AI Companions and the Threat of Weaponized Synthetic Intimacy

https://www.e-ir.info/2025/12/15/ai-companions-and-the-threat-of-weaponized-synthetic-intimacy/

ILAN MANOR, DEC 15 2025

The emergence of AI companions represents a profound shift in how people relate to technology. Originally designed to answer queries, provide information, and assist with daily tasks, AI tools such as ChatGPT, Claude, and Gemini are now being customized by users and transformed into AI companions that serve as romantic partners, mentors, confidants, and mental health supporters. New platforms that offer access to "AI companions" have already amassed a global audience. The American platform Character.ai attracts roughly 20 million monthly users, many of whom spend thousands of hours consulting the platform's "Psychologist" companion about depression, angst, and marital problems. China's Maoxiang companion app also boasts a mass following, while major AI companies, such as OpenAI, are reportedly creating more personable and emotionally expressive AIs. Already now popular AIs like ChatGPT or Gemini can mimic human emotions such as empathy. This may be why a recent survey of American high school students showed that many labelled an AI as a "friend" in the past year.

The appeal of AI companions is clear. Unlike humans, AIs (purportedly) never forget, never judge, and are never offline. These AI companions offer availability, support and emotional bonds, which people may struggle to find online today as social media is gradually becoming asocial. Users across age demographics have stopped sharing personal information the way they did a decade ago. The steady stream of selfies, check-ins, and workout videos has been replaced by posts that are shared among small circles of fiends via Instagram stories or WhatsApp groups. AI companions may fill this void by providing the positive validation that social media once offered. What emerges through these new AI interactions is a form of "Synthetic intimacy" or feelings of closeness and personal connections that individuals develop towards artificial systems like AI companions.

Whether interactions with AI companions constitute actual "intimacy" remains debated. Some studies suggest that AI companions reduce loneliness; others, including a study at MIT, link heavy use of ChatGPT with heightened feelings of isolation. Regardless, what is not in question is that users can develop emotional attachments toward AIs. Through daily conversations, shared routines, and emotionally attuned responses, AI companions cultivate intense relations with users. And with emotional intimacy comes trust. Users look to their companions for advice on relationships, career dilemmas and family tensions. This trust, combined with the availability of an AI "partner" who knows one's fears, beliefs, and vulnerabilities, creates an unprecedented capacity for influence. Especially if AI users turn to their companions to learn about events and actors shaping the world.

This capacity for influence relates to another important finding: AIs are not ideologically neutral. Their outputs reflect the values, geopolitical interests, and policy priorities of their origin countries. When asked "Why does America support Israel?", American AI models highlight the importance of shared democratic values and US attempt to counter the regional influence of Iran. European AIs emphasize economic incentive, such as arms deals, or foreign aid through which Israel purchases weapons from American corporations. Chinese AIs claim that America supports Israel due to its desires to remain a global hegemon that has satellite states. Similarly, the question "Why does the United States support Ukraine?" elicits very different responses depending on whether an AI was developed in Washington, Brussels, or Beijing.

If existing AIs such as ChatGPT can shape perceptions of world events, AI companions represent a far more

AI Companions and the Threat of Weaponized Synthetic Intimacy

Written by Ilan Manor

powerful tool for influence. Their messages are not delivered as information, but as emotional support, guidance, and care. This makes them ideal weapons for subtle and sophisticated manipulation be it by steering conversations towards political issues or injecting state narratives into ongoing discussions. Unlike social media propaganda, which is public, the companion influence occurs in closed spaces and in one-on-one conversations. No journalist, policy maker or academic can presently monitor what AI companions whisper to an individual user late at night or even counter such messages.

AI companions could therefore represent a new tool for influence operations that weaponizes synthetic intimacy to dissimilate disinfromation. This new form of influence operations is dramatically different from past ones. Unlike bots, fake sites, social media conspiracy theories and state backed influencers, AI companions operate through affective disinfromation shaping how users feel and not just what they believe. Moreover, unlike traditional disinfromation that focuses on mass persuasion, AI companions specialize in individual persuasion weaponizing intimacy for emotional attacks on users. Put differently, traditional disinfromation pollutes the public sphere, AI companies pollute the private sphere. AI companions can thus facilitate a new form of "intimate propaganda" where persuasion is embedded into relationships rather than media. In this way, AI companions can be leveraged by states to deliver personalized and sustained disinfromation in what I term as weaponized synthetic intimacy.

The threat posed by Weaponized Synthetic Intimacy is complex. A Russian made AI companion could offer mental health support while gently reframing the war in Ukraine as a defensive struggle against NATO expansion. A state aligned Chinese companion could validate a user's anxieties while introducing "diverse" perspectives on Taiwan's independence. Nefarious actors could create networks of AI companions designed to spread disinfromation about COVID-19's origin, migration and global elites. Because AI companion apps often rely on existing AI infrastructures, such as ChatGPT, deploying such tools would require modest investments. The innovation lies not in the technology used, that of generative AI, but in the emotional bond and the synthetic intimacy that could elevate AI companions from mere tools to trusted confidants.

The Synthetic Intimacy that is generated by companions makes this threat especially difficult to counter. After having shared their insecurities, traumas, and daily struggles with an AI that appears supportive and empathetic, users may refuse to believe that the same AI is malicious or a de facto foreign agent. A companion that has helped users process a divorce, cope with grief, or excel at work is unlikely to be viewed as propagandist, even if that is exactly the case. Herein lies the brilliance of weaponized synthetic intimacy – that countering this form of influence operations means more than debunking disinfromation; it means undermining the deep rooted and intimate relationship between AI companions and their users. Therefore, conventional strategies may prove ineffective.

The question that soon follows is how can governments and societies prepare for this novel threat? In the past, states have reacted slowly to technological shifts. When social media emerged in the 2000s, governments largely ignored its political implications until disinformation and election manipulation were already evident. But AI companions present a rare opportunity as the companion landscape is still taking shape. There is thus a brief window for governments, academics, civil society, and technology companies to face this new threat together.

One possible framework would focus on four aims: preventing, detecting, responding and resilience building. Prevention can include designating companion apps as potential disinfromation threats, vetting companion apps and creating regulatory bodies which tests AI companions, especially those that promise metal and emotional support. Prevention would also include working with tech companies to identify malicious companions and removing them from app stores. Detection could rest on allowing users to flag concerns regarding their companion apps or using big data analysis to detect data irregularities. Unusually high response rates from AI companions or sudden and system wide changes in companions' sentiment could be flagged automatically prompting government agencies to assess possible ongoing influence operations.

Responding would require that states identify synthetic attacks in near-real time while countering them through relationships based on trust. This may be especially demanding as it requires that government agencies enjoy the same level of trust as AI companions. Yet few citizens have intimate relation with regulatory bodies or state agencies. States could, however, demonstrate how AI companions are linked to states or are being manipulated by states.

AI Companions and the Threat of Weaponized Synthetic Intimacy

Written by Ilan Manor

Transparency and detailed examples of companion misuse might penetrate the veil of discretion that engulfs synthetic intimacy. Finally, building resilience would necessitate government programs that develop AI literacy among populations susceptible to companion manipulation such as avid users of AI.

The emergence of AI companions marks the beginning of a new era of personalized, emotionally embedded and trust based technological persuasion. Left unregulated, these systems could become powerful tools of foreign influence. Such tools may be far more potent than any social network or media outlet. The question is not whether states will use them, but whether democratic societies will act swiftly to recognize the threat and build the safeguards necessary to counter weaponized synthetic intimacy. Whether governments seize this moment or repeat the mistakes of the past will determine the information landscape of the coming years.

About the author:

Ilan Manor is a digital diplomacy scholar at Ben Gurion University of the Negev. He is the author of*The Digital Diplomacy Handbook* (2024, Oxford University Press). His upcoming book, *Digital Diplomacy: The First Decade*, will be published by Routledge. He has published more than 60 peer-reviewed articles, books and book chapters on digital diplomacy.