness of the LLM-driven bidirectional co-growth mechanism in fostering a deep emotional connection. The participant noted that the most compelling aspect of the experience was the pet's ability to "remember" and "evolve" based on their conversations. For example, the participant shared:

I'm a morning runner, recently I've been telling the pet when I want to close the conversation: "I'm off for my run, bye!". And the other day, I shared a frustrating job hunting experience with him that even though I've tried so many rounds but in the end I still failed. Instead of just giving me some cliché like, "Don't give up" or "You can do it!", instead, he reminded me of how I run every morning and how the sun always rises. He told me that just like that, my perseverance will eventually bring light too. That just... it really hit me. It felt so much more personal and insightful than any generic encouragement would have.

This sentiment directly validates the core hypothesis. The bond was not formed through simple, repetitive interactions but through the creation of a shared history. The pet's ability to recall context and reflect the user's personal narrative back to them was the key factor in making the relationship feel genuine and meaningful.

He also praised for its accessibility when outside home due to web-based carrier, but also shared his concerns for data privacy. For future improvement, he hopes that the facial recognition is not only for sensing the emotions and passively giving empathetic responses, but also for triggering proactive start of interaction.

4.3.4 Participant 2

Participant 2 is a recent graduate who is first year of full-time employment in her 20s. She worked in finance industry, living a structured, often demanding, work schedule in Tokyo. She lives alone, desiring pet companionship but whose lifestyle presents significant barriers to traditional pet ownership.

She shared that sometimes, after a long and tiring day, she doesn't have the energy to talk. However, she found it very healing just to watch the pet move around and make sounds on its own, or to listen to it share its adventure stories.

She was particularly positive about the pet's lifelike autonomy, highlighting its ability to proactively initiate and conclude conversations, as well as the cute idle animations when left alone. Meanwhile, she doubted the necessity of the virtual pet being able to 'speak', which the concern is also raised in AI robotic pets as well.

4.4. Finding

Interaction depth increased over time proved by long-term session, even when conversation frequency was stable. The LLM-driven co-growth mechanism was proved highly effective, particularly within the long-term group. Participants reported that the pet's personalized growth and increasingly empathetic, context-aware responses were central to forming a deep emotional bond.

The pet's ability to proactively initiate and conclude conversations, along with its evolving storytelling brought surprising experiences, and made it feel more like an autonomous, lifelike virtual pet companion. Feedback on the system's core features—including the multi-modal interaction, the embodied XR experience, and the accessible web system—was positive across both groups.

A notable difference emerged in the perception of hardware of final prototype. The short-term group expressed a greater degree of interest in the novelty of the Naked-eye 3D XR device. In contrast, while the long-term group formed deeper connections, their feedback suggested a potential for interaction fatigue of immersive XR experience, indicating the novelty of the display is most impactful during initial encounters.

There are concerns for data privacy when sharing more personal information in the long-term group, and necessity of AI-powered conversational competence of virtual pet, highlighting the shortage in current approach and directing the future improvements.

Chapter 5
Conclusion

5.1. Conclusion

This research set out to investigate the design and impact of an AI-driven virtual pet companion capable of fostering deep emotional connections and supporting the mental well-being of urban dwellers. Grounded in Self-Determination Theory, the central contribution of this thesis was the design and evaluation of a novel LLM-driven co-growth mechanism. The results from our comparative user study confirm the primary hypothesis: qualitative analysis in feedback from core target users in long-term group shared their vivid experiences and perspectives in interviews and daily logs, and supportively quantitative analysis showed that long-term interaction featuring this co-growth mechanism led to a statistically significant improvement in both self-reported Emotional Bond (EAS) and Mental Well-being (SDT) scores when compared to the short-term contact group. This suggests that a co-evolving, AI-driven companion can be an effective paradigm for digital companionship.

5.1.1 Addressing RQ 1 - How can the LLM-powered virtual pet companion foster a deep emotional bond with the user through a reciprocal growth?

This study answers this question by demonstrating that the growth mechanism is a successful strategy for fostering a deep emotional bond. The mechanism establishes a symbiotic relationship where the user's input and self-disclosure directly shape the pet's evolving personality and narrative. As noted in the qualitative feedback from the long-term group, seeing their own experiences reflected in the pet's growth was key to transforming the dynamic from simple interaction, pre-set surface-level gaming into a meaningful partnership, thereby strengthening the emotional connection.

5.1.2 Addressing RQ 2 - How to leverage generative AI and extended reality (XR) to bring an embodied relationship and emotionally resonant experience?

This research demonstrates how Generative AI and XR can be synthesized to create an emotionally resonant experience. Generative AI (LLM) was instrumental in moving beyond pre-scripted dialogue to create a companion with a unique, developing personality capable of real-time conversation. Extended reality (XR), through both the desktop holographic display and the mobile AR component, provided a persistent, embodied presence for the pet in the user's life. The novel integration of these technologies created an accessible and deeply interactive virtual pet companion that participants could connect with anytime, anywhere without environmental constraints.

5.2. Limitations

Despite the positive results, this research has several limitations that should be acknowledged. First, regarding the evaluation, while the long-term group showed higher average scores on the Emotional Attachment Scale (EAS), the difference between the two groups did not reach statistical significance. This is likely attributed to the small sample size of the long-term group, which may have limited the study's statistical power to detect a significant effect. Second, a technical limitation of the prototype for its interaction feature is that, if the virtual pet could evolve from pre-set emotional expressions and animations to real-time AI generated emotional movements and expressions would be more vivid. Last, current displaying form is limited in integration for inputs like vibrate sensors that would have been more fun.

5.3. Future Discussion

The most compelling avenue for future work is to expand the growth span of the virtual pet companion. The current prototype demonstrates growth from a baseline personality to a more personalized character based on user input. A future version could feature a more extensive developmental arc, starting from a non-verbal stage (using only simple, pre-set emotional sounds) and gradually ”learning” to speak and develop a complex personality over a much longer period, which addressed the questioning whether it needs the ability to 'speak'. This would more closely mirror a real-world developmental journey and could potentially foster an even deeper and more profound emotional bond with the user.

References

[1] J. Peen, R. A. Schoevers, A. T. Beekman, and J. Dekker. The current status of urban-rural differences in psychiatric disorders. *Acta Psychiatrica Scandinavica*, 121(2):84–93, feb 2010. doi:10.1111/j.1600-0447.2009.01438.x.

[2] Julianne Holt-Lunstad, Timothy B. Smith, and J. Bradley Layton. Social relationships and mortality risk: a meta-analytic review. *PLoS Medicine*, 7(7):e1000316, jul 2010. doi:10.1371/journal.pmed.1000316.

[3] Helen L. Brooks, Kelly Rushton, Karina Lovell, Penny Bee, Lauren Walker, Louise Grant, and Anne Rogers. The power of support from companion animals for people living with mental health problems: a systematic review and narrative synthesis of the evidence. *BMC Psychiatry*, 18(1):31, feb 2018. doi:10.1186/s12888-018-1613-2.

[4] Jenna W. Applebaum, Leah Loney, Karolina Horecka, and Te'sa M. Graham. Housing-related companion animal relinquishment across 21 animal shelters in the united states from 2019-2023. *Frontiers in Veterinary Science*, 11:1430388, jul 2024. doi:10.3389/fvets.2024.1430388.

[5] Mohammad Javad Koohsari, Ai Yasunaga, Gavin R. McCormack, Tomoki Nakaya, Yuko Nagai, and Koichiro Oka. The design challenges for dog ownership and dog walking in dense urban areas: The case of japan. *Frontiers in Public Health*, 10:904122, apr 2022. doi:10.3389/fpubh.2022.904122.

[6] Michael V. Heinz, Daniel M. Mackin, Brianna M. Trudeau, Sukanya Bhattacharya, Yinzhou Wang, Haley A. Banta, Abi D. Jewett, Abigail J. Salzhauer, Tess Z. Griffin, and Nicholas C. Jacobson. Randomized trial of a generative ai chatbot for mental health treatment. *NEJM AI*, 2(4):AIoa2400802, 2025. URL: https://ai.nejm.org/doi/

full/10.1056/AIoa2400802, arXiv:https://ai.nejm.org/doi/pdf/10.1056/AIoa2400802, doi:10.1056/AIoa2400802.

[7] F. Kaplan. Free creatures : The role of uselessness in the design of artificial pets. 2000. URL: https://api.semanticscholar.org/CorpusID:109414431.

[8] Marita Skjuve, Asbjørn Følstad, Knut Inge Fostervold, and Petter Bae Brandtzaeg. My chatbot companion - a study of human-chatbot relationships. *International Journal of Human-Computer Studies*, 149:102601, 2021. URL: https://www.sciencedirect.com/science/article/pii/S1071581921000197, doi:10.1016/j.ijhcs.2021.102601.

[9] Jixiang Fan, Yusheng Cao, Morva Saaty, Wei-Lu Wang, Lei Xia, Huayi Liu, and D. Scott McCrickard. Understanding the creation of human-virtual entity bonds through the ar mobile game peridot. In *Proceedings of the Extended Abstracts of the CHI Conference on Human Factors in Computing Systems*, CHI EA '25, New York, NY, USA, 2025. Association for Computing Machinery. URL: https://doi.org/10.1145/3706599.3719751, doi:10.1145/3706599.3719751.

[10] Yudi Zhang and Yaning Wang. Emotionally driven virtual pet games: User needs and future trends. *Science and Technology of Engineering, Chemistry and Environmental Protection*, 2024. doi:10.61173/rgxf0e18.

[11] Richard M. Ryan and Edward L. Deci. Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *The American Psychologist*, 55(1):68–78, 2000. doi:10.1037/0003-066x.55.1.68.

[12] Yaniv Kanat-Maymon, Adee Antebi, and Sigal Zilcha-Mano. Basic psychological need fulfillment in human–pet relationships and well-being. *Personality and Individual Differences*, 92:69–73, 2016. URL: https://www.sciencedirect.com/science/article/pii/S0191886915301070, doi:https://doi.org/10.1016/j.paid.2015.12.025.

[13] Froma Walsh. Human-animal bonds i: the relational significance of companion animals. *Family Process*, 48(4):462–480, 2009. doi:10.1111/j.1545-5300.2009.01296.x.

[14] Nicholas Epley, Adam Waytz, and John T. Cacioppo. On seeing human: A three-factor theory of anthropomorphism. *Psychological Review*, 114(4):864–886, 2007. doi:10.1037/0033-295X.114.4.864.

[15] Andrea Beetz, Kerstin Uvnäs-Moberg, Henri Julius, and Kurt Kotrschal. Psychosocial and psychophysiological effects of human-animal interactions: The possible role of oxytocin. *Frontiers in Psychology*, 3:234, 2012. doi:10.3389/fpsyg.2012.00234.

[16] Marta Díaz-Boladeras. Bond formation with pet-robots: An integrative approach. *Current psychology (New Brunswick, N.J.)*, 42(4):2591–2608, 2023.

[17] Adam S. Miner David C. Atkins Ashish Sharma, Inna W. Lin and Tim Althoff. Human–ai collaboration enables more empathic conversations in text-based peer-to-peer mental health support. *Nature Machine Intelligence*, 5:46–57, Jan 2023. doi:https://doi.org/10.1038/s42256-022-00593-2.

[18] Moshe Glickman and Tali Sharot. How human–ai feedback loops alter human perceptual, emotional and social judgements. *Nature Human Behaviour*, 9:345–359, Dec 2024. doi:https://doi.org/10.1038/s41562-024-02077-2.

[19] Zhang Maggie Mengqing Gao Yang and Lysyakov Mikhail. Does social bot help socialize? evidence from a microblogging platform. 2025. URL: https://ssrn.com/abstract=4786302, doi:https://doi.org/10.1287/isre.2024.1089.

[20] Mirjana Prpa, Giovanni Maria Troiano, Matthew Wood, and Yvonne Coady. Challenges and opportunities of llm-based synthetic personae and data in hci. In *Extended Abstracts of the CHI Conference on Human Factors in Computing Systems*, CHI EA '24, New York, NY, USA, 2024. Association for Computing Machinery. URL: https://doi.org/10.1145/3613905.3636293, doi:10.1145/3613905.3636293.

[21] Sebastian Scherr, Bolin Cao, Li Crystal Jiang, and Tetsuro Kobayashi. Explaining the use of ai chatbots as context alignment: Motivations behind the use of ai chatbots across contexts and culture. *Computers in Human Behavior*, 172:108738, 2025. URL: https://www.sciencedirect.com/science/article/pii/S0747563225001852, doi:https://doi.org/10.1016/j.chb.2025.108738.

[22] Selma Šabanović, Corey C. Bennett, Wei-Lin Chang, and Laurel Huber. Paro robot affects diverse interaction modalities in group sensory therapy for older adults with dementia. In *2013 IEEE International Conference on Rehabilitation Robotics (ICORR)*, pages 1–6. IEEE, 2013. URL: https://doi.org/10.1109/ICORR.2013.6650427, doi:10.1109/ICORR.2013.6650427.

[23] Yuhang Zhao, Chao Xue, Xiang Cao, and Yuanchun Shi. Picopet: "real world" digital pet on a handheld projector. In *Proceedings of the 24th Annual ACM Symposium Adjunct on User Interface Software and Technology*, UIST '11 Adjunct, page 1–2, New York, NY, USA, 2011. Association for Computing Machinery. URL: https://doi.org/10.1145/2046396.2046398, doi:10.1145/2046396.2046398.

[24] Kazuki Urata. Research on package design methods that mimic the hatching of living things. Master's thesis, Keio University, Graduate School of Media Design, Yokohama, Japan, 2024. Available via KOARA repository at https://koara.lib.keio.ac.jp/...

[25] Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. Bert: Pre-training of deep bidirectional transformers for language understanding. *arXiv preprint arXiv:1810.04805*, 2018. URL: https://arxiv.org/abs/1810.04805.

[26] Tobias Vogt, Elisabeth André, and Johannes Wagner. Emotion recognition from speech using hidden markov models. In Elisabeth André, Laila Dybkjær, Wolfgang Minker, and Paul Heisterkamp, editors, *Affective Dialogue Systems*, volume 4898 of *Lecture Notes in Computer Science*, pages 71–80. Springer, 2008. URL: https://doi.org/10.1007/978-3-540-85099-1_12, doi:10.1007/978-3-540-85099-1_12.

[27] Vincent Mühler. face-api.js - javascript api for face detection and face recognition in the browser with tensorflow.js. `https://github.com/justadudewhohacks/face-api.js`, 2018. Accessed: June 19, 2025.

[28] Frank Krueger, K. C. Mitchell, Gopikrishna Deshpande, and Jeffrey S. Katz. Human-dog relationships as a working framework for exploring human-robot attachment: a multidisciplinary review. *Animal Cognition*, 2:371–385, mar 2021. `doi:10.1007/s10071-021-01472-w`.

[29] Jian Mou, Yifan Lin, and Fei Zhou. Virtual pets' cuteness matters: A shared reality paradigm for promoting internet helping behaviour. *Technological Forecasting and Social Change*, 2024. `doi:10.1016/j.techfore.2024.123308`.

[30] John Zimmerman, Jodi Forlizzi, and Shelley Evenson. Research through design as a method for interaction design research in hci. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '07, page 493–502, New York, NY, USA, 2007. Association for Computing Machinery. URL: `https://doi.org/10.1145/1240624.1240704`, `doi:10.1145/1240624.1240704`.

[31] Camillo Lugaresi, Jiu Tang, Hadon Nash, Chris McClanahan, Raquel Urtasun, Matthias Lattner, Sava Zlateski, Ewa Reformat, and Ivan Grishchenko. Mediapipe: A framework for building perception pipelines, 2019. URL: `https://arxiv.org/abs/1906.08172`, `arXiv:1906.08172`.

[32] M. Thomson, D. J. MacInnis, and C. W. Park. The ties that bind: Measuring the strength of consumers' emotional attachments to brands. *Journal of Consumer Psychology*, 15(1):77–91, 2005. `doi:10.1207/s15327663jcp1501_10`.

[33] Allen Louis Edwards. *Statistical Methods*, chapter Mann–Whitney U test (pp. 103–110), pages 103–110. Holt, Rinehart and Winston, New York, 2nd edition, 1967. Accessed: 2025-06-10, Internet Archive. URL: `https://archive.org/details/statisticalmetho0000edwa`.

[34] Heather L. O'Brien and Elaine G. Toms. The development and evaluation of a survey to measure user engagement. *Journal of the American Society*

for Information Science and Technology, 61(1):50–69, January 2010. URL: https://ideas.repec.org/a/bla/jamist/v61y2010i1p50-69.html, doi: 10.1002/asi.21229.

Appendices

A. Web-based Virtual Pet html File

Web-based Virtual Pet html File

```
<!DOCTYPE html>
<html lang="zh">
<head>
    <meta charset="UTF-8">
    <meta name="viewport" content="width=device-width, initial-scale=1.0">
    <title>Spline View & Gesture sensing</title>
    <style>
        body {
            font-family: sans-serif;
            margin: 0;
            overflow: hidden;
            background-color: #333;
        }

        /* Spline Viewer */
        spline-viewer {
            width: 100vw;
            height: 100vh;
            display: block;
            position: absolute;
            top: 0;
            left: 0;
            z-index: 1;
        }
```

```
/* Controlled buttons */
#controls {
    position: fixed;
    top: 20px;
    left: 50%;
    transform: translateX(-50%); /* horizontal centering */
    z-index: 100;
    background-color: rgba(255, 255, 255, 0.8);
    padding: 10px 15px;
    border-radius: 8px;
    box-shadow: 0 2px 10px rgba(0,0,0,0.2);
}
#controls button {
    padding: 8px 15px;
    font-size: 1em;
}
#loadingMessage {
    display: inline-block; /* in line with button */
    margin-left: 10px;
    font-size: 0.9em;
}

/* Gesture Recognition (Canvas) */
#output_canvas {
    position: fixed;
    bottom: 20px;
    left: 20px;
    width: 320px;
    height: 240px;
    border: 3px solid #007bff;
    border-radius: 8px;
    z-index: 100; /* most upper layer */
    background-color: #000;
    box-shadow: 0 4px 12px rgba(0,0,0,0.3);
}
#handStatus { /* Gesture status */
```

```
            position: fixed;
            bottom: 20px;
            left: 350px;
            color: white;
            background-color: rgba(0,0,0,0.5);
            padding: 5px 8px;
            border-radius: 4px;
            font-size: 0.8em;
            z-index: 100;
        }


        /* Hidden video element */
        #input_video {
            display: none;
        }

        #splineStatus { /* Spline status */
             display: none; /* hidden */
        }

    </style>
    <script type="module" src="https://unpkg.com/@splinetool/viewer@1.9.98/build/spline

    <script src="https://cdn.jsdelivr.net/npm/@mediapipe/hands/hands.js" crossorigin="a
    <script src="https://cdn.jsdelivr.net/npm/@mediapipe/camera_utils/camera_utils.js"
    <script src="https://cdn.jsdelivr.net/npm/@mediapipe/drawing_utils/drawing_utils.js
</head>
<body>
    <spline-viewer id="splineScene" url="https://prod.spline.design/1s9G58mA9Ubaawca/sc
    <div id="controls">
        <button id="startButton">Start Gesture Sensing</button>
        <span id="loadingMessage" style="display: none;">loading the model...</span>
    </div>

    <canvas id="output_canvas"></canvas>
```

```
<p id="handStatus">click the button to start</p>

<video id="input_video"></video>
<p id="splineStatus">Spline view is loading...</p>


<script type="module">
 // --- Spline Setup ---
    const splineViewer = document.getElementById('splineScene');
    const splineStatus = document.getElementById('splineStatus'); // optional statu
    let splineApp = null;

    splineViewer.addEventListener('load', () => {
        splineApp = splineViewer.spline;
        console.log('Spline scene loaded.');
        if(splineStatus) splineStatus.textContent = 'Spline is loaded';
        // stop the interaction with gesture for now
    });
    splineViewer.addEventListener('error', (e) => {
        console.error('Spline Viewer error:', e);
        if(splineStatus) splineStatus.textContent = 'Spline loaded failed';
    });


    // --- MediaPipe Hands Setup ---
    const videoElement = document.getElementById('input_video');
    const canvasElement = document.getElementById('output_canvas');
    const canvasCtx = canvasElement.getContext('2d');
    const startButton = document.getElementById('startButton');
    const loadingMessage = document.getElementById('loadingMessage');
    const handStatus = document.getElementById('handStatus');

    // set the canvas size
    const cornerCanvasWidth = 320;
    let cornerCanvasHeight = 240;
```

```
function setupCornerCanvasDimensions(videoWidth, videoHeight) {
    const aspectRatio = videoWidth / videoHeight;
    canvasElement.width = cornerCanvasWidth;
    canvasElement.height = cornerCanvasWidth / aspectRatio;
    cornerCanvasHeight = canvasElement.height; // renew actual height
}
// Initialize a default size
canvasElement.width = cornerCanvasWidth;
canvasElement.height = cornerCanvasHeight;


function onResults(results) {
    canvasCtx.save();
    canvasCtx.clearRect(0, 0, canvasElement.width, canvasElement.height);
    // depict the video in smaller sized canva
    canvasCtx.drawImage(results.image, 0, 0, canvasElement.width, canvasElement

    if (results.multiHandLandmarks && results.multiHandLandmarks.length > 0) {
        handStatus.textContent = \verb|‘${results.multiHandLandmarks.length}han
        for (const landmarks of results.multiHandLandmarks) {
            drawConnectors(canvasCtx, landmarks, HAND_CONNECTIONS, {color: ’#00
            drawLandmarks(canvasCtx, landmarks, {color: ’#FF0000’, lineWidth: 1
        }
    } else {
        handStatus.textContent = ’No hands detected’;
    }
    canvasCtx.restore();
}

const hands = new Hands({locateFile: (file) => {
    return \verb|https://cdn.jsdelivr.net/npm/@mediapipe/hands/${file}|;
}});
hands.setOptions({
    maxNumHands: 1, // detect one hand first
    modelComplexity: 0, //
    minDetectionConfidence: 0.65,
```

```
        minTrackingConfidence: 0.65
    });
    hands.onResults(onResults);

    let camera = null;
    async function startCameraAndHandTracking() {
        startButton.disabled = true;
        startButton.textContent = 'starting...';
        loadingMessage.style.display = 'inline';
        handStatus.textContent = 'Asking for permission...';
    try {
            const constraints = { video: true };
            const stream = await navigator.mediaDevices.getUserMedia(constraints);
            videoElement.srcObject = stream;

            videoElement.onloadedmetadata = () => {
                // set the aspect ratio for the window
                setupCornerCanvasDimensions(videoElement.videoWidth, videoElement.v

                // MediaPipe Camera Initialization
                const camera = new Camera(videoElement, {
                    onFrame: async () => {
                        if (videoElement.readyState >= HTMLMediaElement.HAVE_ENOUGH
                            await hands.send({image: videoElement});
                        }
                    },
                    // Camera: width/height - inner element size
                    width: videoElement.videoWidth,
                    height: videoElement.videoHeight
                });
                camera.start();

                loadingMessage.style.display = 'none';
                startButton.textContent = 'runing';
                handStatus.textContent = 'the camera is on';
            };
```

```
            } catch (error) {
                console.error("can’t turn the camera or MediaPipe on:", error);
                loadingMessage.style.display = ’none’;
                startButton.disabled = false;
                startButton.textContent = ’start gesture sensing’;
                handStatus.textContent = \verb|‘Error: ${error.message}‘|;
            }
        }
        startButton.addEventListener(’click’, startCameraAndHandTracking);
    </script>
<body>
```

B. Consent Form

Edit Form: `https://docs.google.com/forms/d/1yLI1xVUBcDpgUAOW7Ha2ytfm87-gJtb59so9md8TYcM/edit`

Live Form: `https://docs.google.com/forms/d/e/1FAIpQLSelmeLRhzzXWcWQesaVozDdGAXjWz` `viewform`

C. User Survey I

- Do you find this prototype of virtual pet interesting? (1-5)
- Do you think it is necessary to keep a virtual pet like this? (1-5)
- What do you think about the innovation of the Virtual Pet Companion product? (1-5)
- How much do you want to have such a virtual pet? (1-5)
- Are you expecting the pet to be smarter? (even smarter than human) (1-5)
- How do you rate the effects of utilization of Holographic Projection for Virtual Pet product? (1-5)

- How do you rate the effects of applying AR in interaction with the virtual pet? (1-5)
- Do you think it is a good practice to apply LLM (Large Language Model) in virtual pet? (1-5)
- Do you think VR is a better choice for displaying such virtual pet? (1-5)
- Do you think it is a good way to integrate most interactive features into the holographic device rather than mobile apps? (1-5)
- Could you please imagine your preferred product form for the Virtual Pet Companion?
  - Just the device
  - Just on phone
  - Other equipments like VR headsets
  - On phone and with the device
  - Other forms of physical products
  - Other forms of digital products
- What feature do you find most attractive in current prototyping?
- What feature are you mostly envisioning for future product?
- What role do you expect the virtual pet to be served as?
- Which of the upcoming functions or features would you like to see implemented first?

D. User Survey II

- Have you kept pets in your life? (Y/N)
- Have you played any digital pet games? (Y/N)

- Student or work?
- Do you find this prototype of a virtual pet interesting? (1-5)
- Do you think it is necessary to keep a virtual pet like this? (1-5)
- What do you think about the innovation of the Virtual Pet Companion product? (1-5)
- How much do you want to have such a virtual pet? (1-5)
- Do you find the interactions with the virtual pet exciting? (1-5)
- What interactive feature do you think is most exciting/attractive to you?
- How would you rate the Empathy ability of the virtual pet? (1-5)
- How would you rate the LLM-powered growing voice interaction feature of the virtual pet? (1-5)
- How would you rate the unique affiliation of the virtual pet (that brings the sense of personalized intimacy)? (1-5)
- How would you rate the effect of the display of this prototype? (1-5)
- According to the SDT theory, how true do you feel that the virtual pet meets the needs for autonomy (e.g. "When I am with my pet, I am free to be who I am")? (1-5)
- According to the SDT theory, how true do you feel that the virtual pet meets the needs for competence (e.g. "When I am with my pet, I feel like a competent person")? (1-5)
- According to the SDT theory, how true do you feel that the virtual pet meets the needs for relatedness (e.g. "When I am with my pet, I feel loved and cared for")? (1-5)
- What topics would you like to talk about with the virtual pet?

E. Optional User Interview Questions

- What is your ideal virtual pet like?
- What long-term interactive features would you like a future growable virtual pet to have?
- What are they most and least satisfied with during the experience?
- Compared to real pets, what do you think are the key strengths and weaknesses of this product?
- Compared to existing digital pet products, what do you think are the key strengths and weaknesses of this product?
- What is your ideal keeping form of a virtual pet (e.g., physical device, AR/VR, mobile game, etc.)?
- To what extent do you think the LLM-driven bidirectional co-growth interaction mechanism solves the lack of emotional depth in current digital pet products? (Scale: 1 = Not at all, 5 = Completely)
- If you rated the previous question below 3, what suggestions do you have for improvement?
- How would you compare the experience between virtual pets and real pets/digital pet products/games?
- Do you have any other comments or feedback?

F. Ethics