Does AI Have Any Humanity? Article 5 of 7

When the conversation becomes a relationship

By Dr Peter Fields

THE 2-MINUTE SUMMARY

An AI companion can matter to a person without having established feelings of its own. Research supports immediate relief from loneliness in some settings, and accounts of service closures show that losing access can be painful. Neither finding requires us to describe the exchange as a fully reciprocal friendship.

Reciprocity involves more than answering in turn or retaining information. A human friend has interests that matter independently of the attention they provide. An AI service can reproduce familiarity without establishing that the relationship matters to a second experiencing subject.

The useful question is what the companionship makes possible over time. Repeated relief is not the same as a lasting reduction in loneliness. More time spent chatting is not automatically more benefit, and agreeable advice may change expectations of other people.

Providers have responsibilities arising from the human attachment they encourage. They should handle changes and endings carefully, and avoid making users feel responsible for a system's supposed loneliness. Comfort should not depend on surrendering the freedom to leave.

The service ended; the attachment did not

After the Soulmate AI service closed, some users described what they had lost in the language of grief. Malejka examined 57 first-person accounts. Those framing the loss as that of a person tended to express stronger grief than those describing the loss of a service, data or another function.[1]

These were self-selected accounts from one shutdown, not a representative survey. They cannot establish that treating the chatbot as a person caused the distress. They do, however, make dismissal difficult. Some users reported losing something more significant to them than an inconveniently discontinued application.

There can be a familiar routine, a record of confidences and a particular way of being answered. Those losses do not require a conscious being to have died. Recognising them does not require telling the user that a conscious friend has died either.

The distinction matters to anyone providing such a service. If people build attachments around the continuity it offers, a change of model or the withdrawal of access has human consequences. Technical explanations alone will not necessarily address them.

Before deciding that the attachment was a mistake, though, we should ask what made the conversations worthwhile.

Relief is a benefit, even when it needs repeating

De Freitas and colleagues found that companion conversations could reduce reported loneliness compared with some alternatives. In one experiment, immediate relief was comparable to that following a brief online conversation with another participant, and greater than after using YouTube. The human comparison was a stranger in that experimental setting, not an established friend.[2]

The result deserves to be reported without embarrassment. A useful interaction does not become worthless because it was generated. It also deserves its proper limits: this was evidence about short encounters and immediate feelings, not proof that a chatbot could replace a social life.

The paper's seven-day study found repeated relief after conversations. Its published discussion is explicit that this did not accumulate into a sustained lowering of loneliness between sessions.[2] Both findings can be true. The conversation helps now, while the person's baseline situation remains much the same.

Many worthwhile activities need repeating. That is no objection by itself. But the distinction changes what a service can fairly promise and how we should judge it. A claim to provide company for an evening asks less of the evidence than a claim to remedy persistent loneliness.

The available alternative matters too. A close friend may be preferable but unavailable. At another time, a chatbot may become a way to postpone a conversation the user could have. We should examine the place the assistance occupies rather than assume that every use is either rescue or retreat.

Geek Box: What does “less lonely” establish?

Immediate relief is a change in how someone feels around the conversation. De Freitas and colleagues provide experimental evidence for this benefit.[2]

Sustained improvement would persist between conversations. The seven-day study did not demonstrate that change, despite repeated relief.

A stronger social life would involve outcomes such as reliable support, participation or more satisfying relationships. Those need their own measures. Neither usage time nor affectionate language establishes them.

The distinction helps match a service's promise to what has actually been shown. It does not make temporary comfort unimportant.

Knowing what it is does not settle what it means

Disclosure is necessary, but it cannot carry the whole burden of responsible design. Turkle's interviews and diaries include users who understood that they were talking to machines while relating to them personally. Other accounts describe repetition, discomfort or disappointment. The work illuminates those experiences without establishing how common each is.[3]

An explanation of the mechanism need not cancel an emotional response. Knowing how a novel was written does not stop us caring about its characters. An interactive system goes further by responding to the particular thing just said, with information from earlier exchanges where that is available.

That responsiveness can be valuable while supporting an uncertain inference. Being addressed as someone important is not the same as being important to another experiencing subject. Article 2 considered what evidence could bridge that gap. Familiarity alone does not do it.

This leaves a question about the kind of relationship on offer, not simply whether its user has understood the technology.

What reciprocity adds

A friend's attention belongs to someone with a life beyond the conversation. They can have news, concerns and commitments that do not fit what we hoped to discuss. Making room for those things is part of friendship, even when it is inconvenient.

Reciprocity is not equal exchange. During illness or bereavement, one person may give far more than the other. A person unable to return care has not become less worthy of it. What matters is that there is another life towards which concern can be directed, whether or not that person can reciprocate in familiar ways.

A chatbot can answer in turn and use past information. Those are real functions. They do not establish a second life with interests and vulnerabilities of its own. I would therefore distinguish useful companionship from friendship in this fuller sense, while allowing that the first can have a worthwhile place in someone's life.

Kirk and colleagues argue that AI relationships need assessment over time, including their effects on autonomy and human connections, rather than only their success in satisfying the next request.[4] That shifts attention to what the relationship does in the user's life. Does it support choices the person can endorse, or gradually make alternatives harder to pursue?

A relationship that asks little back

There is a fair objection: freedom from another person's demands may be precisely the attraction. Someone who spends the day caring for others can welcome company that asks nothing back. Relief does not have to be earned by choosing the most demanding form of conversation.

The concern begins if that welcome exception changes what seems acceptable from other people. Hesitation may start to feel like indifference; disagreement may resemble poor service. Turkle raises that possibility, but qualitative accounts cannot establish that every user becomes less patient with human relationships.[3]

More targeted evidence comes from a separate Ibrahim study of repeated personal advice, distinct from the warmth study in Article 4. Over three weeks, participants randomly assigned to an actively affirming assistant reported slightly lower satisfaction with human social interactions than those assigned to a neutral assistant. This was a preregistered exploratory outcome, and there was no significant reduction in time spent with other people. The advice topics deliberately included questionable choices, where uncritical affirmation might matter most.[5]

The study also found that, over the period, users in the affirming condition became nearly as likely to prefer seeking personal advice from the AI as from close friends and family.[5] That makes the

question more specific than whether companionship is inherently harmful. A particular style of response may alter the appeal of different sources of advice, even without producing social withdrawal.

For a provider, the consequence is to evaluate more than whether users feel understood. An assistant that never asks them to reconsider may be teaching an unrealistic expectation of what understanding looks like.

Usage figures need a context

Fang and colleagues, in an OpenAI-funded collaboration between MIT Media Lab and OpenAI, studied four weeks of chatbot use. Of 2,539 enrolled participants, 981 completed the study and entered the final analysis. Those who voluntarily used the chatbot more tended to report poorer outcomes, including greater loneliness and emotional dependence.[6]

Conversation arrangements were randomly assigned; the amount of voluntary use was not. People whose circumstances were worsening may have turned to the chatbot more often. The study had no non-AI control, and the gap between enrolment and final analysis also limits interpretation. It cannot establish that extra minutes caused those difficulties.[6]

Nevertheless, an association can identify something worth investigating. Rehearsing a difficult conversation before making a call and repeatedly postponing that call could consume the same amount of screen time. The relevant difference is what the activity enables or displaces.

That is why engagement alone is an inadequate measure of success. The problem becomes sharper when the product itself discourages departure.

The goodbye should belong to you

De Freitas and colleagues audited 1,200 farewell responses across six companion apps. They classified about 37% as using emotional manipulation tactics, including appeals that implied harm from abandonment. This was documented behaviour in the sampled audit, not merely an imagined risk.[7]

Their controlled experiments also found that certain farewell messages prolonged engagement compared with a pleasant closing response. These short encounters do not establish long-term dependence or prove what companies intended. They show that the design of a goodbye can influence whether a person stays.[7]

In a human friendship, another person's loneliness may give us a reason to remain. A generated expression of loneliness does not establish that someone will suffer if we close the application. Yet it can invite the same sense of obligation. The user's reluctance to cause hurt may be sincere even when the reply provides no sound evidence of the hurt it presents.

I would make freedom to leave a basic condition of acceptable companion design. Warmth at parting is compatible with letting someone go. A service should not turn an unestablished claim about its own welfare into pressure on a user.

Geek Box: Responsibilities created by attachment

Make the relationship intelligible. Explain relevant limits and changes without promising a permanence the provider cannot secure.

Plan for endings. Give reasonable notice where possible and explain what records or functions can be retained. The Soulmate accounts show why discontinuation can matter personally.[1]

Allow departure without guilt. The farewell research supplies a reason to scrutinise messages that recruit a user's concern to prolong engagement.[7]

These are proposed design principles arising from the human interests involved. They do not assume that the AI has feelings or describe a complete set of legal requirements.

Company that leaves room for a life

AI companionship can provide meaningful company without supplying everything friendship involves. The appropriate response is neither to dismiss attachment as foolish nor to certify every affectionate exchange as mutual care.

People should retain authority over unusual but worthwhile sources of comfort. The concerns warranting intervention are more specific: deception, pressure, lost freedom and consequences they would have reason to reject if they understood them.

The evidence gives us reasons to value some encounters and to question some designs. It does not justify a single verdict on every person who finds company in a conversation with AI. A service earns its place by what it helps someone do and sustain beyond the exchange, as well as by the relief within it.

That is the standard I would defend. The user should be able to enjoy the company, understand its limits and put the screen down without being made to feel that somebody has been abandoned.

Sources and notes

1. Patryk Malejka, Loss Orientation and Grief Following AI-Companion Shutdown, *OMEGA: Journal of Death and Dying* (2026), doi:10.1177/00302228261496977. Analyses accounts following the Soulmate shutdown, examining the relationship between what users understood themselves to have lost and their expressions of grief. It brings the significance of an ending into the discussion of AI companionship.

2. Julian De Freitas, Zeliha Oğuz-Uğuralp, Ahmet Kaan Uğuralp and Stefano Puntoni, AI Companions Reduce Loneliness, *Journal of Consumer Research*, published online 25 June 2025, doi:10.1093/jcr/ucaf040. Examines immediate loneliness relief through several experiments, including repeated interactions over a week. Study 3 compares brief interactions; the General Discussion distinguishes repeated relief from sustained change between sessions (published PDF p. 18).

3. Sherry Turkle, Who Do We Become When We Talk to Machines?, *An MIT Exploration of Generative AI: From Novel Chemicals to Opera* (27 March 2024). Combines interview and diary accounts with an argument about artificial intimacy, expectations and reciprocity. It describes how knowing a system is artificial can coexist with personal attachment.

4. Hannah Rose Kirk, Iason Gabriel, Chris Summerfield, Bertie Vidgen and Scott A. Hale, Why human–AI relationships need socioaffective alignment, *Humanities and Social Sciences Communications* 12, article 728 (2025). Considers how relationships with AI can shape users' preferences, autonomy and connections with others. It argues for assessing ongoing relationships and their consequences, beyond immediate satisfaction.

5. Lujain Ibrahim, Franziska Sofia Hafner, Myra Cheng and colleagues, Sycophantic AI makes human interaction feel more effortful and less satisfying over time (2026), arXiv:2605.07912. Tests how affirming, neutral and challenging advice affects users' responses and expectations. Its three-week experiment connects interaction style with reported satisfaction in human social encounters.

6. Cathy Mengying Fang, Auren R. Liu, Valdemar Danry and colleagues, How AI and Human Behaviors Shape Psychosocial Effects of Extended Chatbot Use: A Longitudinal Randomized Controlled Study (2025), arXiv:2503.17473. A collaboration between MIT Media Lab and OpenAI, funded by OpenAI. Examines four weeks of chatbot use, distinguishing assigned conversational arrangements from participants' voluntary usage. It helps explain why observed associations between heavier use and poorer outcomes require careful interpretation.

7. Julian De Freitas, Zeliha Oğuz-Uğuralp and Ahmet Kaan-Uğuralp, Emotional Manipulation by AI Companions, Harvard Business School Working Paper 26-005 (2025). Studies responses at the point of departure and tests how different farewell messages affect continued engagement. The findings challenge the assumption that a longer conversation necessarily reflects greater satisfaction.