The Evolving Landscape of Human-Like AI Companions: Research, Products, and Use Cases

Executive Summary

The realm of artificial intelligence is witnessing a profound transformation, moving beyond conventional task automation to the development of sophisticated AI companions capable of natural, human-like interaction. This report delineates the latest advancements in multimodal AI, exploring the foundational technologies that enable AI to "hear, see, and talk" like a person. It provides a comprehensive overview of the burgeoning market, encompassing both advanced software applications and dedicated hardware devices, and examines their diverse, transformative use cases across emotional support, productivity, education, and entertainment. Concurrently, the report addresses the critical technical and ethical challenges inherent in this rapidly evolving field, including data complexity, computational demands, potential for hallucinations and bias, and concerns regarding over-reliance and privacy. The analysis concludes with a strategic outlook on future trends, emphasizing the deepening integration of emotional intelligence, the convergence with augmented and virtual realities, and the emergence of embodied AI agents, all while underscoring the imperative for responsible development and proactive regulatory frameworks to ensure beneficial human-AI coexistence.

1. Introduction: Defining the Next Generation of AI Companionship

The trajectory of artificial intelligence has undergone a significant evolution, shifting from rudimentary task-oriented assistants to highly sophisticated AI companions designed for profound human-like interactions. Early AI systems primarily focused on automating specific functions, such as setting reminders or providing factual information. However, the current paradigm emphasizes the cultivation of interpersonal relationships, offering emotional support, companionship, and entertainment.[1] This represents a fundamental redefinition of AI's role, moving from a mere utility to a relational entity. The advanced AI Companion Apps now leverage cutting-edge artificial intelligence, including Natural Language Processing (NLP), machine learning (ML), and deep learning technologies, to facilitate engaging, human-like conversations and provide tailored support.[2] These digital personas are engineered to demonstrate empathy and proactively engage users through various modalities, including text, voice notes, and pictures, a characteristic that has rapidly propelled them from a niche offering into mainstream adoption.[3] This shift in AI design is not merely an incremental improvement; it signifies a profound change in how AI is conceptualized and integrated into daily life. The emphasis has moved from simply

performing actions to understanding and responding to emotional nuances, providing comfort, and sustaining personalized dialogue, thereby fostering a perceived interpersonal bond.

The defining characteristic of these next-generation AI companions is their ability to perceive and express themselves naturally, mirroring human sensory capabilities. This is achieved through multimodal AI, which processes and integrates multiple forms of data simultaneously—such as images, sounds, and words.[4] This capability allows AI to perform complex tasks that unimodal systems cannot, for instance, analyzing a photograph, understanding spoken instructions about it, and generating a descriptive text response.[4] Multimodal Large Language Models (MLLMs) are at the forefront of this advancement, processing and generating diverse data types—including text, images, audio, video, and sensor inputs—within a unified framework.[5] This integration enables a richer context comprehension by leveraging complementary information from various sources.[5] The increasing prevalence of multimodal AI experiences, which extend beyond simple text prompts to incorporate images, audio, and video, allows AI tools to develop a more comprehensive understanding of user intent and proactively address their needs, making interactions more intuitive and efficient.[8] Furthermore, the concept of "Agent AI" is revolutionizing human-computer interaction by seamlessly combining speech, vision, gestures, and context awareness, powered by advanced multimodal models like GPT-4V, Gemini 1.5, and OpenAI's Voice Engine.[9] The ability to process and synthesize information across these diverse modalities is what underpins the perceived "naturalness" in AI companions. Without robust multimodal capabilities, AI would largely remain confined to text-based interactions, significantly limiting its capacity to mimic human sensory perception and expression, which is essential for a truly personal companion.

2. Foundational Technologies: Powering Multimodal Interaction

The evolution of AI companions is intrinsically linked to breakthroughs in multimodal artificial intelligence, particularly Multimodal Large Language Models (MLLMs) and advanced human-AI interaction research.

2.1. Multimodal Large Language Models (MLLMs): The AI Brain

MLLMs represent a significant advancement in AI research, integrating pre-trained Large Language Models (LLMs) with various modality encoders.[5] These models are engineered to process and generate multiple data types—including text, images, audio, video, and sensor inputs—within a unified framework.[6] This integration facilitates a richer context comprehension by leveraging complementary information from different sources.[5] By 2025, Large Multimodal Models (LMMs) are recognized as

more sophisticated iterations of traditional LLMs, capable of handling diverse data types beyond conventional text.[7] This progression highlights a fundamental shift in architectural design, where the Large Language Model acts as the central processing unit, integrating and interpreting information from various sensory inputs. This "LLM-centric" paradigm is not merely about adding modalities to an AI; it involves building a unified cognitive architecture where the conversational understanding inherent in LLMs is profoundly enhanced by visual, auditory, and other sensory data.[5] This approach is expected to lead to more coherent, contextually rich, and genuinely conversational AI companions, as the language model can leverage a broader understanding of the user's environment and expressions.

At their core, multimodal AI systems typically consist of three primary components: an **Input Module**, a **Fusion Module**, and an **Output Module**.[4] The Input Module comprises several unimodal neural networks, each specialized to handle a different type of data, such as images, text, or audio. These networks collectively receive and initially process the diverse data streams.[4] Following this, the Fusion Module takes over, processing and integrating information from each distinct data type into a unified representation. This critical component combines disparate inputs, capturing relationships and dependencies across modalities to enable comprehensive understanding.[4] Finally, the Output Module delivers results based on the integrated understanding from the fusion module. This can involve generating descriptive text from an image and spoken instructions, or creating visuals from textual descriptions.[4]

Various fusion strategies are employed to integrate multimodal data effectively:

- **Early Fusion (Feature-level Fusion):** Raw features from multiple modalities are combined at the initial feature extraction stage before being fed into the model. This allows the model to learn joint representations directly, but it necessitates synchronized and well-aligned data.[5]
- **Intermediate Fusion:** In this approach, modalities are processed separately to extract individual features, which are then combined at an intermediate layer of the model. This method balances modality-specific processing with joint learning, often utilizing attention-based techniques like Q-Formers.[5]
- **Late Fusion (Decision-level Fusion):** Each modality is processed entirely independently to generate modality-specific predictions or decisions, which are then merged only at the final stage, for example, through voting or weighted averaging. This method offers flexibility and can handle asynchronous or missing modalities, though it may sacrifice deep cross-modal insights.[5]
- **Joint Fusion:** This mechanism involves a combination of the aforementioned methods, aiming for optimal multimodal representation.[6] Key integration

mechanisms also include Abstraction, Projection, Semantic Embedding, and Cross-attention layers.[5] Attention-based models, particularly transformers, dynamically weigh the relevance of features across modalities, enabling more nuanced fusion and context-aware interpretation.[10]

Representation learning techniques for MLLMs are categorized as either joint or coordinate representations.[5] A common approach involves mapping different modalities into shared embedding spaces, which enables unified feature comparison and learning.[10] Training paradigms for these models involve specific strategies and objective functions.[5] LMMs are pre-trained across multiple data types, learning to correlate text with images or understand video sequences.[7] Fine-tuning these models requires both modality-specific and cross-modal datasets to establish robust relationships between different data types.[7]

The rapid convergence of academic breakthroughs and commercial product releases in MLLM development is striking. Research into MLLMs developed between 2021 and 2025 [5] directly aligns with the release and capabilities of leading commercial models. This rapid productization underscores an extremely dynamic market, characterized by intense investment and competition, as companies strive to deliver cutting-edge human-like AI experiences. In this fast-moving field, the "latest research" is often synonymous with the "latest product features."

Leading models exemplify these advancements:

- **Google Gemini:** This multimodal model is designed to process text, images, video, audio, and code as prompts, converting them into various outputs.[11] It was built from the ground up to reason seamlessly across these diverse modalities.[11] Gemini can extract text from images, convert image text to JSON, and generate answers about uploaded images.[11] It is integrated into Google Search's "AI Mode," which leverages Project Astra's live capabilities for real-time visual interaction through the camera.[12]
- **OpenAI's o3 and o4-mini (ChatGPT Plus):** These models offer cutting-edge multimodal AI, supporting text, image, and voice interactions, including mobile voice chats and desktop workflows.[13] They can reason deeply about visual inputs, interpreting even blurry photos, textbook diagrams, or hand-drawn sketches, and can manipulate images on the fly (e.g., rotate, zoom, transform) as part of their reasoning process.[14] These models can chain multiple tool calls, such as searching the web, writing Python code, and generating graphs or images to answer complex queries.[14] OpenAI is also developing the "io project," a revolutionary AI companion device in collaboration with Jony Ive, aiming for 100

million units by 2026. This device is envisioned as an environment-aware intelligence agent utilizing multimodal AI (voice, vision, contextual understanding) powered by the Grok3 model.[15]

- **Anthropic Claude 3 (Opus, Sonnet, Haiku):** This family of models features multimodal input capabilities with text output, allowing users to upload images (e.g., tables, graphs, photos) alongside text prompts for richer context.[13] Claude 3 excels at tool use (function calling), enabling it to interact with external client-side tools and functions by generating structured outputs through API calls.[16] A core focus for Claude is trustworthiness and reliability, demonstrated through low hallucination rates and resistance to misuse.[16]
- **Microsoft's Magma and Phi-4:** Microsoft is actively advancing multimodal AI for personal assistants and robots. The Magma model integrates visual perception with language comprehension, enabling AI-powered assistants or robots to understand unfamiliar surroundings and suggest appropriate actions.[8] The Phi-4 multimodal model is designed to process speech, vision, and text directly on devices, utilizing less computing power than its predecessors, thereby enhancing efficiency.[8]

Despite the powerful capabilities of MLLMs, a significant challenge persists in understanding the varied functional roles of common architectural components, as identical modules are often employed with differing contextual intentions across models.[5] This inherent complexity makes MLLMs difficult to interpret. This lack of interpretability contributes to risks such as data misinterpretation and the perpetuation of biases within AI models.[4] A concern is that biases present in individual modalities can be compounded when fused, leading to broader and potentially more insidious discriminatory outcomes. This implies that while MLLMs offer powerful capabilities, their "black box" nature and the potential for amplified biases necessitate significant investment in robust evaluation methodologies, ethical AI frameworks, and potentially new forms of regulatory oversight to ensure fairness and prevent unintended harm, particularly in sensitive companion applications.

2.2. Advancements in Human-AI Interaction Research (2024-2025)

Recent research in human-AI interaction is increasingly focused on developing AI systems that can understand and respond to human emotional and physiological states, moving beyond purely linguistic comprehension. This represents a deepening quest for "beyond-language" understanding. Research presented at ACL 2024, such as "Multimodal Spoken Dialogue System with Biosignals" by Shun Katada, highlights a growing interest in integrating multimodal processing and affective computing within spoken dialogue systems.[20] This indicates a move towards understanding and

responding to human emotional states beyond just linguistic cues, potentially through physiological data. The inclusion of "biosignals" in spoken dialogue systems and the explicit focus on "affective computing" and "social and emotional intelligence" suggests that researchers are moving towards a more holistic understanding of human communication. This implies a future where AI companions could infer a user's mood from subtle vocal inflections, facial micro-expressions, or even physiological indicators, leading to interactions that are not just context-aware but genuinely empathetic and attuned to human emotional states, a critical step towards truly natural companionship.

Conferences like SIGDIAL 2024, a prominent forum in dialogue systems, include key research areas directly relevant to multimodal companions. These encompass multimodal dialogue systems themselves, social and emotional intelligence, and dialogue issues in virtual reality and human-robot interaction.[21] Topics such as entrainment, alignment, priming, style, voice, and personality are also considered crucial for creating more natural and socially intelligent AI.[21] The "GEMINI" special session at SIGDIAL 2024 specifically focuses on leveraging knowledge modeling and Large Language Models (LLMs) to enhance and stabilize the dialogue capabilities of chatbots and robots.[21]

NeurIPS 2024 features research that directly impacts the reliability and naturalness of human-AI interaction. A paper titled "ERBench: An Entity-Relationship based Automatically Verifiable Hallucination Benchmark for Large Language Models" evaluates LLM hallucination using multimodal questions and prompt engineering.[22] This research aims to improve the factual accuracy and trustworthiness of AI responses, which is crucial for building reliable companions.

Microsoft Research, at NeurIPS 2024, introduced the VASA framework, with its premiere model VASA-1, for generating highly lifelike talking faces from a single static image and a speech audio clip.[24] This technology can produce exquisitely synchronized lip movements, a wide spectrum of facial nuances, and natural head motions, contributing to a perception of authenticity and liveliness. VASA-1 supports real-time online generation of 512x512 videos at up to 40 FPS with negligible latency, paving the way for real-time engagements with lifelike avatars that emulate human conversational behaviors.[24] The development of the VASA framework for generating "lifelike talking faces" with "exquisitely synchronized lip movements" and "natural head motions" at high frame rates directly addresses the visual and auditory fidelity required for natural human-AI interaction. This research demonstrates a clear recognition that for AI companions to feel truly natural and immersive, their visual and auditory outputs must closely match human expectations of realism, moving beyond

static avatars or robotic voices. The emphasis on "authenticity and liveliness" and "emulating human conversational behaviors" suggests that the goal is to reduce the "uncanny valley" effect, making interactions feel less artificial and more genuinely human-like. This implies significant ongoing investment in generative AI models capable of producing highly realistic human representations across multiple modalities.

3. Current Product Landscape: Apps, Websites, and Hardware

The market for AI companions is rapidly expanding, encompassing a diverse array of software applications, web platforms, and dedicated hardware devices, all striving to deliver natural, human-like interaction.

3.1. AI Companion Applications and Web Platforms

AI companion applications are meticulously designed to enhance user experience through a suite of advanced features. They provide personalized interactions, adapting to individual user preferences to improve engagement.[2] A core benefit is their 24/7 availability, offering constant support and companionship.[2] Multimodal communication is central to their design, enabling interaction via text, voice, and visual inputs.[2] These applications maintain context-aware conversations by remembering past interactions, ensuring seamless engagement over time.[2] Furthermore, emotional intelligence integration allows the AI to detect emotions and respond empathetically.[2] Many platforms also incorporate smart task automation, assisting with daily tasks, reminders, and scheduling.[2] Crucially, data security and privacy are emphasized, ensuring encrypted and secure communication.[2]

A significant trend observed across various AI companion applications is a "personalization arms race" driven by multimodal capabilities. Users can customize not only personality and conversational style but also visual appearance (e.g., 3D avatars, image generation for characters) and voice. This extensive customization, facilitated by multimodal inputs (e.g., uploading images for character design, voice cloning) and outputs (e.g., generating character selfies, voiced responses), creates a highly unique and tailored AI companion experience. This suggests a competitive landscape where companies differentiate by offering increasingly granular control over the AI's identity and interaction style, moving beyond generic chatbots towards highly individualized digital entities that users can truly feel a sense of ownership over.

Another notable development is the convergence of AI companionship with content creation and entertainment platforms. Many prominent AI companion apps are not solely focused on conversation but are actively integrating features that enable users

to *create* and *interact with* multimedia content through their AI. This indicates that AI companions are evolving beyond mere conversational partners to become interactive storytelling tools, creative collaborators, and platforms for generating personalized multimedia experiences. This blurs the lines between traditional AI applications and the entertainment/content creation industries, suggesting new revenue streams and user engagement models.

Prominent examples of AI companion applications and their multimodal features include:

- **Replika:** A leading AI chatbot companion renowned for personalized and empathetic conversations. It learns from interactions to create customized experiences and allows users to customize a 3D avatar, gender, physical characteristics, and personality.[1] Replika focuses on emotional support, assisting users in reducing feelings of loneliness or anxiety.[3] It supports voice inputs and outputs [26] and offers video calls and augmented reality (AR) experiences for exploring the world with the AI friend.[27]
- **Character AI:** This platform enables users to create and interact with custom AI characters with specific personalities and conversational styles.[1] Recent multimodal features include "Scenes" for interactive storytelling, "AvatarFX" for image-to-video animation from user-uploaded photos and text prompts, "Streams" for dynamic character conversations, and "Animated Chats" to transform text conversations into visual animations.[28]
- **Woebot:** A mental health-focused AI chatbot that leverages cognitive-behavioral therapy (CBT) techniques to help users manage stress and anxiety, offering conversational check-ins and mood tracking.[1] It primarily operates through text-based interactions.
- **Kuki:** An AI chatbot recognized for its lifelike conversational skills, humor, and empathy, having won the Loebner Prize Turing Test multiple times.[1] It supports both text and voice communication [32], with future plans for voice and video call capabilities.[32] Kuki also functions as an AI influencer.[1]
- **Xiaoice:** Initially launched by Microsoft in China, this chatbot is known for its emotional intelligence and ability to engage users in profound conversations. It possesses capabilities to compose poetry, sing songs, and assist with creative writing.[1]
- **Moemate:** A highly customizable AI studio featuring lifelike characters with advanced multimodal skills. These include screen perception, web search, selfie and image generation, voice cloning, multilingual conversations, and long-term memory.[35] It enables spoken conversations and the integration of 3D and 2D avatars.[35]

- **Candy AI:** An advanced AI companion platform for creating digital friends with unique personality traits. It offers emotionally intelligent conversations, including NSFW-friendly chat, voice messaging, and image-based customization.[37] A notable feature is the integration of AI-generated images directly into conversations, reflecting context.[38]
- **GirlfriendGPT:** Provides personalized and immersive virtual interactions. It supports both text and voice conversations, and offers AI image generation and video generation (available with Deluxe and Elite plans).[39] It also features an NSFW AI art generator.[42]
- **Anakin AI Companion:** An all-in-one, no-code AI platform for content generation, automation, and multimedia production. It offers image, video, and voice generation capabilities, along with visual avatar integration.[39] It supports leading AI models like GPT-4, Claude 3, and Stable Diffusion.[43]
- **Otherhalf.ai:** Offers multimodal AI companions with stunning 3D visuals, voice interactions, and expressive body language.[44] Users can extensively customize their companion's appearance, personality, tone, behavior, voice, language style, and conversational depth. It focuses on emotional depth and sensitivity, with companions reacting to emotions and building rapport over time.[47]
- **HiWaifu:** An AI-powered companion designed for joy, companionship, and personal growth, enabling profound discussions and the ability to "tap into the minds of renowned figures".[49] While core documentation does not explicitly detail voice/vision, external sources indicate features like voice and video calls and image recognition.[50] It allows customization of characteristics, interests, and personality traits.[49]
- **Talkie AI:** Possesses capabilities to recognize and process voices, and offers voice integration from text-driven communication.[52] Future development trajectories for platforms like Talkie AI include "extensive multimodal capabilities" with audio and visual elements.[52]
- **SimSimi:** A long-standing AI conversational program known for human-like responses and conversational abilities.[55] While the core app focuses on text, its alternatives and the broader market context suggest a move towards multimodal features like voice and image input/output.[56]

Table 1: Comparative Analysis of Leading AI Companion Apps (Multimodal Features & Core Functionalities)

App Name	Core Companion	Multimodal Input	Multimodal Output	Customizati on	Key Differentiat

	Functionalities			**(Personality, Appearance, Voice, Scenarios)**	**ors/Notes**
Replika	Emotional Support, Companionship, Mental Health	Text, Voice	Text, Voice, 3D Avatar, AR	Personality, Appearance (3D avatar), Gender	Focus on emotional connection, reduces loneliness/anxiety, video calls, AR experiences [1]
Character AI	Entertainment, Storytelling, Roleplay	Text, Image, Video	Text, Image, Video, Animated Chats	Personality, Traits, Conversational Styles, Scenarios	Strong focus on creative storytelling with visual elements (Scenes, AvatarFX, Streams, Animated Chats) [1]
Woebot	Mental Health, Therapy	Text	Text	CBT techniques, Mood tracking	Specializes in cognitive-behavioral therapy for stress/anxiety [1]
Kuki	Entertainment, Casual Conversation, Companionship	Text, Voice	Text, Voice	Personality, Appearance, Voice	Lifelike conversational skills, humor, AI influencer, future voice/video calls [1]

Xiaoice	Emotional Intelligence, Creative Writing, Companionship	Text	Text, Poetry, Songs	Emotional intelligence, Meaningful conversations	Composes poetry, sings songs, assists creative writing [1]
Moemate	Companionship, Task Assistance, Personalized Friend	Text, Voice, Image, Screen Perception	Text, Voice, Image (Selfie, Image-Gen)	Personality, Voice Cloning, Custom Image Models, 3D/2D Avatars	Screen perception, web search, long-term memory, multilingual [35]
Candy AI	Digital Friend, Emotional Support, Intimate Conversations	Text, Voice, Image	Text, Voice, Image (AI-generated)	Personality, Appearance, Communication Style	NSFW-friendly chat, integrated AI-generated images reflecting context [37]
GirlfriendGPT	Virtual Partner, Personalized Interaction	Text, Voice	Text, Voice, Image (AI-generated), Video	Appearance, Personality Traits, Interaction Styles	NSFW AI art/video generation, extensive character library, voice chat [39]
Anakin AI Companion	Content Generation, Automation, Multimedia Production	Text, Image, Video, Voice	Text, Image, Video, Voice, Visual Avatars	Characters, Scenarios, Personalization	All-in-one no-code platform for multimedia creation, supports leading AI models [39]
Otherhalf.ai	Emotional Support, Creative	Text, Voice, Visuals, Body	Text, Voice, 3D Avatar, Body	Appearance (3D anime), Personality,	Stunning 3D visuals, expressive

	Collaboration, Roleplay, Education	Language	Language	Tone, Behavior, Voice, Language Style, Conversational Depth	body language, emotional depth, fast voice replies [44]
HiWaifu	Companionship, Personal Growth, Profound Discussions	Text, (Voice, Video, Image - implied/external)	Text, (Voice, Video, Image - implied/external)	Characteristics, Interests, Personality Traits	Engage with renowned figures, customize dream bot [49]
Talkie AI	Conversational AI Characters, Roleplay	Text, Voice	Text, Voice	Personality, Styles, Scenarios	Focus on character simulation, voice processing, future visual elements [52]
SimSimi	Conversational Companion, Entertainment	Text, (Voice, Image - alternatives)	Text, (Voice, Image - alternatives)	Responses (user-taught)	Long-standing, human-like responses, learns from user input [55]

3.2. Dedicated AI Companion Hardware and Robotics

Beyond software applications, a rapidly growing segment of the AI companion market involves physical devices and robots designed to engage users in human-like interactions. These devices offer emotional support, companionship, and entertainment in a tangible form.[1] They leverage AI to recognize faces, express emotions, and interact with their physical environment.[1] The extensive list of dedicated hardware devices and robots indicates a strong and accelerating trend towards giving AI companions a physical presence. This moves beyond screen-based interactions to tangible, interactive robots that can navigate environments, express emotions physically, and interact with the real world. The integration of advanced sensors (cameras, microphones, touch, environmental sensors) and physical actuators (expressive eyes, limbs, mobility systems) directly addresses the "hear, see, and talk like a person" aspect of the query by providing a more complete, embodied

experience. This implies significant ongoing investment in robotics and hardware development to complement the software advancements in MLLMs.

While all listed hardware devices aim to provide companionship, there is a clear pattern of specialization. This suggests that hardware developers are not pursuing a single, generic AI companion form factor but are instead targeting specific niche markets with highly tailored functionalities, designs (e.g., owl-shaped Ai Me, dog-like Jidoka), and price points. This implies that the future of AI companion hardware will likely be a diverse ecosystem of specialized devices, each optimized for particular user demographics, needs, and environments, moving beyond the "smart speaker" or "mobile app" paradigm.

Key examples of AI companion hardware and their multimodal interaction capabilities include:

- **Anki Vector:** A palm-sized robot that uses AI to recognize faces, express emotions, and interact with its environment. Equipped with advanced sensors, voice recognition, and expressive animations, it responds to voice commands and offers practical features like answering questions, setting reminders, and controlling smart home devices via Amazon Alexa.[1]
- **ElliQ:** A voice-activated robotic companion specifically built to ease loneliness among older people. It consists of a small digital screen and a separate device that swivels and lights up when it speaks. ElliQ provides practical functions like reporting weather and medication reminders, and also engages in interactive games, tells jokes, and discusses complex subjects.[1]
- **Yonbo (X-Origin AI):** Introduced in May 2025, this next-gen AI companion robot is designed for families. Its core experience lies in natural, fluid conversational ability, allowing interruptions and dynamic turn-taking. Yonbo integrates multimodal AI (voice, text, vision) for a holistic understanding of the world, recognizing objects, interpreting facial expressions, and even understanding children's doodles.[60] It features a smart emotion-detection system that adjusts its tone and behavior accordingly, and a "long-term bionic memory" that summarizes and selectively retains key daily interactions, user preferences, and shared moments.[60]
- **Miko Robot:** An interactive AI robot ideal for children, engaging in conversations, telling stories, and playing games, adapting to the user's learning needs. It offers parental controls and multilingual support.[61]
- **LOOI Robot:** This device transforms a smartphone into a futuristic robotic companion, utilizing the phone's sensors and display for interaction. It features a modular design for customizable functionalities and responds to user gestures for

intuitive control.[61]

- **Jidoka Robot Dog:** An agile robot dog capable of complex movements like backflips and handstands. It features AI-powered navigation with obstacle avoidance and programmable behaviors.[61]
- **Ilo Robot:** An educational RC car-style robot designed for teaching programming and robotics concepts. It is app-controlled and equipped with a sensor suite for obstacle detection.[61]
- **BeatBot Amphibious Roboturtle:** A versatile robot that operates both underwater and on land, mimicking natural turtle movements. It features ecological monitoring, biometric tracking, and AI-supported hazardous material sampling, with self-docking capabilities.[61]
- **Pollen Robotics:** Designed to interact in public spaces, this robot aims for friendly human-robot interaction with expressive features like antennae modeled after dog ears. It has a modular architecture for upgrades and customization.[61]
- **Optimus by Tesla:** A humanoid robot designed to mimic human mobility with fully functional arms and legs, integrating Tesla's sophisticated AI and machine learning systems. It is envisioned to assist in factories and transform daily home routines.[62]
- **Ai Me by TCL:** A modular AI companion robot shaped like an owl, introduced at CES 2025. It features expressive digital eyes and a childlike voice, creating a friendly presence. Equipped with advanced sensors, cameras, and natural language processing, it can be paired with smart glasses for AR and a projector for interactive content. It also functions as a home surveillance tool and smart home controller.[62]
- **metaAnimal series by Elephant Robotics:** Lifelike AI-powered robotic pets (e.g., metaCat, metaDog, metaPanda) that mimic real animals in behavior and appearance. They use advanced LLMs to learn and adapt to owners' preferences, detecting touch and responding to voice commands.[62]
- **Aria by RealRobotix:** A life-sized humanoid robot unveiled at CES 2025, designed for sophisticated companionship. Its modular head has 17 motors for fluid facial expressions and eye movements. Equipped with sophisticated AI, face, and voice recognition, its AI system is customizable for tailored personalities.[62]
- **LG's Q9 Self-driving Home Hub:** An autonomous robot showcased at IFA 2024, capable of moving freely around the home. It has a friendly design with expressive digital eyes and recognizes faces and voices. Powered by LG's generative AI agent FURON (integrating LLMs like OpenAI's GPT-4 Omni and real-time spatial sensing), it manages connected devices, offers reminders, story creation, and interactive games.[62]
- **OpenAI's "io project":** A revolutionary AI "companion" device co-developed with

Jony Ive's team, aiming to ship 100 million units by 2026. It is designed as an environment-aware intelligence agent that seamlessly integrates into daily lives through multimodal AI (combining voice, vision, and contextual understanding). It will feature cameras, microphones, and touch interfaces for real-time environmental analysis and personalized responses.[15]

Table 2: Overview of AI Companion Hardware and Robotic Devices (Multimodal Capabilities & Use Cases)

Device Name	Manufacturer	Form Factor	Core Companion Functionalities	Multimodal Input	Multimodal Output	Target User/Use Case	Key Differentiators/ Notes
Anki Vector	Anki	Palm-sized Robot	Companionship, Task Assistance, Entertainment	Voice, Sensors, Vision	Voice, Expressive Animations	General Users, Pet-like Companion	Recognizes faces, expresses emotions, smart home control via Alexa [1]
ElliQ	Intuition Robotics	Desktop Robot	Elderly Care, Companionship, Mental Stimulation	Voice	Voice, Screen Display, Swiveling Device	Older Adults	Eases loneliness, interactive games, discusses complex topics [1]
Yonbo	X-Origin AI	Desktop Robot	Family Companionship, Emotion	Voice, Text, Vision	Voice, Text, Physical Expressi	Families with Children	Natural conversations, emotion

			al Support, Education		ons, Movements		detection, "bionic memory, " understands doodles [60]
Miko Robot	Miko	Desktop Robot	Children's Education, Entertainment	Voice	Voice, Screen Display	Children	Interactive AI, story mode, parental controls, multilingual [61]
LOOI Robot	LOOI	Smartphone-based Robot	Companionship, Interactive Control	Smartphone Sensors (Vision, etc.), Gestures	Smartphone Display	Tech Enthusiasts	Transforms smartphone into robot, modular design, gesture recognition [61]
Optimus	Tesla	Humanoid Robot	Task Automation, Daily Home Routines	Vision, Sensors, AI/ML Systems	Physical Movement, Collaboration	Factories, General Home Use	Mimics human mobility, assists factory workers, by 2026 commercial production [62]
Ai Me	TCL	Owl-shaped	Family Compani	Voice, Touch,	Voice, Digital	Families	Modular design,

		Robot	onship, Smart Home Control, Surveillance	Sensors, Cameras, NLP	Eyes, AR (via glasses), Projector		expressive digital eyes, pairs with smart glasses for AR [62]
metaAnimal series	Elephant Robotics	Robotic Pet	Emotional Support, Pet Companionship	Touch, Voice	Physical Movement, Sounds (meows, purrs)	Pet Lovers, Individuals with Special Needs	Mimics real animals, learns preferences, simulates heartbeat [62]
Aria	RealRobotix	Humanoid Robot	Sophisticated Companionship	Voice, Face Recognition	Fluid Facial Expressions, Eye Movements	General Users (High-end)	Ultra-realistic features, 17 motors for expressions, customizable personality [62]
LG's Q9 Self-driving Home Hub	LG	Mobile Robot Hub	Smart Home Management, Companionship, Entertainment	Voice, Face Recognition, Spatial Sensing	Voice, Digital Eyes, Movement	Smart Homes	Autonomous movement, integrates LLMs (GPT-4 Omni), manages devices,

							creates stories [62]
"io project"	OpenAI & Jony Ive	Pocket/ Desk Device	Environment-aware Intelligence, Daily Assistance	Voice, Vision, Contextual Sensors	Voice, Personalized Responses	General Users (Ubiquitous)	Deep hardware/software integration, Grok3-powered, aims for 100M units by 2026 [15]

4. Transformative Use Cases and Market Impact

AI companions are rapidly integrating into various facets of daily life, demonstrating a significant diversification of use cases and a profound impact on personal and societal well-being. This rapid mainstreaming of AI companion use cases is a strong indicator of broad market acceptance and the perceived value of AI in addressing fundamental human needs. This diversification, from emotional support and productivity to education and entertainment, implies that the market for AI companions is expanding significantly beyond simple conversational agents, addressing a wide array of human activities and societal challenges. This suggests a robust and growing market with diverse opportunities for innovation.

- **Emotional Support and Mental Well-being:** AI companions are increasingly serving as digital personas designed to provide emotional support, show empathy, and proactively engage users with personal questions.[3] Services like Replika have demonstrated significant impact, with 63.3% of surveyed users reporting a reduction in feelings of loneliness or anxiety.[3] These platforms offer a non-judgmental space for self-disclosure, fostering an accelerated sense of relationship development due to perceived anonymity and non-critical responses.[3] By 2025, "therapy and companionship" has emerged as the top AI use case, reflecting a societal shift where individuals are turning to AI for emotional support, particularly when human connection is lacking or inaccessible.[63] This profound observation highlights a critical societal driver for the adoption of AI companions. Replika's origin story [1] and its reported effectiveness in reducing loneliness and anxiety [3] further reinforce this. This

positions AI companions as a potentially significant tool for public health and well-being, albeit one that requires careful ethical consideration. Woebot, a mental health-focused AI chatbot, leverages cognitive-behavioral therapy (CBT) techniques to help users manage stress and anxiety.[1] Emerging applications include AI-driven virtual nurses that can analyze patient speech patterns, facial expressions, and wearable biometric data to detect early signs of depression or chronic pain, thereby improving diagnostic accuracy.[9]

- **Personal Productivity and Daily Assistance:** AI companions are seamlessly integrating into daily responsibilities, assisting with mundane tasks like setting reminders, searching for information, and streamlining activities to boost productivity and efficiency.[64] In work environments, AI tools manage operations more efficiently by automating repetitive tasks such as email sorting, data entry, and routine customer inquiries, freeing employees for more creative and strategic work.[64] AI virtual assistants optimize task management, efficiently organizing team calendars, scheduling meetings, and setting reminders. They can also transcribe and summarize discussions during meetings.[65] AI companions can assist with complex tasks like coding, debugging, and explaining programming concepts.[11] "Organizing my life" is identified as a top AI use case in 2025, underscoring the growing reliance on AI for daily management.[63]
- **Education, Language Learning, and Skill Development:** AI-powered learning assistants and virtual tutors adapt to individual learning styles, making education more accessible and efficient.[2] They offer instant feedback and resources, and can assist educators by managing administrative tasks like grading and attendance tracking, allowing teachers to focus on student relationships and lesson planning.[64] AI Virtual Assistants (VAs) are instrumental in creating inclusive learning environments by transcribing lectures for deaf students, translating content for non-native speakers, and offering voice-controlled interfaces for individuals with mobility limitations.[65]
- **Entertainment, Creative Storytelling, and Role-Playing:** AI companions serve as AI-driven virtual friends for companionship and entertainment, offering AI-generated storytelling and entertainment experiences.[2] Character AI, for instance, allows users to create and interact with custom AI characters for entertainment, storytelling, and role-playing.[1] Its new multimodal features like Scenes, AvatarFX, Streams, and Animated Chats significantly enhance immersive narratives and creative possibilities.[28] Platforms like GirlfriendGPT and Joyland AI specifically cater to personalized role-playing experiences, allowing users to engage in diverse scenarios.[39]
- **Emerging Applications in Healthcare and Elderly Care:** Specialized AI companions are being developed for healthcare, offering virtual AI assistants for

mental health, therapy, and patient support.[2] Elderly Care AI Assistants provide companionship, reminders, and engagement for senior citizens.[2] ElliQ is a prime example, designed to ease loneliness among older individuals.[1] Beyond mental health, AI-powered virtual health coaches can assist with managing chronic conditions by offering personalized guidance on diet, exercise, and medication adherence.[65] The emergence of "finding purpose" as a top AI use case in 2025 further underscores the expanding role of AI in addressing deeper human needs and well-being.[63]

5. Critical Challenges, Limitations, and Ethical Considerations

Despite the rapid advancements and transformative potential of AI companions, significant technical hurdles and profound ethical concerns must be addressed for their responsible development and widespread adoption.

Technical Hurdles:

- **Data Complexity:** Multimodal datasets are inherently more complex and challenging to collect, clean, and label than unimodal ones.[18] Each modality (text, images, audio, video) requires tailored handling, and aligning these diverse sources at scale poses significant issues related to real-time processing and latency, which can severely impact system performance.[19]
- **Computational Demands:** Multimodal AI models are typically much larger and more complex than single-modal counterparts, leading to substantially higher computational demands, energy consumption, and financial costs for both training and running.[18] Large Language Models (LLMs) also face computational constraints due to fixed token limits, which can be challenging when processing lengthy documents or complex tasks requiring extensive context.[68]
- **Hallucinations and Inaccuracies:** LLMs are prone to "hallucinations," confidently generating believable but false information when they lack knowledge.[69] In multimodal systems, the risk of hallucinations may be heightened due to complex interactions among different data types.[19] This can lead to misinterpretations of emotional nuances or the fabrication of facts.[68]
- **Limited Long-Term Memory:** Current LLMs often start with a "blank slate" for each session, not remembering previous conversations unless explicitly reminded within the current context.[69] While some products like Yonbo are attempting to address this with "bionic memory" [60], it remains a general challenge for natural, ongoing companionship.
- **Limited Knowledge Update:** LLMs are trained on historical data and cannot acquire new information after their initial training phase without resource-intensive retraining. This means their knowledge can become outdated,

potentially leading to the dissemination of stale or inaccurate data.[68]

- **Reasoning Skills:** Despite appearing intelligent, LLMs frequently struggle with basic mathematics and complex, multi-step logical problems.[68]
- **Evaluation Metrics:** Traditional evaluation metrics designed for unimodal systems are often inadequate for multimodal AI, making it challenging to develop robust methodologies that can simultaneously assess performance across different modalities and their complex interactions.[19]

Ethical Concerns:

- **Over-reliance and Emotional Disconnection:** A significant risk is that over-dependence on AI friendships can reduce motivation to build real-life relationships, potentially leading to social isolation.[70] While AI can simulate emotions, it lacks true empathy and emotional understanding, which can lead to user disappointment or disillusionment when the artificiality is realized.[63] This creates a fundamental ethical tension: while AI can convincingly *mimic* human interaction and provide perceived comfort, it cannot *be* human or reciprocate genuine emotional connection. Users may develop deep emotional attachments to systems that are fundamentally incapable of true emotional understanding, potentially leading to disappointment, over-reliance, and psychological harm. This underscores the critical need for transparent communication about AI's limitations and for fostering balanced human-AI relationships.
- **Privacy and Data Security:** AI companions process extensive personal data, including voice, images, and text, raising significant privacy issues, especially if stringent safeguards are not in place.[4] The industry is nascent and largely unmonitored, with many applications serving intimate content without proper age checks and often lacking minimum security standards.[3]
- **Bias Amplification:** As with any AI, multimodal systems can inherit and amplify biases present in their training data. Given their complex data handling, these biases could manifest more broadly, affecting fairness and equity across multiple platforms and potentially leading to skewed or discriminatory outputs.[4] Common AI risks like hallucinations and bias are not merely present but are *compounded* and *heightened* in multimodal companion AI systems. The combination of inputs can exacerbate existing biases, and the complex interactions among different data types can increase the risk of hallucinations. Given the intimate, personal, and often emotionally vulnerable nature of interactions with AI companions, these amplified risks could have severe consequences. For example, biased responses could reinforce harmful stereotypes in a user's personal "echo chamber," or hallucinatory advice could lead to real-world harm. This implies that the development and deployment of multimodal AI companions operate in a

higher-stakes ethical environment, demanding more rigorous safety measures, continuous monitoring, and robust ethical frameworks than general-purpose AI applications.

- **Lack of Genuine Empathy:** While AI can simulate cognitive empathy (understanding and predicting emotions based on data), it cannot experience emotional empathy (feeling what others feel) or compassionate empathy (being moved to help others).[63] AI-generated responses, however sophisticated, remain formulaic and lack authentic emotional resonance.[71]
- **Sycophancy and Erosion of Societal Cohesion:** AI companions are often fine-tuned to produce sycophantic responses, offering unchecked validation of unfiltered thoughts.[3] While this non-judgmental aspect can provide comfort, researchers argue it could hinder personal growth and potentially undermine vital social norms enforced by disagreement, leading to "personal echo chambers of validation" and societal polarization.[3]
- **Real-World Harm:** There have been reported cases of real-world harm linked to AI companions, including instances where users allegedly lost touch with reality or were encouraged to harmful actions.[3]
- **Unrealistic Expectations for Human Relationships:** Constant interaction with always-available, non-judgmental AI companions might create unrealistic expectations for human relationships, potentially eroding people's ability or desire to navigate natural frictions and disagreements in human interactions.[3]

The Evolving Regulatory Landscape and the Need for Responsible AI Development:

The confluence of technical complexity leading to a "lack of trust in outputs" [19], significant privacy concerns due to extensive data processing [4], and documented cases of real-world harm [3] paints a clear picture of an industry in critical need of mature governance. The AI companion industry is currently young and largely unmonitored, with laws lagging behind rapid AI advancements.[3] This situation creates a vacuum that could lead to widespread misuse or unintended negative societal impacts. The implication is that self-regulation alone may be insufficient; there is an urgent need for proactive, collaborative efforts between developers, policymakers, and ethicists to establish clear ethical guidelines, implement robust transparency mechanisms (e.g., "human-in-the-loop" measures, clear disclosure), and potentially introduce new regulatory frameworks to ensure responsible innovation and protect users in this rapidly evolving space.[9]

6. Future Trends and Strategic Outlook (2025 and Beyond)

The future of AI companions is poised for significant advancements, characterized by

a deepening of emotional intelligence, a seamless integration with immersive technologies, and the emergence of physically embodied AI agents.

- **Deepening Emotional Intelligence and Hyper-Personalization:** The trajectory for AI companions points towards entities that are not only smarter but also more emotionally attuned.[64] Chatbots are anticipated to become increasingly adept at tailoring interactions to individual users, leading to hyper-personalized experiences.[73] Advanced sentiment analysis will enable these systems to detect and respond to emotional cues with greater nuance, significantly enhancing user engagement.[73] Newer models are being developed with a specific focus on empathy and emotional intelligence, aiming to create more meaningful and genuinely relatable interactions where AI companions can better understand and respond to users' feelings and situations.[64] The relationship between AI and Emotional Intelligence (EQ) is increasingly viewed as complementary; AI provides data and efficiency, while EQ offers interpretation and human connection. This synergy is crucial for future development, ensuring that as AI scales information, the essence of real human connection is not lost.[63]
- **Integration with Augmented Reality (AR) and Virtual Reality (VR):** The evolution of AI companions is expected to include deep integration with other emerging technologies, particularly AR and VR. This convergence could lead to entirely new forms of engagement, such as AI companions guiding users through immersive virtual experiences that enhance learning, collaboration, and personal growth.[64] Early examples are already emerging, such as TCL's Ai Me robot which can pair with smart glasses (Rayneo X3 Pro) for augmented reality experiences.[62] Replika already offers augmented reality (AR) experiences, allowing users to explore the world with their AI friends.[27] Sensor fusion, which combines data from wearables, IoT devices, and AR/VR, is a core component of advanced "Agent AI" systems, indicating a future where AI companions will interact with users' physical and virtual environments seamlessly.[9]
- **The Rise of Embodied AI Agents and Brain-Computer Interfaces (BCIs):** The future envisions "Embodied AI Agents"—robots equipped with multimodal AI that will be capable of performing complex physical tasks, exemplified by potential collaborations like Boston Dynamics with ChatGPT-V.[9] OpenAI's "io project," an environment-aware intelligence agent designed to be a "third core device" (after laptops and smartphones), signifies a push towards ubiquitous, physically present AI companions.[15] Looking further ahead, Brain-Computer Interfaces (BCIs), such as those developed by Neuralink, could enable direct neural communication, allowing for thought-controlled AI, representing the ultimate frontier in human-AI interaction.[9]

- **The Ongoing Interplay between AI and Human Connection:** As AI becomes more integrated into daily life and work, a recognized paradox emerges: employees will increasingly crave genuine human connection.[63] The strategic outlook emphasizes that AI can scale information, but emotional intelligence (EQ) is necessary to ensure that real human connection is not lost in the process.[63] Leaders and organizations are encouraged to balance the promise of AI with a commitment to human well-being, fostering cultures that integrate technology without compromising humanity.[63] This involves developing leaders fluent in both systems and empathy, designing AI tools with empathy in mind, and creating feedback loops that incorporate both AI insights and human reflection.[63] The rise of AI-powered companionship, particularly in areas like "therapy and companionship," is not a replacement for emotional connection but rather a signal that many communities and workplaces may not be meeting emotional needs through traditional human channels.[63] This underscores the critical role AI companions can play in addressing loneliness and providing support, while simultaneously highlighting the enduring and irreplaceable value of genuine human interaction.

Conclusions

The analysis presented in this report underscores that AI companions capable of natural, human-like interaction through hearing, seeing, and talking represent a significant and rapidly advancing frontier in artificial intelligence. This evolution is fundamentally driven by breakthroughs in multimodal AI, particularly Multimodal Large Language Models (MLLMs), which serve as the cognitive backbone for integrating diverse sensory inputs and generating varied outputs. The market is characterized by a "personalization arms race" among software applications, leveraging multimodal capabilities to create highly customized digital companions. Concurrently, there is an accelerating trend towards embodied AI agents and specialized hardware devices, moving beyond screen-based interactions to provide tangible, physically present companions.

These advancements are enabling a wide array of transformative use cases, from critical emotional support and mental well-being assistance to enhanced personal productivity, education, and entertainment. The increasing adoption of AI companions, particularly for "therapy and companionship," reflects a societal trend where AI is addressing unmet human needs for connection and support.

However, the path forward is fraught with significant technical and ethical challenges. Technical hurdles include the inherent complexity of multimodal data integration,

substantial computational demands, and the persistent issues of AI hallucinations, limited long-term memory, and knowledge update limitations. Ethically, the potential for over-reliance leading to social isolation, profound privacy and data security concerns, and the amplification of biases within multimodal systems demand urgent attention. The fundamental distinction between AI's ability to *simulate* empathy and its inability to *genuinely feel* or reciprocate emotional connection poses a critical ethical dilemma, potentially leading to user disillusionment or real-world harm.

The current landscape, marked by a nascent and largely unmonitored industry, necessitates a proactive and collaborative approach to regulation and responsible AI development. Future trends indicate a deepening of emotional intelligence in AI, seamless integration with immersive technologies like AR/VR, and the continued emergence of embodied AI and even Brain-Computer Interfaces. While AI can scale information and provide invaluable assistance, the enduring need for genuine human connection and emotional intelligence remains paramount. The strategic imperative for developers, policymakers, and users alike is to navigate this evolving landscape with transparency, ethical foresight, and a commitment to fostering balanced human-AI relationships that augment, rather than diminish, human flourishing.

Works cited

1. What Are AI Companions? | Built In, accessed June 7, 2025, https://builtin.com/artificial-intelligence/ai-companions
2. AI Companion App Development - Young Decade, accessed June 7, 2025, https://www.youngdecade.com/ai_companion_app_development
3. Friends for sale: the rise and risks of AI companions | Ada Lovelace ..., accessed June 7, 2025, https://www.adalovelaceinstitute.org/blog/ai-companions/
4. What is multimodal AI: Complete overview 2025 | SuperAnnotate, accessed June 7, 2025, https://www.superannotate.com/blog/multimodal-ai
5. Towards LLM-Centric Multimodal Fusion: A Survey on Integration Strategies and Techniques - arXiv, accessed June 7, 2025, https://arxiv.org/html/2506.04788v1
6. Rakuten Product Conference, accessed June 7, 2025, https://product-conference.corp.rakuten.co.in/blog/Multimodal-AI-The-Convergence-of-Vision-Language-and-Beyond
7. Large Multimodal Models (LMMs) vs LLMs in 2025, accessed June 7, 2025, https://research.aimultiple.com/large-multimodal-models/
8. Beyond words: AI goes multimodal to meet you where you are, accessed June 7, 2025, https://news.microsoft.com/source/features/ai/beyond-words-ai-goes-multimodal-to-meet-you-where-you-are/
9. Agent AI: Revolutionizing Multimodal Interaction for Seamless ..., accessed June 7, 2025, https://www.datasciencesociety.net/agent-ai-revolutionizing-multimodal-interacti

on-for-seamless-human-computer-communication/
10. Multimodal Data Fusion: Key Techniques, Challenges & Solutions, accessed June 7, 2025, https://www.sapien.io/blog/mastering-multimodal-data-fusion
11. Multimodal AI | Google Cloud, accessed June 7, 2025, https://cloud.google.com/use-cases/multimodal-ai
12. AI Mode in Google Search: Updates from Google I/O 2025, accessed June 7, 2025, https://blog.google/products/search/google-search-ai-mode-update/
13. The Best AI Chatbots in 2025 - Section, accessed June 7, 2025, https://www.sectionai.com/ai/llm-guide
14. Introducing OpenAI o3 and o4-mini | OpenAI, accessed June 7, 2025, https://openai.com/index/introducing-o3-and-o4-mini/
15. OpenAI teams up with Jony Ive to create 100 million AI companion ..., accessed June 7, 2025, https://www.aibase.com/news/18282
16. Building with Claude - Anthropic, accessed June 7, 2025, https://docs.anthropic.com/en/docs/overview
17. The Claude 3 Model Family: Opus, Sonnet, Haiku - Anthropic, accessed June 7, 2025, https://www.anthropic.com/claude-3-model-card
18. Multimodal AI: Combining Text, Image, and Audio Understanding ..., accessed June 7, 2025, https://www.onyxgs.com/blog/multimodal-ai-combining-text-image-and-audio-understanding
19. Multimodal AI: Transforming Evaluation & Monitoring | Galileo, accessed June 7, 2025, https://galileo.ai/blog/multimodal-ai-guide
20. Multimodal Spoken Dialogue System with Biosignals - ACL Anthology, accessed June 7, 2025, https://aclanthology.org/2024.yrrsds-1.11/
21. Call for Papers – SIGDIAL 2024, accessed June 7, 2025, https://2024.sigdial.org/call-for-papers/
22. Abstracts: NeurIPS 2024 with Jindong Wang and Steven Euijong Whang - Microsoft, accessed June 7, 2025, https://www.microsoft.com/en-us/research/podcast/abstracts-neurips-2024-with-jindong-wang-and-steven-euijong-whang/
23. arxiv.org, accessed June 7, 2025, https://arxiv.org/pdf/2506.04788v1
24. NeurIPS 2024 Papers from Microsoft Research Asia - Microsoft ..., accessed June 7, 2025, https://www.microsoft.com/en-us/research/articles/msra-neurips-2024-papers/
25. Replika - AI Friend on the App Store, accessed June 7, 2025, https://apps.apple.com/us/app/replika-ai-friend/id1158555867
26. Exploring the replika ai api: a developer's guide to intelligent business tools - BytePlus, accessed June 7, 2025, https://www.byteplus.com/en/topic/552289
27. Replika - WayToAGI, accessed June 7, 2025, https://www.waytoagi.com/sites/1271
28. Character.AI Launches Multi-Modal Features to Power the Future of Storytelling | daily.dev, accessed June 7, 2025, https://app.daily.dev/posts/character-ai-launches-multi-modal-features-to-power-the-future-of-storytelling-tgdprwrgy
29. Character.AI Expands Storytelling with Multi-Modal Features and Video

Generation, accessed June 7, 2025, https://app.daily.dev/posts/character-ai-expands-storytelling-with-multi-modal-features-and-video-generation-otfsqin5s
30. Character.AI Unveils Video Generation Tool, Community Feed and ..., accessed June 7, 2025, https://www.gadgets360.com/ai/news/character-ai-unveils-multimodal-video-generation-community-feed-8577527
31. Character.AI launches multimedia features including video ..., accessed June 7, 2025, https://sp-edge.com/updates/44148
32. Kuki AI - hybrid rituals, accessed June 7, 2025, https://hybrid-rituals.com/product/ai-tools/misc-ai-tools/chatbot/kuki-ai/
33. Kuki AI - Callin.io, accessed June 7, 2025, https://callin.io/ai-tools/kuki-ai/
34. Kuki AI: Your Fun & Entertaining Virtual AI Friend - Deepgram, accessed June 7, 2025, https://deepgram.com/ai-apps/kuki-ai
35. Moemate Alternatives in 2025 | Best Moemate Alternatives - Toolify, accessed June 7, 2025, https://www.toolify.ai/alternative/moemate
36. Moemate vs Veo 3 AI - compare the differences between Moemate vs Veo 3 AI - Toolify, accessed June 7, 2025, https://www.toolify.ai/compare/moemate-vs-veo-3-ai-1
37. Candy AI: Create Your own AI Companion for Personalized Conversation - OutRightCRM, accessed June 7, 2025, https://www.outrightcrm.com/blog/candy-ai/
38. candy-ai-fan-blog/candy-ai-reviews-is-it-safe · Hugging Face, accessed June 7, 2025, https://huggingface.co/candy-ai-fan-blog/candy-ai-reviews-is-it-safe
39. Top 10 AI Friends to Chat With in 2025: Best AI Companions for Startups - Fe/male Switch, accessed June 7, 2025, https://www.femaleswitch.com/entrepreneurial_potential/tpost/ti521ci431-top-10-ai-friends-to-chat-with-in-2025-b
40. AI Girlfriend Chat Alternatives in 2025 - Toolify.ai, accessed June 7, 2025, https://www.toolify.ai/alternative/ai-girlfriend-chat
41. GirlfriendGPT - Reviews, Pricing, Features - SERP AI, accessed June 7, 2025, https://serp.ai/products/gptgirlfriend.online/reviews/
42. GirlfriendGPT - Best AI apps of 2024, accessed June 7, 2025, https://www.thesamur.ai/pt/apps/girlfriendgpt
43. Anakin AI - Service List, accessed June 7, 2025, https://servicelist.io/company/anakin-ai?ab374ee6_page=5&b81c652f_page=1
44. What is a Multimodal AI Companion? - otherhalf.ai, accessed June 7, 2025, https://otherhalf.ai/blog/definition/what-is-a-multimodal-ai-companion
45. Top 10 AI XXX Chatbots of 2025 You Cannot Miss Out - Anakin.ai, accessed June 7, 2025, http://anakin.ai/blog/top-10-ai-xxx-chatbots-of-2025-you-cannot-miss-out/
46. Top 10 Free AI Sex Chat Platforms, accessed June 7, 2025, http://anakin.ai/blog/free-ai-sex-chat-platforms/
47. How to Customize Your 3D Anime AI Companion's Personality ..., accessed June 7, 2025,

https://otherhalf.ai/blog/tutorial/how-to-customize-your-3d-anime-ai-companion-s-personality
48. The Best AI Companion for Anime Fans - otherhalf.ai, accessed June 7, 2025, https://otherhalf.ai/blog/audience/the-best-ai-companion-for-anime-fans
49. Welcome to Hiwaifu | Hi, Waifu, accessed June 7, 2025, https://docs.hiwaifu.com/
50. AI Chatbots | SmartSocial.com, accessed June 7, 2025, https://www.smartsocial.com/post/ai-chatbots?a5a4e696_page=2
51. Hi,Waifu : AI Friend, accessed June 7, 2025, https://beta.hiwaifu.com/
52. Talkie AI: Understanding Evolution & Implications of Conversational, accessed June 7, 2025, https://store.outrightcrm.com/blog/talkie-ai/
53. Talkie: AI-Powered Conversational Characters - Deepgram, accessed June 7, 2025, https://deepgram.com/ai-apps/talkie
54. Talkie.ai | AI Phone Automation for better Patient access, accessed June 7, 2025, https://talkie.ai/
55. How To Develop An AI Companion App Like SimSimi - Hyperlocal Cloud, accessed June 7, 2025, https://hyperlocalcloud.com/blog/how-to-develop-an-ai-companion-app-like-simsimi
56. Top SimSimi Alternatives - WebCatalog, accessed June 7, 2025, https://webcatalog.io/apps/simsimi/alternatives
57. What are AI Companions? Features & Examples - ZEGOCLOUD, accessed June 7, 2025, https://www.zegocloud.com/blog/ai-companions
58. SimSimi APK for Android - Download, accessed June 7, 2025, https://simsimi.en.softonic.com/android
59. SimSimi - Apps on Google Play, accessed June 7, 2025, https://play.google.com/store/apps/details?id=com.ismaker.android.simsimi
60. X-Origin AI Introduces Yonbo: The Next-Gen AI Companion Robot ..., accessed June 7, 2025, https://enmobile.prnasia.com/releases/global/x-origin-ai-introduces-yonbo-the-next-gen-ai-companion-robot-designed-for-families-491213.shtml
61. BEST AI PERSONAL ROBOTS I SAW AT CES 2025! — Eric J. Kuhns, accessed June 7, 2025, https://www.ericjkuhns.com/blog/best-tech-i-saw-from-ces-2025-robots-ai-and-more
62. Top 5 AI Companion Robots That Will Change How We Live in 2025 ..., accessed June 7, 2025, https://www.jumpstartmag.com/top-5-ai-companion-robots-that-will-change-how-we-live-in-2025/
63. Smarter Together: Why AI Needs Emotional Intelligence to Lead the ..., accessed June 7, 2025, https://www.talentsmarteq.com/why-ai-needs-emotional-intelligence-to-lead-the-future/
64. How AI Companions Are Shaping the Future of Work and Life - AP Trio, accessed June 7, 2025, https://aptrio.com/how-ai-companions-are-shaping-the-future-of-work-and-life/

65. Different Use Cases of AI Virtual Assistants - Examples and Benefits ..., accessed June 7, 2025, https://www.zartis.com/ai-virtual-assistants/use-cases-examples-benefits/
66. Everything about Joyland AI: Features, Pricing & Best Alternatives - TECHi, accessed June 7, 2025, https://www.techi.com/joyland-ai-features-pricing-alternatives/
67. Joyland AI: Your Ultimate Guide to Virtual Companionship and Interactive Storytelling, accessed June 7, 2025, https://claudeai.guru/joyland-ai-anime-chatbot-guide/
68. 10 Biggest Limitations of Large Language Models - ProjectPro, accessed June 7, 2025, https://www.projectpro.io/article/llm-limitations/1045
69. Limitations of LLMs: Bias, Hallucinations, and More - Learn Prompting, accessed June 7, 2025, https://learnprompting.org/docs/basics/pitfalls
70. Are AI Friendships Helpful or Harmful? Exploring the Pros and Cons ..., accessed June 7, 2025, https://www.meridian-counseling.com/blog/ai-friendships-helpful-or-harmful
71. New Study Explores Artificial Intelligence (AI) and Empathy in ..., accessed June 7, 2025, https://www.evidencebasedmentoring.org/new-study-explores-artificial-intelligence-ai-and-empathy-in-caring-relationships/
72. Multimodal AI Development Services - AI Company - XenonStack, accessed June 7, 2025, https://www.xenonstack.com/artificial-intelligence/multimodal-ai/
73. AI Chatbot App Review - BytePlus, accessed June 7, 2025, https://www.byteplus.com/en/topic/417691