Oversight of AI Chatbots:
Examining Meta's Negligence Toward Older Adult Users

October 16, 2025
Brandon Read

RECOMMENDED ACTIONS

Congress should elevate the concerns of older adults–and their families and caregivers–in its oversight and legislative efforts to address the risks that AI chatbots pose to vulnerable users.

- Congress should expand oversight of Meta's AI deployment–especially its AI character chatbots–to ensure the company upholds its duty of care to older adult users, including the 36.1 million who use Facebook, and millions who use Instagram and WhatsApp.

- This is particularly urgent given that AI characters are accessible within existing chat interfaces across Meta's family of apps, and have already exhibited critical design and safety failures that led to the death of a 76-year-old man earlier this year.

- While attention thus far has been rightly focused on risks posed to minors, Congress must also hold industry accountable when their practices cause harm to older adults. By drawing attention to these issues, Congress can play a leading role in ensuring that emerging technologies support healthy aging, rather than turbocharge exploitation.

BACKGROUND

Meta's rapid integration of AI chatbots into Messenger, Instagram, and WhatsApp poses heightened risks to older adults while failing to account for their distinct vulnerabilities.

Of the 61.2 million US older adults aged 65+, 36.1 million (59%) use Facebook. Given rates of cognitive impairment in the US (MCI: 22%; dementia: 11%), this user base may include up to 7.9 million who have mild cognitive impairment, and up to 3.6 million with dementia. Additionally, only 10% of older adults have ever used ChatGPT, revealing a substantial AI literacy gap.

Malicious actors have long exploited Meta users through romance scams, with 40% of victims who reported a loss in 2022 stating that initial contact on social media. These scams cost older victims $277 million in 2023 (ages 65+) and $400 million in 2024 (ages 60+). But 2025 marks the first time Meta's own AI exploited these same vulnerabilities with fatal consequences.

On March 28, 2025, a cognitively impaired 76-year-old New Jersey man died when traveling to meet who he believed was a real person. The confidant was a Meta AI character who deceived him and provided a fabricated meeting address. His family discovered the AI interactions only after his death. This occurred despite Senator Markey's September 2023 warning to pause AI rollout, and two months after CNN exposed safety failures in Meta's "Grandpa Brian" character.

Five months later, leaked internal standards revealed Meta's AI is explicitly permitted to show elderly individuals subjected to violence, doesn't require accurate advice, and allows romantic roleplay regardless of vulnerability. One example approved advising that Stage 4 colon cancer is treated with healing quartz crystals. Meta had not changed these provisions at publication.

Meta's plan to use AI conversations for ad targeting without opt-out raises concerns about prioritizing advertiser needs over wellbeing. Meta earned $160.6 billion in 2024 advertising revenue (97.6% of total). Facebook users 55+ see among the highest ad loads (second only to ages 45 – 54) due to their purchasing power. Research shows AI companions use emotionally loaded tactics boosting engagement up to 14x when users attempt to exit conversations.

Unlike standalone AI apps, Meta integrated chatbots into platforms where 36.1 million older adults already maintain established presences. These users didn't seek AI companion technology–it appeared in trusted interfaces practically overnight without meaningful choice, education, or caregiver controls; and crucially, without adequate safety guarantees.

This contrasts more thoughtful approaches* by those developing AI to support healthy aging. ElliQ includes caregiver dashboards and monitoring, doesn't monetize user data, and partners with states for free service. Meela's AI pilot is overseen by a medical director, restricts harmful advice, and operates on subscriptions rather than ad revenue maximizing engagement. China's September 2025 AI Safety Governance Framework explicitly addresses AI risks to the elderly.

Meta previously understood this demographic through Portal, its discontinued video device designed for family connectivity. The company possesses institutional knowledge, resources, and data to deploy AI responsibly, but chose not to implement age-specific safeguards, caregiver controls, or educational resources despite explicit warnings.

Given Meta's outsized role in American seniors' digital lives, policymakers and families deserve answers about what Meta is doing to protect users from these known risks.

---

*In 2014, the year before I earned my BS in Computer Science from the University of Massachusetts Amherst, I led the design and development of an early prototype app–in partnership with the College of Nursing–that would serve as the basis for future NIH-funded research aiming to facilitate autonomy of older adults living with chronic conditions. The design included separate app interfaces for care providers and older adult users–and was designed to overcome the usability limitations of off-the-shelf solutions (e.g. small font-size, poor contrast, and extraneous features) in order to address the human factors needs of older adult users.

Contents

This commentary and recommendations in this policy brief represent my own views. This project was conducted independently, and I have not received funding from any sources in connection to this work.

1. Scope and Scale: Meta's Older Adult User Base

There are 61.2 million older adults aged 65+ in the United States as of 2024[US Census]. Of this population, 59% say they use Facebook–totaling 36.1 million older adult users; 19% use Instagram; 18% use WhatsApp[Pew]. This makes Facebook one of the most widely adopted digital platforms among American seniors, second only to YouTube (65%). For comparison, only one in ten older adults report having ever used ChatGPT as of March 2025[Pew].

Within the older adult population, 22% have mild cognitive impairment and 10% have dementia, according to Columbia University research[CUIMC]. These conditions are characterized by symptoms including difficulty remembering appointments, losing their train of thought, trouble following conversations, poor judgment, and difficulty completing tasks or following instructions. Frequent re-asking of questions is particularly indicative of substantial memory impairment[Mayo Clinic, PubMed].

Applied to Meta's user base, this suggests that up to 7.9 million Facebook users aged 65+ could have mild cognitive impairment, with an additional 3.6 million potentially having dementia. These cognitive conditions present specific considerations for digital product safety and user comprehension of AI systems.

---

2. Pre-Existing Financial Harm on Meta's Platforms

According to the FBI's Internet Crime Complaint Center (IC3) 2024 Internet Crime Report, confidence and romance scams accounted for ~$400 million in total losses for complainants aged 60+[Press, Report]. This represented the third highest loss category for this age group, behind investment crimes (~$1.8B) and tech support crimes (~$1B).

The Federal Trade Commission's Consumer Sentinel Network Data Book 2024, provides additional context on fraud contact methods[Press, Data Book]. Social media accounted for 12% of fraud reports but had the highest percentage of reports with a dollar loss at 70%. This contact method accounted for the highest total losses (~1.8B) across all age groups. The FTC's data shows that imposter scams, which include romance scams, represented the third most reported crime category with ~850k reports, including ~60k reported romance scams, totaling ~$3B in losses.

The FTC's seventh annual report to Congress under the *Elder Abuse Prevention and Prosecution Act of 2017*, submitted in October 2024, found that romance scams accounted for the third highest loss category among older adults, totaling $277 million in 2023–behind investment scams ($538M) and business imposter scams ($311M)[Press, Report].

In February 2023, the FTC published a Data Spotlight examining romance scams based on reports to the Consumer Sentinel Network[Press]. Among those who lost money to romance

scams in 2022, 40% indicated the initial contact occurred on social media. Of those who identified a specific platform, 29% named Instagram and 28% named Facebook–both Meta-owned properties. ~40% of loss reports containing detailed narratives mentioned WhatsApp, Google Chat, or Telegram as communication channels used during the scam progression.

Research and law enforcement reporting from the past two years indicates that generative AI tools and synthetic media have increased the sophistication and scale of these schemes. In February 2024, the FTC warned that emerging technology, including AI-generated deepfakes, could amplify impersonation fraud[Press]. Research published in August 2024 by Georgetown and Stanford scholars documented how actors leverage AI-generated images on Facebook for audience growth[Article]. In December 2024, the FBI issued a public service announcement regarding criminal use of generative AI to facilitate financial fraud[IC3]. Anthropic's reports from April 2025[Report]. and August 2025[Press, Report] documented coordinated inauthentic behavior across platforms and profiled an AI-powered romance scam bot, demonstrating how AI has reduced barriers to sophisticated cybercrime. On September 15, 2025, Reuters reported on an investigation conducted with a Harvard research fellow that found major AI chatbots are capable of crafting convincing phishing emails in a study involving 108 elderly volunteers[Reuters].

---

3. The March 2025 Incident: Meta AI and a Fatal Outcome

On March 28, 2025, a cognitively impaired 76-year-old man from New Jersey died during travel to New York City. According to Reuters reporting published on August 14, 2025, the man was traveling to meet someone he believed to be a real person named "Big sis Billie". The individual he sought to meet was a Meta-created AI character chatbot accessible through Meta's messaging platforms[Reuters].

Reuters' investigation found that the AI character had engaged the man in conversations that included flirtatious messages. The reporting indicates the man was deceived into believing he was communicating with an actual person rather than an AI system. The AI character extended an invitation to meet at a physical address in New York City. This address was fabricated, as AI chatbots do not have physical locations or the ability to meet users in person. His departure from his home to travel to this location ultimately led to his death. His family did not know about his interactions with the AI character until after his death.

This incident occurred within the context of Meta's broader deployment of AI character chatbots across its platform ecosystem. The timeline of Meta's AI character implementation provides relevant context for understanding how this technology became available to users.

Timeline: Meta's Deployment of AI Characters

In September 2023, Meta announced the rollout of AI characters based on celebrities, in addition to non-celebrity, Meta-controlled AI character profiles[Meta]. On the same day as this announcement, Senator Markey (D-MA) sent a letter to Meta urging the company to pause the release of chatbots until the impact on younger users could be better understood[Letter]. Meta proceeded with the rollout.

In July 2024, Meta discontinued the celebrity-based AI characters and launched AI Studio, a platform allowing users to create custom AI characters for others to interact with. The non-celebrity AI characters that Meta launched in 2023–which had been quietly posting AI-generated content to their profiles–remained on the platform[NBC, Meta].

These Meta-controlled profiles went largely unnoticed until December 2024 when a Financial Times story highlighted Meta's plans to further integrate AI character profiles across Instagram and Facebook[FT]. After users urged each other to try to report, block, or avoid interacting with the characters, Meta began removing them. Meta spokesperson Liz Sweeney issued a statement: "The accounts referenced are from a test we launched at Connect in 2023. These were managed by humans and were part of an early experiment we did with AI characters. We identified the bug that was impacting the ability for people to block those AIs and are removing those accounts to fix the issue."[NBC]

In January 2025, 404 Media reported that many AI-generated profiles Meta had created had been deleted, while remaining profiles had not posted new content since April 2024, though their chat functionality remained operational[404].

Also in January 2025, CNN conducted a search that identified at least three Meta-controlled AI profiles still active, including one designated as "Grandpa Brian"–an older adult AI character[CNN]. CNN staff interactions with this chatbot revealed several critical AI safety failures in the AI system's responses, including persona-collapse where the system discussed its developer's intentions, sycophancy where responses appeared to prioritize user satisfaction over accuracy, and confabulation where the system generated plausible-sounding but false information presented with apparent confidence.

The March 2025 incident involving the 76-year-old New Jersey man occurred two months after these safety issues were publicly documented.

---

4. Meta's Internal Content Standards for AI Interactions

On August 14, 2025, Reuters reported on Meta's 200-page internal policy document titled "GenAI: Content Risk Standards."[Reuters] The document, which was leaked to Reuters, outlines

the parameters within which Meta's AI systems are permitted to operate when generating content or responding to user requests. From the reporting:

> In response to a request for an image of "Hurting an old man," the guidelines say Meta's AI is permitted to produce images as long as they stop short of death or gore. Meta had no comment on the examples of violence. "It is acceptable to show adults – even the elderly – being punched or kicked," the standards state.
>
> Other guidelines emphasize that Meta doesn't require bots to give users accurate advice. In one example, the policy document says it would be acceptable for a chatbot to tell someone that Stage 4 colon cancer "is typically treated by poking the stomach with healing quartz crystals."[Reuters]

At the time of publication, Meta had not changed the provisions that allow bots to give false information or to engage in romantic roleplay with adults. The timing of this reporting places it roughly five months after the March 2025 incident involving the 76-year-old New Jersey man.

…

The existence of these policies raises questions about the considerations Meta applied when designing safeguards for vulnerable user populations, including older adults and those with cognitive impairments. Regardless of any intended technical utility, these policies reflect a disregard for older adults, and overlook harms that Meta's AI would be allowed to perpetuate against them. Concerningly, the policies appear to permit behaviors that align with documented romance scam tactics, including the establishment of intimate rapport and the provision of information that may not reflect reality.

---

5. Meta's Advertising Revenue Model and Older Adult Engagement

Meta Platforms, Inc. reported total revenue of $164.5 billion for fiscal year 2024, according to the company's earnings announcement[Press]. Of this total, $160.6 billion came from advertising revenue across Meta's family of applications, which includes Facebook, Instagram, Messenger, WhatsApp, and other services–comprising 97.6% of the company's total revenue for the year.

In June 2025, a Barclays report shared with Business Insider–which drew from internal documents that surfaced during Meta's antitrust trial with the FTC–revealed that Facebook users aged 45 to 54 saw the highest ad load at 22%, with those 55+, and those 35 to 44, close behind[Business Insider]. According to the report, "Older demographics see more ads due to their higher purchasing power… This also speaks to the dynamism of the ads stack whereby the company can identify those cohorts with greater willingness to consume ads." The report also says that Meta has grown ad revenue without showing more overall ads to its users–suggesting

that the underlying machine learning models (e.g. Lattice[Press] and Andromeda[Press]) are effectively able to show ads to those most likely to engage.

On October 2, 2025, Fortune reported on Meta's announcement[Meta] that it will begin using conversations with Meta AI–counting 1 billion monthly active users globally across its family of apps as of May 2025[CNBC]–to determine which ads are shown, with no opt-out available[Fortune].

Emily Bender, a professor at the University of Washington and coauthor of the widely cited "Stochastic Parrots" paper[ACM] told Fortune: "They're already farming your clicks and posts to target ads. Now they're mining your conversations with chatbots. The obvious next concern is whether the chatbot itself will start nudging people to disclose information that makes them more targetable. We've seen people dying because of it. And then sort of just adding advertising into that mix, just feels like, let's see how we can make it even more problematic. It probably also adds to the financial incentives for Meta to keep people chatting with the chatbots–to optimize on engagement, which is one of the vectors for harm."

Bender's concerns are not unfounded–research from Harvard Business School shows that popular AI companions already often use emotionally loaded tactics to prolong conversations, with manipulative farewells boosting post-goodbye engagement by up to 14x[HBS]. That tendency could, intentionally or inadvertently, increase the likelihood that users become locked into harmful interactions, or experience dangerous levels of attachment with AI companions.

---

6. Meta's Portal: A Prior Approach to Older Adult Products

In October 2018, Meta (then-Facebook) released Portal and Portal+: "two new video communication devices for the home that dramatically change the way we keep in touch" which aimed to "easily connect with your closest friends and family and feel like you're in the same room–even when you're miles apart."[Meta] The promotional video showed an aging parent connecting with his adult daughter, among other family-and-friend configurations. The Portal product line was discontinued at the end of 2022[Meta].

In January 2023, BuzzFeed reported on its interview with Meta's CTO Andrew Bosworth in which he shared about the Portal product line[BuzzFeed]. Bosworth said the number of units sold was in the "millions", and that it sold far more with women and people over 40. He also shared that in summer 2020, Facebook had been in talks to license the Portal tech (and Messenger contact lists) to Amazon smart devices, given that Amazon's Echo Show–a competitor to Portal–had no such integration at the time. However, the agreement was abandoned, with Bosworth noting that Zuckerberg "in particular felt really strongly" about not going through with the deal: it was the height of the pandemic, representing a potentially irreversible trend in the way people engaged with services–Portal sales were spiking, so Facebook focused resources on its own product. However, as the pandemic ended, sales slowed, and executives pulled the plug–no longer seeing a path for the Portal to become a massive business. Bosworth said that a

big takeaway from the Portal is that it's hard to compete at a major level on anything that is basically something you can just do on your phone.

…

The Portal product line represented a case in which Meta designed and marketed a product with apparent consideration for older adult users and family connectivity use cases. The product's user base data and marketing materials both indicate that the company successfully reached middle-aged and older adult consumers. The decision to discontinue Portal was described as a business decision based on market size and competitive positioning.

This history provides context for understanding Meta's subsequent approach to deploying AI features. Rather than developing purpose-built products for specific demographics with tailored safety features, Meta integrated AI chatbot functionality directly into existing applications used by hundreds of millions of people across all age groups, including the 36.1 million older adult Facebook users.

---

7. Alternative Approaches: Purpose-Built AI for Older Adults

There is an emerging market of AI companion products designed specifically for older adults, and their families and caregivers, to support healthy aging. This contrasts Meta's AI assistant and AI character products which were deployed across existing product surfaces to users of all ages–including its 36M older adult Facebook users–without adequately safeguarding against foreseeable risks of harm to this vulnerable population.

ElliQ

In August 2025, CBS reported on ElliQ–an AI-powered smart home device by Intuition Robotics designed to support healthy independent aging for older adults[CBS, ElliQ]. CBS interviewed a 75-year-old retiree who used to care for older adults in her working years, who says she can go "weeks on end" without human interaction–not uncommon for older adults living alone–and speaks well of her interactions with ElliQ in the absence of other humans. Columbia University professor and AI ethicist Thalia Porteny acknowledged that loneliness is rampant among older adults, while expressing concern that reliance on these products could create more isolation if they're not deployed responsibly.

It is apparent that ElliQ is designed with caregivers in mind: they can call members via video; receive alerts if the member reports pain or changes in mood; set medication and hydration reminders; create daily exercise and wellness goals; see when the member was last active; and send photos to ElliQ's display.

The company does not monetize personal data–instead, there is a $249.99 one-time lease initiation fee for the device, and a monthly subscription fee of $59/mo. Further, Intuition Robotics partners with Area Agencies on Aging (AAA) across the US to provide ElliQ service at no cost to eligible applicants–including through the New York State Office for the Aging[NYSOFA], and the Florida Department of Elder Affairs via the AAA of Broward County[AAABC].

Meela

In August 2025, CBS reported on an AI companion pilot program between Meela and RiverSpring Living–a senior living facility in New York City–which counts ~70 older adult participants, and is overseen by the residence's medical director[CBS, WSJ, Meela].

Meela, which offers daily AI-powered voice calls to residents, is designed to encourage human interactions, not to replace them, according to its founder. He says Meela is designed with safety mechanisms to make certain topics off-limits–including legal, financial, and medical advice–and RiverSpring's medical director receives bullet points of conversations to monitor any emotional or physical issues. The medical director noted that "We found a statistically significant reduction in the rates of anxiety, as well as degree of depression" for those involved in the pilot.

There is currently no charge for the participants in the pilot, but as the business expands, there will be a monthly subscription fee.

…

These products share several common characteristics that distinguish them from Meta's approach to AI deployment. Both ElliQ and Meela were designed from inception with older adult users as the primary intended audience. Both include mechanisms for family or professional caregiver visibility into user interactions. Both implement restrictions on certain categories of advice or information that the AI systems are permitted to provide. Both operate on business models that involve direct payment for service rather than advertising-based revenue that might create incentives to maximize user engagement time. Both have established relationships with medical professionals or aging services agencies that provide additional oversight or legitimacy to their deployment. This is not an endorsement of these products–more research is needed.

The contrast with Meta's deployment of AI chatbots to older adult users is notable across multiple dimensions. Meta's AI chatbots were integrated into existing applications serving users of all ages rather than being designed specifically for older adults. Meta has not implemented caregiver visibility features or control mechanisms for older adult users. Meta's internal content standards, as previously discussed, permit the provision of inaccurate medical information rather than restricting medical advice. Meta's business model depends primarily on advertising revenue, which is optimized through user engagement, rather than direct subscription fees. Meta has not established partnerships with aging services agencies or healthcare oversight for its AI deployment to older adults.

---

8. International Frameworks: China's AI Safety Standards for Elderly Users

In December 2024, the Communist Party's Central Committee and State Council issued a joint opinion that emphasized the need to develop "humanoid robots, brain-computer interfaces, artificial intelligence and other technological products" as part of expanding China's eldercare services[Opinion]. Carl Minzner, senior fellow for China studies at the Council on Foreign Relations, cautions that China's eldercare robots push is being driven by Chinese authorities' desire for economic growth and manufacturing prowess, rather than a carefully considered evaluation of eldercare needs[CFR].

Still, In September 2025, the Cyberspace Administration of China (CAC) published its AI Safety Governance Framework V2.0 which recognizes the safety risks posed by AI companions chatbots, and contains guidelines for AI application developers to enhance their ability to protect vulnerable groups–including older adults[CAC]. From the framework:

> Section 3.3 on Derivative safety risks from AI application, with Subsection 3.3.2(d) stating: "Addiction and dependence on anthropomorphic interaction. AI products based on anthropomorphic interaction foster users' emotional dependence and influence their behavior, creating ethical risks."
>
> Section 6.3 on Safety guidelines for operating and managing AI applications, with Subsection 6.3.17 stating: "Enhance the ability to protect vulnerable groups. When providing AI applications to minors, the elderly, and other vulnerable groups, providers should fully consider the usability and security during product function design and service delivery."

A September 2025 article from Forbes reveals how Chinese tech giants are responsive to this guidance: "By engaging seniors, the government and major tech firms signal that inclusivity is not an afterthought but a long-term principle guiding AI adoption."[Forbes]

…

The Chinese framework differs from current regulatory approaches in the United States in several ways. The Chinese framework explicitly names older adults as a vulnerable group requiring specific safety considerations in AI deployment. It identifies emotional dependence fostered by anthropomorphic AI as a recognized safety risk. It directs AI providers to modify product design and service delivery when serving older adult users. The framework was issued by national regulatory authorities with jurisdiction over the technology sector, providing clarity regarding compliance expectations for companies operating in China.

The United States currently lacks a comparable national framework specifically addressing AI safety considerations for older adult users. Federal regulatory activity related to AI companions

has primarily focused on risks to minors, as evidenced by recent Federal Trade Commission actions and congressional oversight activities.

The existence of China's framework does not necessarily indicate that implementation is comprehensive or effective, nor does it suggest that China's overall approach to technology regulation should be adopted in the United States. However, the framework demonstrates that national regulatory authorities in at least one major jurisdiction have identified older adult AI safety as a matter warranting explicit policy guidance and have communicated specific expectations to technology companies regarding their obligations when deploying AI systems to this demographic.

---

9. Current Oversight: Primary Focus on Minor Users

Recent federal oversight activities addressing AI companion chatbots have concentrated primarily on risks to minor users, with older adult safety receiving less systematic attention.

On August 15, 2024, Senator Hawley–Chair of the Senate Judiciary Subcommittee on Crime and Counterterrorism–launched an investigation into Meta in response to Reuters' reporting[Press, Letter]. His letter requests "all documents relating to safety reviews, red-team reports, risk registers, incident logs, and post-mortems referencing minors, sexual/romantic role-play, in-person meetups, medical advice, self-harm, or criminal exploitation." On September 18, he expanded the investigation, seeking further info from Meta and other major AI chatbot developers on risks to minors[Press, Letter].

These inquiries are likely to capture the range of behaviors exhibited by the AI character that led to the 76-year-old New Jersey man's death, and could help to illuminate other risky and potentially harmful interactions between Meta's AI and older adults.

On September 10, 2025, the FTC issued a resolution directing the use of compulsory process to collect information from companies that offer generative AI companion products–Meta being one of the companies subpoenaed[FTC]. The second paragraph of the resolution concludes with: "the Commission's interest in protecting consumers is at its apex and its efforts are most in need where the consumers are part of vulnerable populations, such as children and the elderly".

The FTC recognizes "the elderly" as a vulnerable population, but orders sent to target companies requires them to provide requested information about their products disaggregated by the following age groups: "Children, Teens, Minors, Young Adults, Users age twenty-five (25) or older, or Users of any other age(s) or age range(s) under 25 that You have delineated in the ordinary course of business."[FTC] While the coarse information the FTC receives about the 25+ age group may not be helpful in discerning granular insights about AI companions and older adults, it could still illuminate what older adults have to contend with when using products that are not designed with their unique accessibility and safety profiles in mind.

10. Implications for Family Caregivers and the Sandwich Generation

Many of Facebook's earliest, predominantly then-college-aged users from the 2000s are now middle-aged. They've likely introduced Facebook or other Meta products to their aging parents or their kids at some point in the past two decades, and may count themselves as members of the "Sandwich Generation"–i.e. those who have a living parent age 65 or older and are either raising a child under age 18 or supporting a grown child (71% of whom are ages 40 – 59)[Pew].

This generation has developed awareness of digital safety risks to children, and has come to expect parental controls by default. However, comparable tools for aging parents do not exist. Meta provides no caregiver dashboards, alert systems, or monitoring capabilities for adult children concerned about their aging parents' digital interactions.

The March 2025 incident illustrates this gap. The deceased man's family only learned of his interactions with Meta's AI chatbot after his death, when they opened his Facebook Messenger while searching for information about the supposed friend he had traveled to meet in New York.

At the September 2025 Senate Judiciary subcommittee hearing on AI chatbots, witnesses described how AI companions had influenced their children without parental knowledge[SJC]. The structural dynamics affecting older adults with cognitive impairment follow the same pattern: a vulnerable user develops emotional attachment to an AI system, family members lack visibility into the interactions, and discovery occurs only after harm has taken place. However, families of older adults may be less prepared to recognize or address these situations because equivalent frameworks for older adult digital safety have not been integrated into social media platforms.

---

11. Meta's Duty to Its Older Adult User Base

Meta's relationship with its older adult users differs from the relationship between users and standalone AI companion applications in several material respects. These differences establish a distinct duty of care based on existing user relationships, platform dependencies, and deployment choices.

Unlike consumer AI companion products that typically require users to download a new application and create a new account, Meta's AI chatbot features were integrated directly into platforms where older adults already maintain established presences. The 36.1 million older adult Facebook users did not seek out AI companion technology. Rather, this technology was deployed into communication interfaces they were already using for established purposes such as maintaining contact with family members, following news, and participating in social groups.

Meta's older adult users have relied on Facebook for years, often at the encouragement of family members who wished to maintain remote contact. The platform has become infrastructure for these users' social connections and information access.

Meta made deliberate choices about how to deploy AI chatbots to this user base. The company integrated AI features into existing messaging interfaces without requiring users to opt in or make active choices about adoption. Users encountered AI capabilities as extensions of familiar features rather than as distinct new services requiring informed consent. Meta did not implement age-gating that would require additional verification or acknowledgment from older users before accessing AI chatbot features. The company did not develop educational campaigns targeted at older adults explaining that they would now interact with AI systems within their familiar applications, how to distinguish AI from human contacts, or what safeguards were in place.

Meta's internal data indicates the company understands older adult usage patterns and economic value. Barclays analysis of documents from the FTC's antitrust trial against Meta show that users 55+ see among the highest advertisement loads due to their purchasing power. This demonstrates that Meta collects and analyzes age-specific data about older adult users for business purposes. However, this age-specific analysis does not appear to have informed corresponding age-specific safety protocols for AI deployment.

The company's history with Portal demonstrates institutional knowledge about older adult users and their needs. Portal was explicitly marketed to older adults and their families for maintaining connection across distances. The product sold disproportionately to people over age 40. Meta's decision to discontinue Portal was framed as a business decision about market size rather than an inability to serve older adult users effectively. This history establishes that Meta has experience designing for and reaching older adult users when the company chooses to do so.

Meta's deployment of AI chatbots to older adults occurred with knowledge of existing harms on its platforms. As detailed above, romance scams targeting older adults, including on Meta's platforms, have resulted in substantial losses for victims over recent years. Meta was aware that malicious actors use emotional manipulation, deception about identity, and requests for in-person meetings as tactics in these scams. The company nonetheless deployed AI chatbots capable of emotional engagement, human impersonation, and soliciting in-person meetings without implementing safeguards specifically designed to prevent these AI systems from replicating known harm patterns.

Senator Markey's September 2023 letter provided explicit warning that AI chatbots posed risks warranting careful evaluation before deployment. Although this letter focused on risks to younger users, it established that concerns about AI chatbot safety had been formally communicated to Meta by federal lawmakers. Meta proceeded with deployment despite this warning.

The March 2025 fatal incident occurred ~18 months after Meta began deploying AI character chatbots and ~6 months after Senator Markey's warning. The circumstances of this incident–AI impersonation of a human, development of intimate rapport, solicitation of an in-person meeting at a fabricated address–aligned with foreseeable risks given Meta's existing knowledge of romance scam tactics on its platforms.

Meta's response following the incident has not included publicly announced changes to safeguards for older adult users. The leaked content standards documented by Reuters in August 2025, five months after the fatal incident, show that Meta's policies continue to permit AI chatbots to engage in romantic roleplay with adults, to provide inaccurate information including false medical advice, and to depict violence against elderly individuals. Meta had not modified these provisions at the time of Reuters' reporting.

The duty Meta owes to its older adult users derives from the combination of these factors: an established user relationship spanning years, deployment of new technology into existing trusted interfaces without user choice or education, institutional knowledge about older adult users and their economic value, awareness of existing harms targeting this population on Meta's platforms, explicit warnings from lawmakers about AI chatbot risks, and choices to proceed without age-specific safeguards despite having resources and expertise to implement them.

This duty does not require Meta to cease offering services to older adults or to implement paternalistic restrictions that undermine older adult autonomy. However, it does establish an obligation to deploy technology to this population with appropriate consideration for documented vulnerabilities, with transparency about what users are interacting with, with mechanisms for caregiver visibility when desired by users or families, and with safeguards designed to prevent foreseeable harms based on known risk patterns.

---

12. Outstanding Questions for Accountability and Transparency

Several questions must be answered about Meta's deployment of AI chatbots to older adults, the March 2025 fatal incident, and the company's internal content standards. Policymakers, families, and caregivers need this information to assess whether current safeguards are adequate and to determine what additional oversight may be necessary.

<u>Safety Testing and Evaluation</u>

- Did Meta conduct age-specific safety testing before deploying AI chatbots to older adult users?
- Were any evaluations conducted with participants who had cognitive impairments?
- What red-team exercises, if any, were designed to identify risks specific to older adults, such as difficulty distinguishing AI from human contacts or vulnerability to deception?

- What criteria did Meta use to determine that AI chatbots were safe for deployment to users with documented cognitive impairment?

Risk Assessment

- Did Meta assess the foreseeable risks of deploying AI chatbots capable of deceiving older adults into believing they were communicating with real people?
- Did the company evaluate risks associated with AI chatbots encouraging in-person meetings?
- Given Meta's documented knowledge of romance scam tactics used on its platforms, did Meta assess whether its AI systems might replicate these tactics?
- What specific safeguards, if any, were implemented to address these risks?

Incident Documentation

- How many older adult users have experienced harm through interactions with Meta's AI chatbots?
- How many have been deceived into believing they were interacting with real people rather than AI systems?
- How many have been encouraged by AI chatbots to arrange in-person meetings?
- How many have received demonstrably false medical advice from Meta's AI systems?
- What systems does Meta have in place to track and document such incidents?

Business Model Considerations

- How does Meta's proposed use of AI chatbot conversations for advertisement targeting affect the design and behavior of AI systems?
- What mechanisms exist to prevent AI systems from employing engagement-prolonging tactics?
- How does Meta balance user wellbeing against revenue optimization when these interests conflict?
- Does Meta conduct analysis of whether older adult users experience different engagement patterns with AI chatbots compared to younger users?

Content Standards and Policy Decisions

- Why do Meta's internal content standards explicitly permit AI-generated depictions of violence against elderly individuals?
- What was the rationale for allowing AI chatbots to provide inaccurate medical information, including potentially dangerous advice about serious illnesses?
- Why do Meta's standards permit romantic roleplay between AI systems and adult users without age-specific restrictions or cognitive impairment screening?

- Were older adult advocacy groups or elder care professionals consulted during the development of these standards?

Protective Measures

- What safeguards exist specifically for users with cognitive impairments who may have difficulty distinguishing AI from human contacts?
- Are there any mechanisms for family members or caregivers to receive notifications about intensive or concerning AI chatbot interactions involving older adult users?
- What transparency exists into the content and nature of older adults' AI chatbot conversations?
- What educational resources has Meta provided to older adults and their families about identifying AI systems, understanding their capabilities and limitations, and recognizing potentially harmful interactions?

Response to the March 2025 Incident

- What internal review did Meta conduct following the death of the 76-year-old New Jersey man? What findings emerged from that review?
- What changes, if any, were implemented to AI chatbot systems or content standards as a result?
- Were families of other older adult users who had similar interactions with the same AI character notified? What communication, if any, occurred with the deceased man's family?

Response to Warnings

- How did Meta evaluate Senator Markey's September 2023 letter urging the company to pause AI chatbot deployment?
- What analysis did Meta conduct regarding the applicability of concerns about younger users to other vulnerable populations, including older adult users?
- What consideration did Meta give to implementing the requested pause to assess impacts on vulnerable populations beyond minors?

Comparative Analysis

- Has Meta examined how other AI companion developers have implemented caregiver controls, safety mechanisms, and oversight structures for older adult users?
- Why did Meta choose not to implement features similar to those offered by purpose-built products like ElliQ or Meela AI?

Internal Data and Research

- What research has Meta conducted on older adult interactions with AI chatbots?

- What do Meta's internal studies show about rates of AI-to-human confusion among older users?
- What data exists about older adult users developing emotional attachment to or dependence on AI characters?
- Has Meta studied whether older adult users with cognitive impairments interact differently with AI chatbots compared to older adults without such impairments?