POLICY BRIEF

Recommendations on Companion AI (CAI)

Four areas of action with concrete regulatory, enforcement, and funding policy measures.

Terminology

- CAI Apps: Standalone companion AI applications such as Replika, Character.AI, Nomi.
- CAI Function in LLM: Companion-like interaction in general-purpose language models such as ChatGPT, Gemini, and Claude.

01

Protecting individual users

Effectively enforce data protection

The use of sensitive Data derived from intimate self-disclosure must be strictly limited to what is necessary for the application's core functionality. (purpose limitation). The use of sensitive data for advertising or other commercial purposes within the framework of an "intimacy economy" must be prohibited.

Mandate context-based crisis response mechanisms

Providers should be required to intervene promptly, actively, and proportionately at the first signs of a crisis or suicidal statements, and to provide easily accessible pathways to professional support services.

Separate companion functions in LLMs

Risk assessment, risk management, and the fulfillment of legal obligations require a defined purpose of use. Companion functions, understood as persistent, persona-based conversations with simulated emotional attachment, should be offered in LLMs in clearly separated modes with their own risk management, separate data processing, and their own age verification.

Prioritize dialogue- and context-based youth protection

A dialogue-based protective measure is recommended as the method of choice for implementing youth protection measures. When indications that a user is a minor are detected during the dialogue, providers should gently interrupt the interaction and refer users to age-appropriate alternatives and support persons. Such a measure would be less intrusive than an upfront identity or age verification, as it does not require additional data collection and only responds to recognizable protection needs as they arise. Article 28(3) DSA clarifies for platforms that there is no obligation to process additional data, but at the same time requires a high level of

protection. This level of protection should therefore be achieved not primarily through the collection of additional data, but through risk-appropriate protective measures within the interaction.

Expand Annex III of the AI Act to include AI-driven Manipulation ● ●

Annex III should include a separate section for AI systems whose purpose is to manipulate human decision-making, behavior, or emotions. The Commission should, pursuant to Art. 112(2)(a), ensure that there is sufficient time for a corresponding adjustment, as the negotiations on the Digital Omnibus have provided for an extension of the implementation deadline for the relevant obligations until December 2, 2027, through new decisions.[3]

Enforce existing AI bans ●

Companion AI systems must be immediately reviewed by the enforcement bodies of the member states and the AI Office for prohibited practices under Art. 5 AI Act. Immediate action must be taken in the event of violations.

Correct EU guidelines on prohibited AI practices ● ●

The companion AI example in para. 134 of the guidelines on Art. 5 AI Act (Commission 2025) should be deleted, as it incorrectly classifies companion AI as harmless. The entry in para. 88, which classifies these systems as harmful, should be retained.

Verification of compliance with the GAI Code of Conduct ●

The Federal Network Agency and the Commission should systematically verify whether signatories to the GPAI Code of Conduct, such as OpenAI with ChatGPT, Google with Gemini, and Microsoft with Copilot, have complied with the measures required to implement this code.

Resubmission of the AI Liability Directive ● ●

A new proposal should provide for a lower burden of proof and a presumption of causation in favor of those harmed by AI, and build on the negotiation results already achieved. Anyone who introduces potentially dangerous AI and accepts the materialization of known risks should, in the event of harm, be required to prove that a deterioration in health or a suicide is not attributable to the conversation with the companion app.

Consulting the CAI incident database ● ●

Incidents involving Companion AI, which have come to light primarily through ongoing legal proceedings, are compiled in a CAI incident database. This database is

[3] See press release dated May 7, 2026.

continuously updated to facilitate risk assessment and improve access to ongoing proceedings for litigation work by consumer protection agencies, regulatory authorities, and NGOs.

02

Ensuring information integrity and decision-making autonomy

Clearly separate promotional content visually

Advertising content should be clearly separated from outputs through consistent visual distinction, such as a color-coded box next to or below the text. Tests show that currently emerging labeling practices are often overlooked. Integration into the output text must be ruled out. Advertising must not influence the generation of outputs.

Classify ChatGPT as a VLOSE

ChatGPT should be classified immediately as a very large online search engine within the meaning of Art. 33(1) DSA.

Supplementary review of classification as a VLOP

Additionally, it should be examined whether ChatGPT should be classified as a very large online platform under Art. 33 DSA to ensure the application of Art. 25(1) DSA for protection against manipulative design elements, as well as Art. 28 DSA for enhanced youth protection.

Enact the Digital Fairness Act

The Digital Fairness Act (DFA) could provide additional protection against risks posed by Companion AI by addressing addictive design, dark patterns, manipulative personalization, and AI-supported forms of interaction such as chatbots.

03

Prevent a reduction in the existing level of protection

Urgently stop the reduction of protection for sensitive data

The changes proposed in the Digital Omnibus to reduce the protection of sensitive data in the context of AI should be rejected. Intimate conversation data from minors and adults is obtained through continuous encouragement to self-disclose and is already being exploited for commercial purposes. With the announced introduction of advertising in LLM systems, a further expansion of this exploitation is imminent.

04

Protect fair competition

Update the blacklist of unfair business practices

The appendix to Section 3(3) of the Unfair Competition Act (UWG) lists practices that are always considered unfair without the need for a case-by-case review. AI-driven manipulative practices should be added to the UWG's blacklist to protect competitors who do not engage in such practices.