AI Companions for Mental Health: Regulatory Qualification Between Intended Use and Functional Reality. Comparative Remarks Between the EU and the US

Francesca Gennari, Sofia Palmieri*

ABSTRACT: This article examines the regulatory landscape concerning AI Companions between the EU and the US. It concludes that in both legal systems the intended use of a product or service expressed by the manufacturer is used as a decisive element whenever it is dubious as to where a device is or not a medical device. It is argued that in many cases, the prevalence of the manufacturer's intended use in dubious cases is used to market products as "simple" consumer products whereas, instead, they would need a clinical validation under the medical devices regulations of the two legal systems. This becomes more evident with AI Companions, LLM-based (for now) services which offer companionship to users that are generally not categorized as a medical device. The growing cases of suicide, self-harm, depression, and addiction by young people highlight the limits of the primacy of the manufacturer's intended use when defining or not an object as a medical device. As authors, we argue in favour of a functional qualification of the intended use, as both the current EU and US regulatory landscape do not ensure a sufficiently high level of care and protection of consumers and patients alike.

KEYWORDS: AI Companions; intended purpose; functional qualification; medical devices; EU-US

SUMMARY: 1. Introduction – 2. The Rise of AI Companions: a Social and Technical State of the Art – 3. What Are AI Companions? A Methodology Clarification – 4. Regulation of AI Companions in the EU – 4.1. AI Companions Under EU Consumer and Digital Services Law – 4.2. AI Companions Under Sector-Specific Product Regulations – 4.2.1. AI

* *Francesca Gennari: Research affiliate at Sant'Anna School of Advanced Studies, Pisa, Italy. Mail: francesca.gennari@santannapisa.it; Sofia Palmieri: Post-Doctoral Fellow in Medicine, Artificial Intelligence, and the Law, Harvard Law School, Cambridge, US. Mail: spalmieri@law.harvard.edu. The authors contributed equally to the creation, draft and review of this essay. More specifically, the sections were divided among the authors as follows: Section 1: PALMIERI – Section 2: PALMIERI – Section: 3 -GENNARI – Subsection 3.1: PALMIERI – From subsection 3.2. to subsection 3.2.2: GENNARI— Section 4: PALMIERI– Section 5: PALMIERI – GENNARI. The Introduction and Conclusions are common.* *Acknowledgment and Funding: Palmieri's research for this essay was supported by a Novo Nordisk Foundation Grant for a scientifically independent International Collaborative Bioscience Innovation & Law Programme (InterCeBIL Programme — grant number NNF23SA0087056). Francesca Gennari's research for this essay was supported by the Fit4MedRob project. This work was supported by the Italian Ministry of University and Research, under the complementary actions to the NRRP "Fit4MedRob – Fit for Medical Robotics" Grant (#PNC0000007). This article was subject to double blind peer review process.*

Companions Under the Medical Devices Regulation (MDR) – 4.2.2. AI Companions Under the AI Act – 5. Regulation of AI Companions in the US – 6. Towards a Functional Approach to Regulatory Qualification – 7. Conclusions.

1. Introduction

AI companions are increasingly deployed as tools for emotional support, yet their legal qualification remains uncertain despite their growing role in mental-health-adjacent interactions. Most tools used for mental health in this space, particularly among young people, are based on generative AI (or, more correctly, foundation models) [1] and delivered as chatbots apps.[2] A growing body of literature has argued that AI systems performing therapeutic functions should fall within medical device regulation. However, a prior and unresolved question remains: how to identify when such therapy-like functions arise in systems that are deliberately framed as non-medical?

This article uses the term "AI companion" (hereinafter abbreviated with AC) as intended by Rogge.[3] While a universally accepted definition of artificial AC has historically remained elusive due to the interdisciplinary nature of the field, a common denominator can be found in the core qualities of the agent's social design. Borrowing from this recent literature, an AC is defined as a social agent characterized by an adaptive and engaging design that pursues the formation of emotional bonds and long-term relationships with its users. In this framework, adaptivity refers to the agent's ability to recognize and react to a user's individual preferences and situational context, while engagement involves the agent's proactive ability to initiate interactions and respond to the user on an emotional level, effectively reducing its status as a mere "tool" and repositioning it as a social actor.

We adopt this broader definition for three main reasons. The first one is that, given the strides AI has had recently, we cannot exclude further advancement on this technology, or if it will merge with other available technologies, such as Augmented Reality (AR), Virtual Reality (VR) or Extended Reality (ER).[4] Using only the expression AI chatbot risks limiting our analysis to just one form of technological support, although we recognize that it is the most used at the moment of writing.

---

[1] R. BOMMASANI *et al.*, *On the Opportunities and Risks of Foundation Models*, in *ArXiv*, 2021.

[2] A. ABD-ALRAZAQ, M. ALAJLANI, A. ALALWAN, B. M. BEWICK, P. GARDNER, M. HOUSEH, *An overview of the features of chatbots in mental health: A scoping review*, in *International Journal of Medical Informatics*, 132, 2019, 103978.

[3] In that literature, the term generally refers to AI systems designed to sustain ongoing, personalized interactions with users and to simulate elements of companionship through conversational, affective, or relational engagement. The choice of this terminology is deliberate. While different expressions – such as *chatbots*, *conversational agents*, or *digital therapeutics* – appear across disciplines, "AI companion" captures a broader class of systems characterized not merely by task-oriented interaction but by the construction of a quasi-relational dynamic with the user. Because this concept is more widely discussed in the human-computer interaction and psychology literature than in legal scholarship, its use here warrants explicit clarification. For this reason, the article adopts the term while acknowledging that alternative terminologies exist and that the concept has not yet been fully consolidated within legal doctrine. For more insights on the matter, see A. ROGGE, *Defining, Designing and Distinguishing Artificial Companions: A Systematic Literature Review*, in *International Journal of Social Robotics*, 15, 2023, 1557-1579.

[4] M. CHESSA, L. GERINI, R. HUSSAIN, M. MARTINI, M. PIZZO, F. SOLARI, E. VIOLA, *Extended Reality and Artificial Intelligence for Exergaming: Opportunities and Open Challenges for Rehabilitation and Cognitive Training*, in *8th IEEE Forum on Research and Technologies for Society and Industry Innovation* (RTSI 2024), 2024, 357-362.

Secondly, AI companion is already used in several general audience documents to refer to the object of our analysis.[5] In our view, it can facilitate the association of our reflections to real-life cases and example. Thirdly, and, we might add, most importantly, the ambiguity with which these tools are released in the market: while typically framed as tools for companionship or well-being, they are frequently used by individuals for purposes that resemble mental-health support.

This article addresses that gap by examining how AI Companions are defined, marketed, and used in the EU and the United States. It argues that current regulatory frameworks rely too heavily on manufacturers' declared intended use, allowing systems that effectively perform mental-health-related functions to avoid classification as medical devices. While some AI companions may possess what could be described as an "implicit medical function," others may operate primarily as personified interfaces or self-projective avatars, similar to digital avatars that have long been used for social or expressive purposes. Nonetheless, in both scenarios – whether an AI companion carries an undeclared therapeutic role or functions primarily as a conversational or expressive tool – regulatory frameworks should ensure that providers remain accountable, and where appropriate liable, for deploying systems that may prove unsafe, particularly for vulnerable populations such as teenagers and young adults. The broader claim developed in this article is that AI Companions expose a structural weakness in contemporary technology regulation. Both EU and US frameworks continue to rely heavily on manufacturers' declared intended use when determining whether a product qualifies as a medical device. This approach assumes that the relevant function of a technology can be identified ex ante and communicated through product labelling and marketing. AI Companions challenge that assumption. Through personalization, emotional responsiveness, memory, and long-term interaction, these systems may evolve into quasi-therapeutic actors whose practical role differs substantially from the role formally declared by their developers. The central question is therefore whether intended-use doctrines remain adequate when the most important functions of a technology emerge through interaction with users rather than through design alone.

The essay proceeds in four steps. First, it considers the social and technological drivers behind the rise of AI Companions and their use in mental-health contexts. Second, it clarifies the taxonomy and concept of "AI Companion" as used in this analysis. The same level of detail cannot be applied to the United States, largely due to the absence of comprehensive federal legislation governing AI. However, new national initiatives and emerging laws are beginning to address how AI Companions may be marketed and presented to users. The essay therefore concludes by proposing a functional approach to regulatory qualification – one that focuses on the actual effects and uses of these systems rather than solely on their declared intended purpose – in order to better capture the risks they may pose.

[5] A. WIENER, *Love in the Time of AI Companions*, in *The New Yorker*, 09/03/2026, https://www.newyorker.com/magazine/2026/03/16/love-in-the-time-of-ai-companions (Last accessed: 02/04/2026); J. BERNARDI, *Friends for Sale: the Rise and Risks of AI Companions*, in *Ada Lovelace Institute Blog*, 23/01/2025, https://www.adalovelaceinstitute.org/blog/ai-companions/ (Last accessed: 02/04/2026; J. BECK, *Friendship on Demand*, in *The Atlantic*, 18/03/2026, https://www.theatlantic.com/family/2026/03/ai-friendship-chatbot/686345/ (Last accessed: 02/04/2026); J. O'DONNELL, *I companion IA sono l'ultimo stadio della dipendenza digitale, e i legislatori prendono la mira*, in *MIT Technology Review Italia*, https://www.technologyreview.it/i-companion-ia-sono-lultimo-stadio-della-dipendenza-digitale-e-i-legislatori-prendono-la-mira/ (Last accessed: 02/04/2026).B.CALDERINI, *Agenda Digitale*, *AI companion e dipendenza emotiva: una sfida normativa urgente*, www.agendadigitale.eu (Last accessed: 02/04/2026), A. THIENE, *La generazione Z nel baratro digitale*, in *Persona e mercato*, 2025, 489 ff., https://www.personaemercato.it/wp-content/uploads/2025/07/Thiene.pdf (Last accessed: 02/04/2026).

Finally, the essay closes on a cautiously optimistic note, pointing to recent developments in both jurisdictions. Ongoing investigations by the European Commission and recent jury decisions in the United States concerning the harms associated with addictive design in digital platforms suggest a growing willingness to scrutinize technologies based on their real-world impact on users. The fact that regulators in the EU and courts in the US are increasingly prepared to hold everyday platforms liable despite their formally declared intended use gives reason to believe that a similar functional reasoning may soon be applied to AI Companions as well.

2. The Rise of AI Companions: a Social and Technical State of the Art

Across the European Union, mental health disorders – particularly depression, anxiety, and stress-related conditions – have reached levels that strain the capacity of existing health systems.[6] This trend is especially pronounced among adolescents and young adults, for whom social isolation, economic insecurity, and the cumulative effects of the COVID-19 pandemic have intensified pre-existing vulnerabilities.[7]

Importantly, the persistence of unmet mental health needs in Europe is not primarily a problem of formal entitlement. Even in countries with universal healthcare coverage, access to timely and appropriate mental health care is constrained by shortages of professionals, long waiting lists, uneven territorial distribution of services, high indirect costs, and persistent cultural stigma. These barriers disproportionately affect those already in vulnerable positions: adolescents, young adults, migrants, and individuals living in socio-economically disadvantaged or remote areas.

Conversational artificial intelligence systems marketed as *AI Companions*, *mental wellness chatbots*, or *emotional support tools* – which rapidly proliferated across the European digital market[8] – increasingly operate in ways that resemble substitutes for inaccessible mental-health support, even where they avoid explicit therapeutic claims.[9]

These tools are not necessarily conceived as mental-health interventions, and the role they ultimately play is often quite blurry. Many are marketed simply as conversational chatbots – something closer, at least on the surface, to a digital companion. These systems often occupy a social role analogous to that of a "friend", engaging user in sustained and emotionally responsive interactions. Yet unlike interactions with human beings, there is no clear code-switching between friend-like behaviour and therapeutic engagement. As we noted in Cohen and Palmieri, when speaking with a friend who happens to be a therapist, the relationship can shift depending on context.[10] In human interactions, a distinction typically exists

[6] N. AHMED *et al.*, *Mental health in Europe during the COVID-19 pandemic: a systematic review*, in *The Lancet Psychiatry*, 10(7), 2023, 537-556; C.S. GRØNKJÆR, R.H.B. CHRISTENSEN, D. KONDZIELLA, M.E. BENROS, *Mental health disorders before, during and after the COVID-19 pandemic: a nationwide study*, in *Brain*, 148(5), 2025, 1829-1840.

[7] S. BRANJE, *The impact of the COVID-19 pandemic on adolescent mental health across the world*, in *Current Opinion in Psychology*, 53, 2023, 101665; World Health Organization, *Mental Health Atlas 2023*, Geneva, 2023; European Commission, *Communication on a Comprehensive Approach to Mental Health*, COM(2023) 298 final, Bruxelles, 2023.

[8] M.L. WAINBERG, P. SCORZA, J.M. SHULTZ, L. HELPMAN, J.J. MOOTZ, K.A. JOHNSON, Y. NERIA, J.E. BRADFORD, M.A. OQUENDO, M.R. ARBUCKLE, *Challenges and Opportunities in Global Mental Health: a Research-to-Practice Perspective*, in *Current Psychiatry Reports*, 19(5), 2017, 28.

[9] Fortune Business Insights, *AI Companion Market*, www.fortunebusinessinsights.com (Last accessed: 02/04/2026).

[10] G.I. COHEN, S. PALMIERI, *AI Therapists vs Companions: Wellness, Licensure, and Liability*, in *AJOB – American Journal of Bioethics*, 26(2), 2026.

between informal social exchange and professional therapeutic engagement, often signalled explicitly or implicitly. Both parties understand that the relational frame has changed.

With AI chatbots, those signals are far less clear. The boundaries between companionship, advice, and something resembling therapy can easily blur. And those boundaries matter. When an interaction is framed as mental-health care, there are normally safeguards, expectations, and professional responsibilities that help mitigate potential harms and power asymmetries. When those same dynamics occur in an ambiguous conversational setting, many of those protections simply do not exist.

More broadly, there remains considerable confusion about what these chatbots are meant to be, what functions they are supposed to perform, and under what conditions they might – if ever – serve as meaningful allies in expanding access to care or improving the coverage of mental-health services. This blurring of roles creates what can be described as a functional–formal mismatch: systems that are formally positioned as non-medical tools may nevertheless perform functions that are substantively similar to mental-health interventions.

3. What Are AI Companions? A Methodology Clarification

The concept of an AI companion encompasses a broad and heterogeneous set of technologies. While public and regulatory discussions often refer to "AI Companions" as a distinct category, the boundaries between companion systems, conversational agents, wellness applications, and medical technologies remain fluid. The term is frequently used to describe products specifically marketed for companionship, such as Replika, yet companionship functions may also emerge in systems that were not originally designed or commercialized for that purpose. For this reason, it is useful to distinguish between "Companion AI" as a commercial category and AI Companions as a broader functional phenomenon. The latter refers to AI systems that establish sustained, personalized, and emotionally significant interactions with users, regardless of their formal regulatory classification or commercial positioning.

A useful starting point is the broader category of conversational AI systems. These systems differ considerably in their design objectives, intended uses, and regulatory implications. Nevertheless, many share technical characteristics that enable users to develop ongoing relational dynamics with them, including conversational continuity, memory, personalization, and emotional responsiveness. Consequently, companionship should be understood not as a discrete technological feature but as a spectrum that may emerge across different categories of AI systems.

Although the categories are neither mutually exclusive nor exhaustive, four broad typologies of AI Companions can be identified.[11]

First, general-purpose AI systems may function as companions despite not being designed primarily for companionship. Large language models such as ChatGPT or other conversational AI assistants are typically marketed as productivity, information, or creative-support tools. However, their ability to engage in open-ended dialogue, remember user preferences, and provide personalized responses can lead users to es-

[11] These categories were developed during the drafting of this article through desktop research using the search term "AI Companion". Given the nascent and rapidly evolving nature of this field, the authors consider this taxonomy a useful and workable framework for analysis. However, it should be regarded as provisional and may require refinement as the market, technologies, and scholarly understanding of AI companions continue to develop.

tablish relationships that extend beyond purely instrumental interactions. Increasingly, individuals report using such systems for emotional reflection, personal advice, or social interaction, even when these functions are not part of the provider's declared purpose. This demonstrates how companionship may emerge organically through patterns of use rather than through explicit design choices.

Second, there are companion-oriented AI systems specifically designed to simulate friendship, intimacy, or emotional relationships. These systems represent what is commonly understood as "Companion AI" in the commercial sense. Examples include applications such as Replika and similar platforms that encourage users to create persistent relationships with virtual interlocutors. Unlike general-purpose systems, their core value proposition is not information provision or task completion but emotional engagement itself. Such systems often incorporate anthropomorphic features, persistent memory, personality development, gamification mechanisms, and relational narratives intended to foster long-term user attachment. While they are generally marketed as entertainment or companionship products, they may nevertheless perform functions that overlap with emotional support or mental well-being.

Third, wellness-oriented AI applications occupy an intermediate position between companionship and healthcare. These systems are typically marketed as tools for emotional well-being, stress management, self-reflection, mindfulness, or psychological support rather than as medical products. Depending on their design and interface, however, they may facilitate interactions that closely resemble therapeutic conversations. Many wellness applications employ conversational interfaces, personalized coaching, mood tracking, or cognitive-behavioral techniques. Although they usually avoid explicit medical claims, the distinction between wellness support and mental-health intervention may become increasingly difficult to sustain where users rely on such systems to address emotional distress or psychological vulnerabilities.

Fourth, certain AI systems may be developed and marketed as medical devices intended to support individuals with diagnosed medical or psychological conditions. In this context, companionship functions may form part of a broader therapeutic intervention. Examples may include AI-enabled digital therapeutics, software as a medical device, or conversational systems designed to support patients experiencing depression, anxiety, loneliness, cognitive decline, or other health-related conditions. Unlike the previous categories, these systems are intended to produce medically relevant outcomes and therefore operate within established healthcare regulatory frameworks. Nevertheless, they often rely on many of the same interactional mechanisms found in non-medical companion systems, including personalization, conversational continuity, and emotional engagement.

These categories should not be understood as rigid classifications as they illustrate the continuum that exists between ordinary conversational systems, companion technologies, wellness applications, and regulated medical products. Indeed, one of the central challenges addressed in this article is that AI systems may migrate across these categories over time or simultaneously exhibit characteristics associated with more than one category. As a result, regulatory qualification based solely on a provider's declared intended purpose may fail to capture the actual functions performed by the system and the risks generated through its use. This ambiguity provides the foundation for the functional approach to regulatory qualification developed in the following sections.

4. Regulation of AI Companions in the EU

According to EU law,[12] AI Companions must be regulated and classified based on their function-namely, whether they qualify as medical devices or remain consumer products. Current regulatory frameworks rely predominantly on the notion of intended use, as defined by the manufacturer. The broad range of functions that AI Companions can perform in their current form – largely enabled by LLM and foundation-model architectures – is reflected in the complex regulatory landscape that applies to them. From a legal perspective, two distinct regulatory layers must be considered. First, regulators must determine whether the system qualifies as a medical device or remains a general consumer product, depending on its intended function.[13] Second, once the presence of artificial intelligence is established, AI-specific regulatory frameworks may also be triggered, such as the EU AI Act or related national legislation. We hereunder explore how EU and US deal with this.

Before considering whether an AI companion qualifies as a medical device – which remains the core argument of this paper – it is important to recognize that a number of EU horizontal regulatory instruments may already apply regardless of such classification. Consumer protection law, and increasingly digital-services regulation, focus less on the declared purpose of a product and more on the way it is marketed, designed, and deployed toward users. This perspective is particularly relevant for AI companions, whose risks often arise not from diagnostic or therapeutic claims but from the emotional and behavioural influence they may exert over users. The significance of EU consumer and digital-services law extends beyond the specific obligations they impose. These instruments increasingly evaluate technologies according to their foreseeable effects on users rather than solely according to their declared purpose. In this respect, they offer an important example of functional regulatory reasoning that may become relevant for the governance of AI companions. Hereunder, without the presumption of offering a complete account of the regulation applicable to AI companions, we hint at consumer law as a possible lens to read the AI Companions phenomenon.

4.1. AI Companions Under EU Consumer and Digital Services Law

From a consumer-law perspective, AI Companions occupy an inherently ambiguous position between entertainment products, companionship services, emotional-support tools, and mental-health-adjacent applications. Providers frequently market these systems as instruments for well-being, self-reflection, or companionship while simultaneously disclaiming any therapeutic purpose. Yet where the overall presen-

[12] For reasons of space and time, the authors decided not to focus on the legislation involving data sharing (such as GDPR and the upcoming EHDS) or all the recently approved cybersecurity regulations.

[13] It must be specified that, according to EU law, AI companions could also be considered games. This however depends, yet again, on the intended purpose of the manufacturer. The authors regret not having the time to investigate this research path further in this article. However, for completeness's sake, the "game" regulatory path must be mentioned and eventually be warned about. The discipline of games, especially the ones involving gambling, in the EU law is almost all delegated to member states, hence with the possibility of AI companions to use another legal loophole. For more info on the (non)regulation of gambling in the EU see A. LITTLER, *Regulatory Perspectives on the Future of Interactive Gambling in the Internal Market*, in *European Law Review*, 33(2), 2008, 211, available at SSRN: https://www.ssrn.com/ssrn/ (Last accessed: 02/04/2026).Nevertheless, this does not change the fact that, even if catalogued as games, AI companions would still fall within the large category of consumer products, although far less regulated than other products and services.

tation of a service creates the impression that it can meaningfully assist users in coping with loneliness, anxiety, emotional distress, or other psychologically sensitive situations, questions arise as to whether such representations accurately reflect the nature, capabilities, and limitations of the product. This issue is particularly significant where vulnerable users are involved and when products foreseeably affect mental well-being. Under the Unfair Commercial Practices Directive[14] commercial practices directed at vulnerable groups must be assessed in light of the characteristics of the targeted audience. Consequently, disclaimers buried in terms of service may not always be sufficient to neutralise expectations generated through a system's design, marketing, or operation. The regulatory choice to treat these systems as ordinary digital services rather than health-related products has distributive consequences. It shields developers from accountability while exposing users to risks that are neither clearly communicated nor adequately governed. It is indeed true that, when it comes to market AI Companions in the EU there is simply no match between following the AI rules concerning AI Companions and the plethora of norms (and its present uncertainty) combining the AIA with the MDR. It is hoped that, given the impossibility to force manufacturer to make AI Companions medical devices, it is advisable that the Commission, on its own or because of an advice from the scientific panel labels AI Companions GPAIs as high risk.

These concerns are reinforced by the Digital Services Act (DSA),[15] which reflects a broader European effort to address risks arising from digital services that affect users' autonomy, safety, and well-being. In particular, the Regulation places increasing emphasis on the protection of minors, the mitigation of systemic risks, and the prevention of manipulative interface design. Such concerns are especially pertinent in the case of AI Companions, whose business models often depend upon sustained engagement, emotional attachment, and highly personalized interactions. The relevance of these principles has become increasingly apparent through the European Commission's scrutiny of digital services employing addictive or manipulative design features capable of influencing user behaviour. Moreover, the Commission's 2025 Guidelines on the Protection of Minors under the DSA[16] emphasize the need to account for risks associated with emotional dependency, persuasive design, and prolonged engagement among children and adolescents. While AI Companions may not always fall neatly within the regulatory categories envisaged by the DSA, the Regulation nevertheless signals an important evolution in EU law: digital services are increasingly evaluated not only according to their declared purpose but also according to their foreseeable effects on users. This development is particularly relevant for AI companions because many of the risks they generate arise not from explicit therapeutic claims, but from their capacity to shape behaviour, influence decision-making, and foster emotionally significant relationships over time.

[14] Directive 2005/29/EC of the European Parliament and of the Council of 11 May 2005 concerning unfair business-to-consumer commercial practices in the internal market and amending Council Directive 84/450/EEC, Directives 97/7/EC, 98/27/EC and 2002/65/EC of the European Parliament and of the Council and Regulation (EC) No 2006/2004 of the European Parliament and of the Council (Unfair Commercial Practices Directive), http://data.europa.eu/eli/dir/2005/29/2022-05-28.

[15] Regulation (EU) 2022/2065 of the European Parliament and of the Council of 19 October 2022 on a Single Market For Digital Services and amending Directive 2000/31/EC (Digital Services Act) PE/30/2022/REV/1 OJ L 277, 27.10.2022, 1-102, http://data.europa.eu/eli/reg/2022/2065/oj.

[16] Communication from the Commission – Guidelines on measures to ensure a high level of privacy, safety and security for minors online, pursuant to Article 28(4) of Regulation (EU) 2022/2065 C/2025/6826 OJ C, C/2025/5519, 10.10.2025, http://data.europa.eu/eli/C/2025/5519/oj.

In a way, the UCPD and the DSA demonstrate that EU law already contains mechanisms capable of addressing some of the concerns raised by AI Companions even when such systems are not classified as medical devices. At the same time, these instruments remain only partially adapted to technologies that blur the distinction between consumer products, wellness services, and therapeutic interventions and thus ambiguity becomes particularly evident when considering whether AI Companions should instead be regulated as consumer products or, under certain circumstances, as medical devices.

4.2. AI Companions Under Sector-Specific Product Regulations

In terms of functional qualification, AI Companions – despite the 4-tier taxonomy mentioned earlier – may generally fall into two categories: (a) AI-powered consumer products or (b) AI-enabled medical devices. When AI Companions are marketed primarily for entertainment, companionship, or general well-being purposes, they are likely to be treated as consumer products placed on the EU market. In such cases, the applicable regulatory framework would in principle be the General Product Safety Regulation (GPSR).[17] However, the GPSR only mentions software as a security element or as a part of a product, which is considered *ex lege*, as an object with interconnected software content. Applying this framework to fully digital, disembodied systems such as AI Companions therefore raises interpretative challenges, as the regulation was not originally designed with standalone conversational AI services in mind.[18] The limitations of the GPSR are particularly visible in the context of AI Companions because the primary risks generated by these systems are not necessarily physical. Rather, they may involve psychological dependency, emotional manipulation, behavioural steering, or deterioration of mental well-being. Such harms do not fit easily within traditional product-safety paradigms. As a result, a regulatory gap emerges: the GPSR does not adequately address software systems that actively shape users' behaviour and decision-making in psychologically sensitive contexts.

In practice, however, users may turn to AI Companions for mental-health-related advice without explicitly seeking therapy – for example, when asking a system how to cope with distress or personal difficulties. This behaviour is understandable given the limited access to effective talking therapies, one of the major bottlenecks in modern psychiatry, where waiting lists can extend for months even in high-income coun-

[17] Regulation (EU) 2023/988 of the European Parliament and of the Council of 10 May 2023 on general product safety, amending Regulation (EU) No 1025/2012 of the European Parliament and of the Council and Directive (EU) 2020/1828 of the European Parliament and the Council, and repealing Directive 2001/95/EC of the European Parliament and of the Council and Council Directive 87/357/EEC, PE/79/2022/REV/1 OJ L 135, 23.5.2023, 1-51, http://data.europa.eu/eli/reg/2023/988/oj.

[18] Reference to Art. 3(1) General Product Safety Regulation. Notice that as the GPSR was approved just a few months before the AI Act, there is no explicit references to it. Only by reading Article 6 GPSR concerning the safety of products one can notice the indirect reference to AI. See Art. 6(1)(b) mentioning the effect of the product on other products also by means of interconnection. For a better assessment of the connections of the AI Act and the GPSR, see F. GENNARI, *Mixed-Functions IoT Devices: Regulatory and Liability Requirements – A First Overview*, in *European Journal of Risk Regulation*, 15(4), 2024, 919-921.

tries.[19]In low- and middle-income countries, an AI might even be the only possible access point to therapy.[20]

4.2.1. AI Companions Under the Medical Devices Regulation (MDR)

The central difficulty is that the MDR continues to rely on intended purpose as the primary gateway to regulation. The question is whether this approach remains defensible when AI Companions may perform quasi-therapeutic functions irrespective of the claims made by their developers. Where an AI system is intended to diagnose, prevent, monitor, predict, treat, or alleviate disease, the applicable regulatory framework in the EU becomes the Medical Devices Regulation (MDR) (EU) 2017/745.[21] Under Article 2(1) of the MDR, software qualifies as a medical device when it is intended by the manufacturer to perform such medical functions, meaning that AI systems providing therapeutic or decision-support functions in mental health contexts may be classified as Software as a Medical Device (SaMD).[22] Yet determining whether an AI companion falls within the medical device category is far from straightforward.[23] EU jurisprudence and regulatory guidance emphasise the central role of the manufacturer's intended use in the classification process. The Court of Justice of the European Union (CJEU) has clarified that when the medical nature of a product is uncertain, the manufacturer's stated purpose plays a decisive role in the classification process.[24]

This creates a divergence between declared intended use and foreseeable functional use. Where users reasonably rely on a system for mental-health support, its exclusion from medical-device regulation becomes increasingly difficult to justify from a risk-based perspective.

Developers frequently frame such systems as tools for well-being, reflection, or companionship, thereby avoiding the regulatory burdens associated with medical-device qualification. This phenomenon is not unique to AI Companions and has previously been observed in other digital health contexts, particularly with mixed-function IoT systems.[25] Developers typically avoid explicit therapeutic claims, framing their

---

[19] S. SINGER *et al.*, *Effects of a statutory reform on waiting times for outpatient psychotherapy: a multicentre cohort study*, in *Counselling and Psychotherapy Research*, 22, 2022, 982-997.

[20] UNICEF, *Mental health a human right, but only 1 psychiatrist per 1,000,000 people in sub-Saharan Africa*, https://www.unicef.org/ (Last accessed: 02/04/2026).

[21] Regulation (EU) 2017/745 of the European Parliament and of the Council of 5 April 2017 on medical devices, amending Directive 2001/83/EC, Regulation (EC) No 178/2002 and Regulation (EC) No 1223/2009 and repealing Council Directives 90/385/EEC and 93/42/EEC, OJ L 117, 5.5.2017, 1-175, http://data.europa.eu/eli/reg/2017/745/oj.

[22] K. LUDVIGSEN, S. NAGARAJA, A. DALY, *When Is Software a Medical Device? Understanding and Determining the "Intention" and Requirements for Software as a Medical Device in European Union Law*, in *European Journal of Risk Regulation*, 13(1), 2022, 78-93.

[23] M. OSTERMANN, O. FREYER, F.G. VERHEES *et al.*, *If a Therapy Bot Walks Like a Duck and Talks Like a Duck Then It Is a Medically Regulated Duck*, in *npj Digital Medicine*, 8, 2025, 741.

[24] See judgments *Snitem* and *BrainProducts*, Judgment of the Court (Fourth Chamber) of 7 December 2017. Syndicat national de l'industrie des technologies médicales (Snitem) and Philips France v Premier ministre and Ministre des Affaires sociales et de la Santé, Case C-329/16, ECLI:EU:C:2017:947; Judgment of the Court (Third Chamber), 22 November 2012. Brain Products GmbH v BioSemi VOF and Others. Case C-219/11 ECLI:EU:C:2012:742. For commentary refer to T. MINSSEN, M. MIMLER, V. MAK, *When Does Stand-Alone Software Qualify as a Medical Device in the European Union? The CJEU's Decision in Snitem and What It Implies for the Next Generation of Medical Devices*, in *Medical Law Review*, 28(3), 2020, 615-624.

[25] F. GENNARI, *op. cit.*, 912-927.

products instead as tools for well-being, emotional reflection, or companionship.[26] Sometimes they might use some psychological therapy methods, such as CBT, but are not registered as medical devices.[27] The recent European Health Data Space Regulation[28] (Regulation (EU) 2025/327) may also become relevant where AI Companions process electronic health data or provide functionalities closely connected to health and wellness services. Although the EHDS is not primarily a product-safety instrument, it reflects the broader European trend toward treating digital health applications within an integrated governance framework centred on data quality, transparency, and user protection.

Having established how AI Companions may be classified depending on their intended function, a second regulatory layer must also be considered. Regardless of whether a system qualifies as a medical device or remains a consumer product, the presence of artificial intelligence itself may trigger additional regulatory obligations – or a different discipline – particularly under the EU AI Act (hereinafter abbreviated with AIA)[29] and US States' AI frameworks.

4.2.2. AI Companions Under the AI Act

The AI Act introduces a distinct regulatory layer that operates both dependently and independently from consumer-product and medical-device qualification. We will say more about that below. Consequently, even where an AI companion is not classified as a medical device, its underlying AI model may still trigger obligations under the AI Act.

a) Preliminary considerations: are AI Companions a prohibited practice under Article 5?

The European Union has adopted a horizontal regulatory framework intended to address risks arising from artificial intelligence. The AI Act does not explicitly refer to AI Companions; however, such systems may fall within the definition of a general-purpose AI (GPAI) model under Article 3(63) of the AIA. More specifically they are considered as General Purpose AI models, in short, GPAIs. In synthesis they are fuelled by generative AI and they are models "[…] trained with a large amount of data using self-supervision at scale, that displays significant generality and is capable of competently performing a wide range of distinct tasks regardless of the way the model is placed on the market and that can be integrated into a variety of downstream systems or applications, except AI models that are used for research, development or prototyping activities before they are placed on the market".

Any developer of an AI system – referred to as an "AI provider" under the AI Act[30] – must first verify whether the system falls within the scope of Article 5 AIA. In fact, Article 5 lists all the prohibited practices.

---

[26] X. ZHAO, *Top AI Mental Health Apps 2026*, https://www.myflourish.ai/ (Last accessed: 02/04/2026).

[27] P. DERENZO, *7 Best Mental Health AI Chatbots*, www.iask.ai (Last accessed: 02/04/2026).

[28] Regulation (EU) 2025/327 of the European Parliament and of the Council of 11 February 2025 on the European Health Data Space and amending Directive 2011/24/EU and Regulation (EU) 2024/2847, PE/76/2024/REV/1, http://data.europa.eu/eli/reg/2025/327/oj.

[29] Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence and amending Regulations (EC) No 300/2008, (EU) No 167/2013, (EU) No 168/2013, (EU) 2018/858, (EU) 2018/1139 and (EU) 2019/2144 and Directives 2014/90/EU, (EU) 2016/797 and (EU) 2020/1828 (Artificial Intelligence Act) PE/24/2024/REV/1 OJ L, 2024/1689, 12.7.2024, ELI: http://data.europa.eu/eli/reg/2024/1689/oj.

[30] According to Article 3(3) AIA, provider "[…] means a natural or legal person, public authority, agency or other body that develops an AI system or a general-purpose AI model or that has an AI system or a general-purpose AI model

That means a series of ways in which an AI system operates, regardless of its characteristics, hence whether it works with generative AI or not, that are considered incompatible with fundamental rights and societal interests.

Nevertheless, however, the AIA does not have an article on prohibited practices for GPAI models. In this article we adopt a teleological interpretation of the AIA which if the AI act's rationales are both the creation of safe AI products but also the protection of human rights according to Article 1 AI Act (AIA), then, it would be inconsistent to exclude GPAI models – including AI Companions – from the scope of Article 5. This seems to be logical, in order to avoid absurd legal effects such as that an illicit way of interacting with a user is sanctioned if one has an AI system and not sanctioned whether there is a GPAI model. This however does highlight a huge gap in protection as there is not a formal recall to Article 5 in the AIA chapter V, which is specifically devoted to GPAIs.

In particular, according to Article 5(1), (a) Article 5(1)(a) prohibits AI systems that "[…] deploys subliminal techniques beyond a person's consciousness or purposefully manipulative or deceptive techniques, with the objective, or the effect of materially distorting the behaviour of a person or a group of persons by appreciably impairing their ability to make an informed decision, thereby causing them to take a decision that they would not have otherwise taken in a manner that causes or is reasonably likely to cause that person, another person or group of persons significant harm".[31] Similarly, Article 5(1)(b) states that also an "AI system that simply any of the vulnerabilities of a natural person or a specific group of persons due to their age, disability or a specific social or economic situation, with the objective, or the effect, of materially distorting the behaviour of that person or a person belonging to that group in a manner that causes or is reasonably likely to cause that person or another person significant harm".[32]

One has to wonder, also with the help of the European Commission analysis guidelines,[33] first whether the AI Companions for mental health actually use techniques that we can call subliminal or that exploit vulnerabilities, and, secondly, if the reply is positive, that it cannot cause significant harm. It has been proved that the way that AI Companions work can impact sensibly on the behaviour of teen-agers, which are legally minors and still considered a vulnerable group. The increasing number of teenagers' suicide on both sides of the Ocean seems to be a witness of this danger.[34] However, in legal terms it would be the complainant that will still need to prove in a liability judgment, that the death of a loved one could be avoided if the AI system had not worked in the way it did. In the EU, a little help might be provided by the

developed and places it on the market or puts the AI system into service under its own name or trademark, whether for payment or free of charge".

[31] Emphasis added.

[32] Emphasis added.

[33] COMMUNICATION FROM THE COMMISSION Commission Guidelines on prohibited artificial intelligence practices established by Regulation (EU) 2024/1689 (AI Act), Brussels, 29.7.2025 C(2025) 5052 final.

[34] One of the many cases we report is the story of a teenager committing suicide to "reach" their loved Character.ai companion, in the US. L. KUENSSBERG, *"A Predator in Your Home": Mothers Say Chatbots Encouraged Their Sons to Kill Themselves*, in *BBC News*, https://www.bbc.com/news/articles/ce3xgwyywe4o (Last accessed: 02/04/2026). Or, in the EU, a similar case concerned the Nomi AI companion E. GUO, *An AI chatbot told a user how to kill himself – but the company doesn't want to "censor" it*, in *MIT Technology Review*, 06/02/2025, https://www.technologyreview.com/2025/02/06/1111077/nomi-ai-chatbot-told-user-to-kill-himself/ (Last accessed: 02/04/2026).

enactment of the New Product Liability Directive (NPLD).[35] According to Article 10 (2) NPLD in cases of extreme technical difficulty, the judge can presume the existence of a causal link and of defectiveness concerning the product. According to Article 4(1) NPLD a product can also be software. If we consider the definition of AI system at Art. 3(1) of the AI act, software is a much larger term, hence it is included. Going back to the evaluation that an AI provider has to do to decide whether the AI system or GPAI that it is designing falls in the two above-cited cases, must document how the AI system does not fall within the pre-determined conditions.

The interpretation of these obligations has been further clarified by the European Commission's Guidelines on the Scope of Obligations for Providers of General-Purpose AI Models under the AI Act (2025), which emphasize a functional assessment of providers' responsibilities across the AI value chain.

b) If AI Companions are not prohibited practices and are not medical devices, what rules should they comply with according to the AI Act?

If an AC neither qualifies as a prohibited practice nor is classified as a medical device by the AI provider, the question arises as to which regulatory regime under the AI Act applies to an AC, understood as a GPAI model. First, it has to be determined whether the GPAI entails or not a systemic risk. It must be said that there is a core regime, common to all the GPAIs, regardless, of whether they could entail or not systemic risk. The first default rule for GPAI is set at Article 53. AI providers of GPAI models must observe and that will become the main rule for compliance for AI Companions which are not marketed as medical devices. The following paragraphs will synthetically describe the obligations that are requested to AI providers of GPAIs (regardless of the risk that they entail).

c) The GPAIs core regime

The first thing is to fill in and update the "technical documentation of the model, including its training and testing process and the results of its evaluation, which shall contain, at a minimum, the information set out in Annex XI for the purpose of providing it, upon request, to the AI Office and the national competent authorities".[36] Secondly, Article 53 AIA requires developers of general-purpose AI models to prepare, maintain, and supply up-to-date documentation to providers integrating those models into downstream AI systems.[37] This duty must be fulfilled while safeguarding intellectual property, confidential business information, and trade secrets under Union and national law. This appears to be of difficult implementation, and it is expected that the AI Office should give guidance to MS as far as how IPRs must be safe-

[35]Directive (EU) 2024/2853 of the European Parliament and of the Council of 23 October 2024 on liability for defective products and repealing Council Directive 85/374/EEC (Text with EEA relevance) PE/7/2024/REV/1 OJ L, 2024/2853, 18.11.2024, ELI: http://data.europa.eu/eli/dir/2024/2853/oj. Another tool could be available, as the EU Commission had proposed a directive to harmonize the non contractual aspects of AI liability (Proposal for a DIRECTIVE OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL on adapting non-contractual civil liability rules to artificial intelligence (AI Liability Directive) COM/2022/496 final)
. Unfortunately, the proposal has been withdrawn in February 2025 due to an impossibility to reach an agreement among the Union Members. See here for more info https://eur-lex.europa.eu/legal-content/EN/HIS/?uri=celex:52022PC0496 (Last accessed 02/04/2026). That is why this document is not included in the present analysis.

[36] Art 53(1)(a) AIA.

[37] These duties have been exemplified in the following guidelines COMMUNICATION FROM THE COMMISSION Commission Guidelines on the scope of the obligations for providers of general-purpose AI models established by Regulation (EU) 2024/1689 (AI Act).

guarded. The precited documentation must also (i) equip downstream providers with a clear understanding of the model's capabilities and limitations sufficient to ensure their own regulatory compliance, and (ii) include, at a minimum, the elements specified in Annex XII.[38] Thirdly Article 53(1)(c) of the AI Act requires providers of general-purpose AI models to establish and implement a policy ensuring compliance with Union copyright and related-rights law. This policy must specifically enable the identification and observance of rights reservations made pursuant to Article 4(3) of Directive (EU) 2019/790, including through the use of state-of-the-art technological tools capable of detecting such reservations.[39] This refers to the text and data mining exemptions for commercial entities. It might be quite difficult to implement as some states are more favourable to data mining unless the rights are expressly reserved, such as Italy, whereas others are more protective of the content creators' copyright, such as France.[40] One of the most important obligations to implement would be the following one: to "draw up and make publicly available a sufficiently detailed summary about the content used for training of the general-purpose AI model". This will be done in accordance with a template provided by the AI Office. For whoever is thinking of creating an AI companion it might be worthwhile to point out the exemption for the previous obligations for whoever releases their GPAI models under a free and open-source license.[41] Fourthly, an obligation that covers both open-source and not GPAIs there is a duty of cooperation with the EU authorities[42]. This is also true for non-EU based AI providers thanks to the obligations of having an authorised representative.[43] Finally, Art. 53(4) gives the possibility of relying on codes of practices. In particular, since July 2025 it is possible to adhere to Commission's promoted codes of practice for GPAI for three main themes: transparency, copyright and safety and security.[44] They are not mandatory, but it is implied that abiding by them will give a proof of respect of "contractual fairness" or "good faith"[45] according to the different EU legal traditions. That would mean that one could presuppose that the GPAI, the AI companion in this case, was produced with fairness.[46] They are a concrete example of the codes of practice mentioned at Article 57 AIA.

d) The systemic risk GPAIs regime

[38] Art. 53 (1)(b) AIA.

[39] Art. 53 (1)(c) AIA.

[40] For a comparative analysis of AI and copyright please refer to Henri Capitant, Journées Sud – Coréeennes rapports nationaux "Intelligence Artificielle et droit d'auteur". Association Henri Capitant, *Journées sud-coréennes. L'intelligence artificielle*, https://www.henricapitant.org/ (Last accessed: 02/04/2026).

[41] Art. 53(2) AIA.

[42] Art 53(3) AIA.

[43] Art 54 AIA.

[44] European Commission, *The General-Purpose AI Code of Practice*, https://digital-strategy.ec.europa.eu/en/policies/contents-code-gpai (Last accessed: 02/04/2026).

[45] For the Common law tradition see, *ex multis*, J. BEATSON, D. FRIEDMAN, *Good Faith and Fault in Contract Law*, Oxford, 1997; S. WEATHERHILL, *Contract Law of the Internal Market*, Cambridge, 2016; For the continental law tradition, and Italy in particular, see *ex multis*, C.M. BIANCA, *Dell'adempimento delle obbligazioni*, Bologna, 1967; L. LAMBO, *Obblighi di protezione*, Padova, 2007, C. CASTRONOVO, *La responsabilità civile*, Milano, 2018.

[46] The principle of Fairness in EU law depends from specific fields of law, but it has been specifically studied in terms of competition policy and data protection, subject matters that are applicable as well to AI companions. D. GERARD, *Fairness in EU Competition Policy: Significance and Implications*, in *Journal of European Competition Law & Practice*, 9(4), 2018, 211-212; D. CLIFFORD, J. AUSLOOS, *Data Protection and the Role of Fairness*, in *Yearbook of European Law*, 37, 2018, 130-187.

Alternatively, if the GPAI is defined as implying a systemic risk, there is a further set of obligations to comply with. To start, the notion of systemic risk is defined at article 3(65) and is and better specified in recitals from 97 to 173 AIA as "risk that is specific to the high-impact capabilities of general-purpose AI models, having a significant impact on the Union market due to their reach, or due to actual or reasonably foreseeable negative effects on public health, safety, public security, fundamental rights, or the society as a whole, that can be propagated at scale across the value chain". By using this definition and combining it with the one of GPAI, one can wonder: if teenagers in the EU use a certain model of AI Companions and there is an increased number of depression cases, or, even worse, suicides, could GPAIs such as AI Companions assessed as having a systemic risk? According to Article 51 AIA, the GPAI either has to have "high impact capabilities"[47] or if, according to the parameters established in Annex XIII, the Commission, either ex-officio or because of a qualified alert of a scientific panel[48]. At the moment, AI Companions or GPAIs which are used for companionship – but are not marketed as medical devices- are not mentioned directly by the precited annex but they could be included. In fact, if one analyses the criteria set by the Annex there are two that might be applied in the case of AI Companions. Letter d) highlights the multimodality character of the model which is connected to the fact that some AI Companions have a "virtual personification" using images such as with Character AI, and voice effects, or soon could. Moreover, letter g) of the Annex mentions the number of users as a possible an indicator of systemic risk. If a specific AI companion provider acquires a dominant or quasi dominant position on the market under the meaning of Article 102 TFEU and its users are mostly a vulnerable group, such as teenagers, that could be an element that could make the EU Commission consider labelling an AI companion a systemic risk GPAI model. Nevertheless, it exists a (rebuttable) presumption for a GPAI to entail systemic risk. That is only whether the cumulative amount of computation used for its training measured in Floating Point of Operations (FLOPs) is "greater than 10^{25}FLOPS".[49] Nevertheless, Article 51(3) leaves the Commission the possibility to adopt delegated acts to either change the list of hypotheses of systemic risk but also to update what is considered to be a high impact capability GPAI. What further obligations having a systemic risk GPAI entail compared to non-systemic risk GPAIs? Just the ones of Article 55 which are the following. Article 55 sets out additional obligations for providers of general-purpose, supplementing the duties contained in Articles 53 and 54. Providers must carry out state-of-the-art evaluations of their models, including structured and adversarial testing, in order to identify and mitigate systemic risks arising from their development, market placement, or use.[50] They are also required to assess and reduce systemic risks[51] at the level of the European Union, maintain continuous oversight of such risks, and promptly document and report serious incidents – together with any corrective measures – to the AI Office and, where appropriate, national authorities.[52] Providers must ensure that both the model and its underlying physical infrastructure benefit from an adequate level of cybersecurity protection.[53] To demonstrate compliance, providers may rely on codes of practice established under Article 56 until harmonised standards are adopted. Conformity with such

[47] Article 51(1) (a) AIA.
[48] Article 51(1)(b) AIA.
[49] Article 51(2) AIA. It is possible for the AI provider of the GPAI model to rebut this presumption.
[50] Art 55(1)(a) AIA.
[51] Art 55(1)(b) AIA.
[52] Art 55(1)(c) AIA.
[53] Art 55(1)(d) AIA.

standards creates a presumption of compliance, whereas providers who do not rely on them must demonstrate alternative adequate measures to the European Commission. All information generated under Article 55, including sensitive commercial material, must be handled in accordance with the confidentiality requirements set out in Article 78.

e) AI Companions as medical GPAIs

Whether AI Companions used for mental health qualify as medical devices depends entirely on their 'intended use' under EU law. Because classification as a medical device requires compliance with the stringent MDR, AI providers/manufacturers often exploit this grey area by deliberately marketing their products as non-medical wellness tools. To provide an overview of why an AI companion developer might prefer the traditional 'consumer product' pathway, it is instructive to examine how these systems would be classified if marketed as medical devices. Firstly, according to the combinations of Article 2(1) and Annex VIII, rule 11, they would fall into the old definition of Software as a Medical Device (SaMD), and, specifically in the subsection of Medical AI (MDAI) according to the recent Medical Devices Coordination Group (MDCG) and the AI Board (AIB) FAQs.[54] This subdivision implies that if a software with a declared medical function set in Article 2, and whose manufacturer wants to market it as a medical device, is a SaMD. Moreover, if the software used, also falls within the definition of AI system (or in this case, GPAI model) then it is a MDAI. As well as "traditional SaMD" It can be placed in all risk categories, meaning from I (low risk) to III (highest risk level).[55] This applies to AI as well, including AI systems and GPAIs including AI Companions. It must be mentioned that, according to rule n. 3 Annex VII MDR the class of the SaMD (including the MDAI) is determined on its own if the software is a stand-alone program. On the contrary, if it is incorporated within a medical device hardware part as a component, then it will follow the risk class of the hardware part. If we come back to the AI Companions, they will most certainly fall within the former category. As far as the risk assessment, one must check the effects that might have on patients. If death cannot be excluded, then they will fall within class III, the highest risk class and be subjected not only to a conformity procedure assessed by a third-party assessment body called Notified Body (NB) but also to compelling post-market compliance duties.[56] However, how do the MDR rules on MDAI combine with the ones of the AIA? If the software used falls within one of the categories of the AI Act (included the one of GPAI) then there must be some kind of harmonisation. In principle, Article 8(2) AIA sets that if a harmonisation discipline is included in the Annex I section A (and the MDR, concerning medical software is) then the AI provider-MDR manufacturer can avoid duplications by keeping as a baseline the old conformity procedure (in this case, the MDR one) and to add the AI act rules that are pertinent to the case (AKA chapter III on high risk AI systems, not even the rules concerning "simple" GPAIs). This

[54] European Commission – Medical Device Coordination Group (MDCG), *MDCG 2025-6 FAQ on the interplay between the Medical Devices Regulation and the In Vitro Diagnostic Medical Devices Regulation*, https://health.ec.europa.eu/latest-updates/mdcg-2025-6-faq-interplay-between-medical-devices-regulation-vitro-diagnostic-medical-devices-2025-06-19_en (Last accessed: 02/04/2026).

[55] F. GENNARI, *O Complementarity, Where Art Thou? Wading through the Medical Device Regulation and the AI Act Compliance: The Case of Software as a Medical Device. A Primer*, in *BioLaw Journal – Rivista di BioDiritto*, 3, 2024, 411-453; S. PALMIERI, *The Renewed EU Legal Framework for Medical AI*, in *European Journal of Law and Technology*, 15(3), 2024.

[56] For example, for the III MD risk class, the Post Market Surveillance Report, PSUR, must be submitted by the manufacturer to the Notified Body every year. See Art 86 MDR.

process of clarification is still ongoing, but a first set of answers (even though quite a general one) was provided by the pre-cited FAQs in June 2025. However, in the meantime, the Digital Omnibus and the MDR and IVDR simplification proposals have been presented by the EU commission at the end of the last year.[57] In particular, in the simplification of the MDR document, the most relevant change concerning MDAI (a category in which AI Companions would belong) is the shift from Annex I list A to list B of the MDR. Were this modification approved, it would likely mean that AI providers/MDR manufacturers of AI Companions as medical devices would have to carry out two different compliance procedures, which will likely overlap in several points. Some of the most evident examples of the similarity of the AI and MDR conformity procedures are the following: the creation of a quality management system and data curation to support the medical safety and efficiency of the AI companion to give some examples. It is difficult to see an MDR simplification if there is a clear multiplication of procedures whenever AI is used as a medical device, unless the MDR is modified in order to contain a specific discipline for AI as a medical device.

5. Regulation of AI Companions in the US

AI Companions have raised attention also in US legal literature. Although the regulatory architecture differs significantly, both systems ultimately depend on manufacturers' descriptions of the product's purpose. Consequently, both jurisdictions face the same challenge: systems that function as quasi-therapeutic tools may remain outside medical-device regulation because their providers deliberately frame them as wellness, lifestyle, or companionship products. Under the Federal Food, Drug, and Cosmetic Act (FDCA),[58] software qualifies as a medical device when it is intended for use in the diagnosis, cure, mitigation, treatment, or prevention of disease. As in the European Union, the notion of intended use – derived from manufacturers' labelling, promotional materials, and overall presentation – functions as the primary trigger for regulatory oversight. This creates strong incentives for developers to frame their systems as general wellness or lifestyle tools, even where their functionalities may approach health-related uses. This phenomenon has been described in the digital health literature as "skating the regulatory line", whereby products are deliberately positioned at the boundary of medical-device regulation without formally entering it.[59]

The 21st Century Cures Act reinforces this dynamic by explicitly excluding certain categories of software, including specific clinical decision-support functions, from the statutory definition of a medical device.[60]

[57] Proposal for a REGULATION OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL amending Regulations (EU) 2016/679, (EU) 2018/1724, (EU) 2018/1725, (EU) 2023/2854 and Directives 2002/58/EC, (EU) 2022/2555 and (EU) 2022/2557 as regards the simplification of the digital legislative framework, and repealing Regulations (EU) 2018/1807, (EU) 2019/1150, (EU) 2022/868, and Directive (EU) 2019/1024 (Digital Omnibus) COM/2025/837 final; European Commission, *Proposal for a Regulation to Simplify Rules for Medical and In Vitro Diagnostic Devices*, https://health.ec.europa.eu/publications/proposal-regulation-simplify-rules-medical-and-vitro-diagnostic-devices_en (Last accessed: 02/04/2026).

[58] U.S. Food and Drug Administration, Federal Food, Drug, and Cosmetic Act (FD&C Act), https://www.fda.gov/regulatory-information/laws-enforced-fda/federal-food-drug-and-cosmetic-act-fdc-act (Last accessed: 02/04/2026).

[59] D.A. SIMON, C. SHACHAR, I.G. COHEN, *Skating the Line between General Wellness Products and Regulated Devices: Strategies and Implications*, in *Journal of Law and the Biosciences*, 9(2), 2022, art. lsac015.

[60] U.S. Congress, *Public Law No. 114-255 (21st Century Cures Act),* https://www.congreff.gov/114/plaws/publ255/PLAW-114publ255.pdf (Last accessed: 02/04/2026).

While the Act was designed to promote innovation, it further entrenches a regulatory model in which classification depends heavily on how software functions and claims are articulated. As a result, many conversational AI systems operate in close proximity to healthcare functions without triggering medical-device requirements. From a Federal regulatory perspective, there is not a binding document but just a series of recommendations over the years concerning AI and Machine Learning techniques. Only in 2025, the FDA published a draft guidance on the use of AI in medicine[61] but it does not make the case for AI Companions. It is still difficult to evaluate whether this will happen soon, but it seems likely. The FDA has recently designed a "breakthrough status" to an LLM-based tool for surgical patients. If this medical device will be finally authorised, it will provide the "mold" for the FDA to evaluate other LLM based medical devices, AI Companions included.[62]

The result, on both sides of the Atlantic, is a regulatory landscape in which functionally similar systems may fall on different sides of the regulatory divide based primarily on their declared intended use. AI Companions thus occupy an unstable position at the intersection of consumer software, wellness technologies, and potential medical interventions. Only a limited number of systems approaching this category have undergone clinical validation. One often-cited example is Woebot, a conversational mental-health application marketed in the United States, although such cases remain relatively rare and have faced challenges in sustaining long-term commercial deployment.[63]

With regard to discussing AI Companions not based on function but based on the fact that they are containing an AI component, different are the considerations to be made. While the United States does not yet have a comprehensive federal framework governing artificial intelligence, regulatory activity is increasingly emerging at the state level, often through sector-specific legislation. This decentralized landscape means that AI Companions – particularly those used for emotional support or mental-health-adjacent functions – may fall within the scope of different state laws depending on how their functions are framed and deployed. Two recent legislative initiatives are particularly illustrative in this regard: California's SB 243, which directly targets "companion chatbots,"[64] and Illinois' Wellness and Oversight for Psychological Resources Act (WOPR Act)[65], which focuses on AI-based mental health applications.

---

[61] U.S. Food and Drug Administration, *Artificial Intelligence-Enabled Device Software Functions: Lifecycle Management and Marketing Submission Recommendations. Draft Guidance for Industry and Food and Drug Administration Staff*, January 2025, https://www.fda.gov/regulatory-information/search-fda-guidance-documents/artificial-intelligence-enabled-device-software-functions-lifecycle-management-and-marketing (Last accessed: 02/04/2026).

[62] K. PALMER, *FDA Offers Clues to AI Regulation with RecovryAI Designation*, in *STAT News*, 03/03/2026, www.statnews.com (Last accessed: 02/04/2026).For an analysis updated to 2023 concerning the AI/ML approved medical devices see G. JOSHI, A. JAIN, S. REDDY ARAVEETI, S. ADHIKARI, H. GARG, M. BHANDARI, *FDA-Approved Artificial Intelligence and Machine Learning (AI/ML)-Enabled Medical Devices: An Updated Landscape*, in *Electronics*, 13(3), 2024, art. 498. On the specifics of US medical devices approval procedures see P.J. ZETTLER, E. LIETZAN, *Regulating Medical Devices in the United States*, in D. ORENTLICHER, T. HERVEY (eds.), The *Oxford Handbook of Comparative Health Law*, Oxford, 2020, 755-777.

[63] B. THOMPSON, *Therapy Shock: Pioneering Mental Health Chatbot Is Shutting Down*, in *New Atlas*, 28/05/2025, https://newatlas.com/mental-health/woebot-closing/ (Last accessed: 02/04/2026).

[64] California Legislature, *Senate Bill No. 243 (2025–2026)*, https://leginfo.legislature.ca.gov/faces/billNavClient.xhtml?bill_id=202520260SB243 (Last accessed: 02/04/2026).

[65] State of Illinois, [Press release], https://www.illinois.gov/news/release.html?releaseid=31573 (Last accessed: 02/04/2026). During the submission phase, it is provided only the press release as it appears that the law text is not available anymore on the Illinois General Assembly website.

California's SB 243 is notable because it is one of the first statutes specifically designed to regulate so-called "companion chatbots," defined as AI systems intended to simulate human companionship or emotional relationships. The law focuses primarily on transparency and harm-mitigation obligations rather than on classifying such systems as medical technologies. Its regulatory approach can be broadly understood through four pillars.

First, the statute introduces a human-versus-AI disclosure requirement. When a reasonable user could believe that they are interacting with a human being, operators must provide a clear and conspicuous notice that the interlocutor is an artificial intelligence system. The objective is to reduce the risk of emotional deception, which could be particularly problematic in situations involving vulnerable users seeking support or companionship. Second, SB 243 requires operators to develop a suicide and self-harm prevention protocol. Companies must establish and publicly describe procedures addressing the possibility that the system might generate or respond to suicidal ideation or self-harm content. This includes mechanisms for directing users to crisis hotlines or text-based support services when relevant signals arise during the interaction. Third, the law introduces heightened safeguards for minors. When an operator knows that a user is under the age of eighteen, it must ensure disclosure that the chatbot is an AI system, provide periodic reminders of this fact during extended interactions, and take reasonable measures to prevent the delivery of sexually explicit material. Finally, the statute establishes reporting and enforcement mechanisms, including annual reporting obligations beginning in 2027 and civil liability with statutory damages for violations. Importantly, SB 243 is largely procedural rather than performance based. The law requires operators to develop and publish risk-management protocols but does not impose concrete effectiveness standards regarding how systems must handle mental-health-related interactions. In addition, several of its strongest protections hinge on whether the operator knows that a user is a minor, potentially creating incentives to avoid active age verification or inference. These design choices illustrate the regulatory balancing act between protecting vulnerable users and avoiding overly burdensome compliance obligations for emerging technologies.

Illinois has adopted a different regulatory strategy. Rather than targeting AI Companions as such, the Wellness and Oversight for Psychological Resources Act (WOPR Act) focuses on AI-powered mental health and wellness applications, including those providing psychological advice, therapeutic guidance, or mental health support through automated systems. The law was introduced in response to the rapid proliferation of digital mental health applications using artificial intelligence to simulate elements of counselling or therapeutic interaction. Under the WOPR framework, certain AI-driven mental health tools must comply with transparency and oversight obligations, including disclosure that the service is not a licensed mental health professional and the provision of clear information about the nature and limits of the system's capabilities. The law also emphasizes safeguards for vulnerable users and mechanisms to redirect individuals experiencing mental health crises toward appropriate professional resources.

For AI Companions, the applicability of the Illinois law therefore depends largely on how the system's functions are framed and marketed. A chatbot designed primarily for companionship, entertainment, or general emotional reflection may fall outside the scope of the statute if it does not present itself as offering mental health advice or psychological treatment. Conversely, if an AI companion is promoted as providing therapeutic support, counselling-like interactions, or structured mental health guidance, it may fall within the regulatory perimeter established by the WOPR Act. This illustrates a broader feature of the

U.S. regulatory landscape: classification often depends less on the underlying technical architecture of the system than on how its intended function is articulated to users. As with medical device regulation at the federal level, developers may therefore strategically frame AI Companions as tools for well-being or emotional reflection rather than mental health treatment in order to avoid triggering stricter regulatory oversight.

6. Towards a Functional Approach to Regulatory Qualification

The preceding analysis highlights a structural limitation common to both the European and United States regulatory frameworks: the centrality of the manufacturer's declared intended use in determining whether a system falls within the scope of medical-device regulation or remains classified as a general consumer product. While this criterion serves an important role in ensuring legal certainty, it becomes increasingly inadequate when applied to AI Companions whose interactional design and foreseeable use enable them to perform functions that are substantively adjacent to mental-health support.

AI companions are rarely marketed as diagnostic or therapeutic tools. Instead, they are typically framed as instruments for companionship, emotional reflection, or general well-being. However, this formal positioning often diverges from the functional reality of the interaction. Many of these systems engage users in sustained, personalized, and emotionally responsive exchanges; respond to expressions of distress, anxiety, or loneliness; and offer reassurance, coping suggestions, or forms of cognitive reframing. In contexts of unmet mental-health needs, such systems may become a regular and, at times, primary point of reliance. In these circumstances, limiting regulatory qualification to declared intended use risks overlooking the actual role these systems play in users' lives. We also recognize that it is possible that AI Companions which truly do not have any medical function, but just a not better identifiable need for companionships (such as, in different ways avatars) should not be constrained to market their AI companion as a medical device. Still, it is very important, especially if no medical function is thought to be the rationale of the AI companion, that the AI providers are made accountable and liable of the psychological consequences that their services or products, depending on the reference regulatory framework.

For this reason, this article argues for a functional approach to regulatory qualification, under which classification does not depend exclusively on how a system is described by its provider, but also on what it does in practice, how it is reasonably expected to be used, and the risks that arise from that use. This approach does not seek to displace intended use altogether. Rather, it aims to recalibrate its role, ensuring that it does not operate as an absolute determinant where the functional characteristics of the system clearly point toward health-related uses.

Such a shift is justified on several grounds. First, it aligns more closely with the risk-based logic underpinning both medical-device regulation and contemporary AI governance. Regulatory frameworks are designed to ensure that technologies capable of affecting health, safety, or fundamental rights are subject to proportionate safeguards. Allowing systems to escape scrutiny solely because they are framed as non-medical, despite performing health-relevant functions, undermines this objective. This concern is particularly acute in contexts characterised by informational asymmetries and user vulnerability, where individuals may not distinguish between a formally non-medical tool and a system that effectively provides mental-health support.

Second, a functional approach better reflects the relational nature of AI Companions. Unlike traditional software, these systems are designed to simulate social interaction, maintain conversational continuity, and adapt to users' emotional states. Their significance lies not only in their technical capabilities but in how they are experienced: as responsive interlocutors capable of eliciting disclosure, fostering emotional dependence, and shaping coping practices. Where such dynamics are present, it becomes increasingly artificial to treat these systems as ordinary consumer products devoid of health-related implications.

Third, a function-sensitive framework would reduce opportunities for regulatory arbitrage. Both EU and U.S. regimes currently incentivize developers to position their products as close as possible to therapeutic functionality without triggering medical-device obligations–a practice often described as "skating the regulatory line." The result is a misalignment between regulatory burden and actual risk, whereby providers benefit from reduced compliance obligations while users bear the consequences of uncertainty, error, or harm. A functional approach would help correct this imbalance by tying regulatory scrutiny more closely to real-world effects rather than formal classification alone.

The preceding considerations can be translated into a more operational set of indicators capable of guiding regulatory assessment.

To operationalize the functional approach outlined above, a structured set of indicators can assist in determining when an AI companion should trigger health-related regulatory scrutiny.

For regulatory purposes, an AI companion should be presumed to warrant health-related qualification, and, where appropriate, assessment under medical-device frameworks, where a combination of the following elements is present:

(1) Mental-health–relevant interaction: the system responds to user inputs concerning emotional distress, anxiety, depression, trauma, or self-harm with outputs that go beyond generic conversation and provide tailored guidance, reassurance, or behavioural suggestions;

(2) Therapeutic-like engagement: the system engages in sustained dialogue that incorporates elements functionally analogous to counselling or psychological support, including reframing, coping strategies, or structured emotional assistance;

(3) Personalisation and continuity: the system adapts its responses based on prior interactions, user-specific data, or inferred psychological states, thereby fostering an ongoing and relational form of engagement;

(4) Foreseeable reliance in lieu of care: it is reasonably foreseeable that users may rely on the system as a substitute for, or complement to, professional mental-health support, particularly in contexts of limited access to care;

(5) Vulnerability-sensitive deployment: the system is used, or is likely to be used, by individuals or groups in situations of heightened vulnerability, including minors or persons experiencing psychological distress.

These elements should not be understood as rigid cumulative criteria, but as converging indicators supporting a functional assessment. Their presence signals that the system operates within a health-sensitive domain and should not be excluded from appropriate regulatory scrutiny solely on the basis of its formally declared intended use.

These indicators suggest that regulatory qualification should be informed by the interaction between design, function, context, and foreseeable effects, rather than by formal categorisation alone. This does not imply that all AI Companions should be treated as medical devices. Rather, it establishes that where sys-

tems cross a threshold of functional proximity to mental-health intervention, they should not be permitted to remain entirely outside health-related regulatory oversight.

From an EU perspective, such an approach could be implemented through a more substantive interpretation of existing concepts, particularly the notion of intended purpose under the Medical Devices Regulation, informed by foreseeable use and functional characteristics. Moreover, as the primacy of the manufacturer's intended purpose was first decided in the pre-cited EU Cases, a preliminary question could be raised by a national judge to the CJEU in ambiguous cases. Hopefully the CJEU could employ the above-cited criteria for functional regulatory classification. That would ensure that national courts would follow this interpretation in the light of the primacy principle of EU law[66]. In the meantime, especially for AI Companions whose medical function might be only hinted but not declared, the Commission could foresee to declare AI companion as entailing a systemic risk, given that categories of vulnerable people, such as teenagers, have proved to be harmed by the use of AI Companions at least a further means of protection. By reading the AI Act, there are not more obligations for systemic risk GPAIs than for GPAIs that do not have a specific risk. Complementary instruments, including consumer protection law, could address residual risks where systems remain formally outside the medical-device framework but nonetheless influence user behaviour in psychologically sensitive contexts. In the United States, a similar shift would require greater attention to functional outputs and interactional dynamics, rather than reliance on formal claims alone, particularly in areas where products encourage emotional or mental-health–related reliance.

Ultimately, a functional approach to regulatory qualification is not a call for indiscriminate expansion of medical-device regulation. It is a call for doctrinal coherence. Where AI Companions operate in ways that users reasonably experience as forms of mental-health support, regulatory classification should reflect that reality. Aligning legal qualification with functional impact is essential to ensuring that the level of oversight corresponds to the risks these systems generate.

However, some encouraging hints might come from both sides of the Atlantic which directly challenge the status quo of the intended purpose/ use supremacy with similar, although different, digital products. The Commission's take on TikTok on addictive design[67] under the Digital Services Act and the recent jury verdicts in New York and New Mexico which enforced Meta and other social media platforms[68] pay compensation because of the effects of their addictive design on teenagers or young adults seem to imply that the intended purpose of the manufacturer cannot be a legal shield to avoid accountability for (serious) damage to users. This makes the functional regulatory approach all the more relevant in both cases AI companions have or not de facto a medical function as a new notion of digital well-being is starting to

[66] Among the many valuable contributions on the definition of the EU law primacy status, see B. DE WITTE, *Direct Effect, Primacy, and the Nature of the Legal Order*, in P. CRAIG, G. DE BÚRCA (eds.), *The Evolution of EU Law*, Oxford, 2021, 187–227.

[67] European Commission, *Commission Preliminarily Finds TikTok's Addictive Design to Breach the Digital Services Act*, https://digital-strategy.ec.europa.eu/en/news/commission-preliminarily-finds-tiktoks-addictive-design-breach-digital-services-act (Last accessed: 02/04/2026).

[68] M. ISAAC, *A Jury Finds Social Media Giants Liable for Creating Addictive Products*, in *The New York Times*, 25/03/2026, https://www.nytimes.com/2026/03/25/technology/social-media-trial-verdict.html (Last accessed: 02/04/2026).

emerge.[69] For the AI companion provider, it will not likely possible to opt for the "simpler route" of a "consumer product" marketing without being accountable or liable, if damage is caused (especially) to vulnerable users.

7. Conclusions

AI Companions expose a fundamental tension at the heart of contemporary digital health regulation. Both in the European Union and in the United States, legal qualification continues to depend primarily on the notion of intended use as articulated by the manufacturer. While this approach offers clarity and flexibility, it becomes increasingly inadequate in the face of technologies whose functions are fluid, relational, and shaped as much by user interaction as by initial design.

As this article has shown, AI Companions frequently operate in a space where formal classification and functional reality diverge. Systems that are framed as tools for companionship or well-being may, in practice, perform roles that resemble mental-health support, particularly in contexts of unmet care needs and user vulnerability. This functional–formal mismatch allows technologies capable of influencing users' psychological well-being to remain outside regulatory frameworks designed to ensure safety, accountability, and effectiveness.

The consequence is not merely conceptual inconsistency, but a regulatory gap with tangible implications. Users may rely on these systems in situations of distress without the protections associated with health-related interventions, while providers may avoid the obligations that would otherwise attach to technologies operating in comparable domains. The resulting asymmetry risks undermining both user protection and the integrity of the regulatory system.

Addressing this challenge requires a shift from a purely formal understanding of intended use toward a more functional approach to regulatory qualification. Such an approach would not eliminate the role of intended use but would situate it within a broader assessment that considers interactional design, foreseeable reliance, and the context in which systems are deployed. By doing so, it would better align regulatory oversight with the actual risks posed by AI Companions.

Ultimately, the question is not whether AI Companions should be categorically treated as medical devices, but whether regulatory frameworks can adequately capture technologies that blur the boundary between social interaction and health-related support. Ensuring that legal qualification reflects functional reality is a necessary step toward developing a coherent and effective governance model for AI systems operating in this increasingly sensitive domain.

[69] N. BARANOWSKA, G. MALGIERI, *Towards a Legal Concept of Digital Well-Being: Insights from the EU Commission's Preliminary Findings on TikTok's Addictive Design*, in *Verfassungsblog*, 26/03/2026, https://verfassungsblog.de/tik-tok-design-eu/ (Last accessed: 02/04/2026).