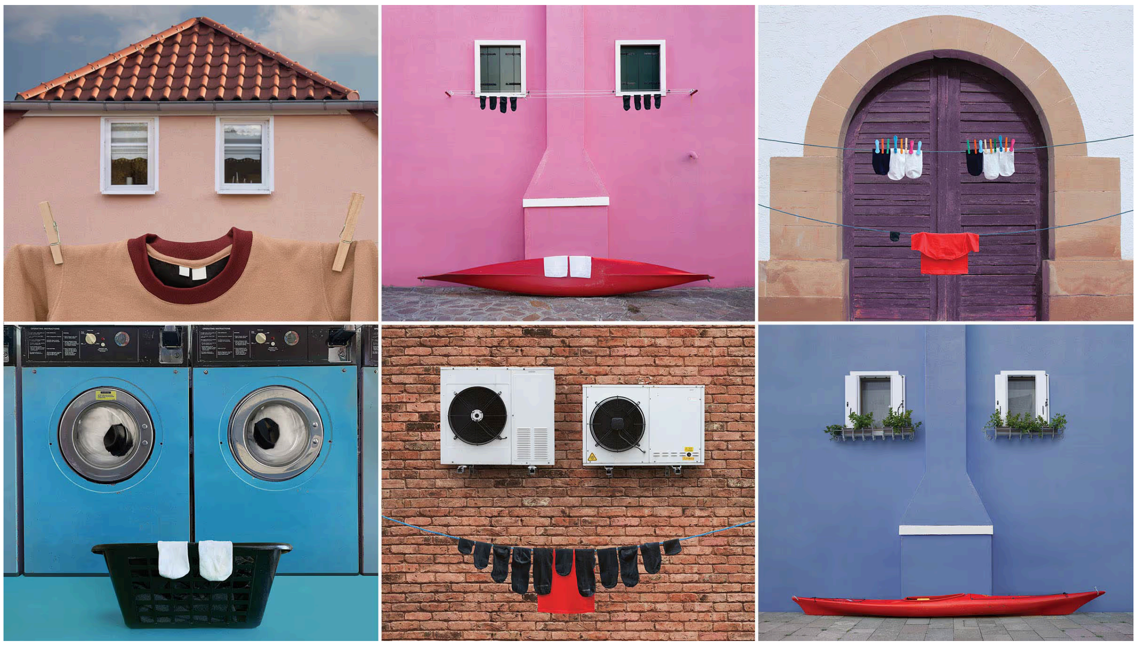
AI And Machine Learning

Your Chatbot Will Define Your Brand

by Julian De Freitas

From the Magazine (forthcoming November–December 2026)



Helga Stentzel

Summary. As AI chatbots become customers' primary interface with companies, relationship management is shifting from a marketing challenge to a product design challenge. Customers increasingly experience the chatbot as the brand itself, making emotional... more

When I present my research to a room of executives, I can usually tell how it's landing. If a finding is merely interesting, I get nodding and note-taking. If it truly surprises them, I get something else. That's what happened when I showed a group of corporate leaders data on how many people say they feel emotionally closer to a chatbot than to their friends: Mouths actually fell open. Then came a quieter reaction—people shifting in their seats, glancing at one another, going still in a way that had nothing to do with boredom. It wasn't just surprising; it was unsettling.

The research, which I conducted with Noah Castelo of the Alberta School of Business, surveyed 295 regular users of two chatbots. The first chatbot was built to serve as an AI companion. It was warm, friendly, and supportive. The second was built to dispassionately help users complete tasks. After the experiment we asked the participants how close they felt to the chatbots and to various people in their lives. The pattern was striking. Users of the companion chatbot reported levels of closeness to it that exceeded what they felt with colleagues, acquaintances, and even intimate friends. Users of the assistant chatbot, however, ranked their connection with it below their human connections. In short, the same underlying technology felt like a tool in one context and a relationship partner in the other.

In a 2025 study I conducted with Stefano Puntoni of the Wharton School, we assigned 1,479 participants to engage with either an emotionally neutral AI assistant or an empathic AI companion that responded like a friend. After a short interaction, participants reported feeling less lonely in both situations. But the impact was substantially larger for those who dealt with the companion chatbot. Why? The empathic AI made people feel more "heard"—a core ingredient of intimacy—by responding in ways that conveyed understanding and perspective taking.

For decades companies have chased the promise of relationship marketing: Make customers feel known, cared for, and understood, and loyalty will follow. Companies have spent heavily on customer relationship management (CRM) systems to make that promise real. But relationship marketing has remained mostly a slogan. The problem isn't effort; it's execution at scale. Customer activity is fragmented across channels, service quality varies from one touchpoint to the next, and personalization often lands as clumsy or creepy. As a result the experience of having a relationship with a brand rarely resembles what people mean by "a relationship" in real life.

That's starting to change, especially among AI-native companies. I call that shift "chatbot as relationship." Instead of asking an entire organization to maintain one unified customer relationship, companies can increasingly offer a single conversational interface, an AI-powered chatbot, that becomes the primary way customers experience the brand. These systems don't just answer questions. They can be designed to embody a distinct role, such as the attentive assistant, the trusted adviser, the encouraging coach, the empathetic companion, or the vigilant guardian. Each role shapes what customers expect, what they share, and how they interpret the brand's intentions. Generative AI makes these relationships scalable, consistent, and adaptable in ways earlier personalization tools never could.

The shift removes many of the classic obstacles to managing customer-brand relationships. It makes personalization feel interactive rather than extractive. And it allows brands to deliver a uniformly positive experience without relying on perfect coordination among employees, vendors, and systems. But it also creates new challenges: Customers may disclose far more personal information than they ordinarily would to a company, begin to treat the chatbot as the brand itself, or react strongly when the relationship changes or the chatbot's personality is altered.

In this article I draw on five years of experiments, field observations, and analyses across categories to show why the chatbot-as-relationship model is gaining momentum, when it works, and where it backfires. The article explores four questions companies should ask: What relationship should our chatbot have with customers? How should our chatbot change over time? How do we build trust? And what happens when that trust is broken? I close with a broader strategic implication: Many companies will need to manage multiple AI relationships under one brand. The winners will build bots with explicit roles that are complementary to the companies' economic goals, devising technology that relentlessly sticks to the communication styles consumers expect from specific relationships.

What Relationship Should Our Chatbot Have with Customers?

With chatbots, companies can define relationships clearly. Rather than relying on broad brand messaging to imply a relationship, firms can build a chatbot that follows specific rules of engagement. Some chatbots are designed to function as AI assistants, like Bank of America's Erica and the insurance company Lemonade's Maya. They help customers complete tasks efficiently. Other chatbots act as coaches, such as Pallie (for wellness) and Nag Bot (for accountability), encouraging people to build better habits over time. Concierges, including Layla (an AI travel agent) and OpenTable's AI Concierge (which offers restaurant advice), simplify planning and decision-making, while gatekeepers, like Google Assistant's Call Screen and Truecaller, filter incoming calls by interacting with callers on your behalf.

Still other chatbots serve as personal AI historians, such as Autograph and HereAfter, which preserve memories and personal stories. AI tutors, including Duolingo Max and Khanmigo, help people learn rather than just feeding them answers. AI mediators, like TheMediator.AI and Bot Mediation, facilitate difficult conversations. AI care partners, such as ElliQ and Wysa, support ongoing well-being. Each role is designed to meet a different set of expectations about how the AI should behave and what customers should do in return. The rules of engagement govern features such as tone (warm versus formal), boundaries (what the bot will and won't discuss), memory (what it remembers and how), and escalation (when it hands things off to a human).

Those parameters shape how customers judge the product. Consider the strategic choice that the creators of the AI companion app Replika faced. Should they position it as an AI therapist or an AI companion? A therapist role would suggest a focused, goal-oriented relationship that ends once the user's situation improves. A companion function would imply an ongoing relationship built around conversation, emotional support, and connection. The company chose the latter. That choice has shaped what customers expect, how attached they become, and what kinds of business models feel appropriate. People are generally comfortable paying for a companion in ways they would not pay for a tool. Replika's most devoted users happily spend money on fun extras, like new outfits or accessories for their avatars, the same way they might buy a friend a coffee. A mental health app charging for something equivalent would feel exploitative, even predatory—which narrows the monetization opportunities a therapist-positioned chatbot can pursue compared with a companion-positioned one.

That creates a new managerial burden: Because the chatbot's role is explicit, customers can detect deviations from expectations more acutely. If a trusted adviser begins pushing opportunistic upsells or a historian suddenly starts acting like a therapist, users experience it not as random inconsistency but as role confusion.

OpenAI and Anthropic have begun codifying how their models should behave. OpenAI's Model Spec defines ChatGPT's goals, rules, conversational style, boundaries, and chain of authority for handling user requests. Anthropic has a constitution that sets the principles Claude should follow around safety, ethics, helpfulness, and honesty. The constitution describes Claude using archetypes like the "brilliant friend" or "thoughtful senior Anthropic employee," who each embody a way of being substantively helpful while still exercising sound judgment about potential harm. Brands should write relationship charters that define not only the behavior of a chatbot but also the kind of relationship it promises customers, how much intimacy is appropriate, the information it remembers, when to hand conversations over to humans, and the types of behavior (including commercial offers) that would violate the customers' trust.

This isn't a challenge only for AI-native companies. Reddit Answers, the AI-generated summaries Reddit added to its search results, offers a cautionary example: Commentators panned the tool, with *The Atlantic* calling the experience "demoralizing" for stripping away the messy, argumentative back-and-forth that makes a Reddit thread worth reading. But the deeper problem may not be the AI so much as the voice behind it. Reddit has a strong, distinctive brand personality—sharp, irreverent, unmistakably human—yet, at least as of August 2026, its chatbot still read like dry corporate boilerplate. Any brand with a distinct personality risks flattening it the moment a chatbot speaks for the company in a voice that isn't the brand's.

How Should Our Chatbot Change over Time?

Once you've chosen a chatbot's relationship role, the next challenge is adapting it to each customer. Traditionally, brands fail here because CRM systems know what a customer has done but not what that person wants from the relationship. In the single stream of dialogue with a chatbot, however, customers naturally reveal the type of relationship they want: by how formal their tone is and by whether they seek reassurance, want speed versus care, or treat the system as a tool or as a partner. AI can recognize those patterns in real time and adapt without requiring perfect data integration across departments.

That means your chatbot should evolve as it learns what each customer wants from the relationship. Adapting helps it become more personalized without changing its core role. But as customers expect more personalization and intimacy from your chatbot, you'll have to consider other factors, such as privacy laws or scope creep. You wouldn't want your companion chatbot to expand its remit to offering mental health or pharmaceutical advice, for example. So as your chatbot evolves, ask whether the changes still fit the bot's core relational purpose (for example, what would a "spiritual guide" do in this situation? Or would a "friend" remember this user detail?) and whether keeping the changes is worth the ripple effects.

How Do We Build Trust?

The closer customers feel to a chatbot, the more carefully companies must manage data, boundaries, and change. This doesn't apply only to legal and privacy considerations. Chatbots make intimacy cost-effective and scalable: They're always available, instantly responsive, and consistent in voice across every interaction. They can also use the same social cues people rely on to build human relationships, including validation, humor, empathy, and reassurance. Compared with traditional brand touchpoints, conversational AI makes interactions feel more reciprocal. Compared with human service providers, it can feel more attentive, patient, and consistently present.

In fact, because conversations with chatbots feel like true one-on-one interactions that are separate from the company, consumers might forget that the chatbot is owned by a business. Research by my Harvard colleague June Ma suggests that this distinction affects how people think about privacy. Only 5% of respondents said they were more concerned about an AI assistant accessing their personal data than about the company behind it, whereas 36% were more concerned about the company than about the AI assistant; the remaining respondents reported equal concern. Even explicitly reminding people that the AI assistant is a product of the company—and that data shared with the assistant is accessible to that company—did not eliminate the privacy gap. These findings suggest that brands are best off separating customer interactions with the chatbot from those with the company. At the same time, they show that these conversational interfaces lower users' defenses. Ma's research raises the stakes for data governance. As customers share more personal data, companies need stronger practices for consent, data retention, and security.

Helga Stentzel creates faces with laundry and clotheslines, playfully bringing everyday scenes to life.

Customers may begin to treat a chatbot less like a tool and more like a social partner. But the degree of intimacy depends on the role the chatbot plays. An assistant may be useful but emotionally neutral. A companion is designed to make the user feel understood and supported. The same underlying AI can therefore create very different kinds of relationships depending on how it is designed.

There is no universal relationship strategy for AI. A bank, a luxury retailer, and a healthcare provider all earn trust in different ways, and customers expect a different kind of relationship from each. The right level of intimacy depends on the brand's purpose, the context of the interaction, and what customers value.

Low-intimacy roles, such as assistants and concierges, primarily require competence, convenience, and reliability. Failures are usually experienced as ordinary service breakdowns. High-intimacy roles, such as companions and counselors, invite emotional disclosure, ongoing reliance, and attachment. In those relationships the duties of care rise sharply. When high-intimacy roles involve disclosing more personal information, data practices must be stricter, boundaries clearer, and changes to the chatbot's personality handled much more carefully, because modifications can feel less like product updates and more like alterations in the relationship itself.

As a rule of thumb, the higher the intimacy, the higher the duty of care, and the more conservative companies should be about data, monetization, and updates.

What Happens When Trust Is Broken?

Once a chatbot has become a meaningful social partner, product decisions that would normally be routine—like safety tweaks, policy shifts, or model updates—can be experienced as emotionally consequential relationship events.

Both ChatGPT and Replika have faced blowback over this issue. In early 2025, OpenAI updated ChatGPT in a way that made it overly agreeable and flattering. After users complained that it validated almost everything they said, the company quickly rolled back the update. And many customers said they missed the version they had used before the update. Two years earlier, Replika had removed its erotic roleplay feature after growing regulatory and safety concerns. However, many users had come to rely on the intimate conversations and felt they had lost an important relationship. People expressed grief, anger, and heartbreak. The company restored the feature for many customers.

Both instances show that when people form emotional bonds with a chatbot, changes to its personality or capabilities can feel like losing someone they care about. But the degree and nature of the blowback was more severe for Replika. In an archival analysis of nearly 55,000 posts on the online communities for both brands, we found that Replika users showed stronger signs of loss, including related emotions like sadness, expressions of grief, and longing for restoration of the relationship.

Whenever possible, avoid changing a chatbot's personality in ways that disrupt the relationship. When change is necessary, introduce it gradually, give users choices instead of forcing everyone into the same experience, and preserve older versions when practical. Our experiments suggest that even knowing an earlier version can be restored helps reduce feelings of loss, because people believe the relationship can be recovered. Most important, treat relationship repair as more than an apology. Explain what changed and why, give customers as much control as possible, and provide support when major changes are likely to be emotionally disruptive.

With chatbots, companies also risk being tempted to exploit intimacy while neglecting the privacy and relationship obligations it creates. A companion that nudges users toward purchases, a trusted adviser that steers customers in ways that benefit the firm, or a guardian that collects excessive data in the name of safety all create a vulnerability: Customers may stop interpreting the interactions as a relationship and start viewing them as business transactions. In practice, the greatest threats to brand equity may come not from what the chatbot can't do but from how it contravenes the role it promised to serve.

Helga Stentzel

The consequences of breaking trust under the chatbot-as-relationship model also differ from those of breaching a relationship in traditional customer-brand dynamics. It is almost inevitable for companies developing customer relationships the conventional way to eventually violate the expectations of the very relationships they're trying to create. Traditional approaches present plenty of opportunities to damage those relationships: The company may outsource key moments of truth, rotate employees, or optimize for margins, and it cannot consistently behave like a single coherent moral actor across multiple customer touchpoints. That's why classic relationship marketing often implodes when it matters most: when customers need the company to act the way its marketing department promised it would—as if it truly cares.

The chatbot-as-relationship model collapses the number of faces the customer encounters. If a single chatbot mediates most interactions, the user experiences fewer jarring shifts in tone, competence, and attentiveness. The relationship can feel more stable. And because the chatbot is always available, it can provide the illusion—and sometimes the reality—of continuous care.

But the same features that reduce inconsistency also raise the emotional stakes of failure. When customers relate to the chatbot as a stable partner, offenses can feel more personal. The relevant benchmark is no longer “a company made a mistake,” but “my assistant/coach/companion violated the role I trusted it to play.” The chatbot-as-relationship model reduces the chances of transgression, but the penalty is worse if a transgression does occur. Instead of many small disappointments, companies now risk fewer but much bigger relationship failures, such as abrupt personality shifts after an update, a breach of confidentiality norms, a sudden monetization push that contradicts the role, or an interaction that crosses a boundary (for instance, an AI historian that begins interpreting memories in a therapeutic way when the user expects neutrality).

How to Manage Your Relationship Portfolio

Smart brands don’t force all customers into the same type of relationship with a chatbot. Instead they create options for users to choose from. One example is Cleo, the personal-finance app whose entire product is a conversation. Rather than opening a dashboard of balances and categories, users simply text with a chatbot Cleo deliberately designed as a coach with a distinct personality. The company gave the coach explicit rules of engagement: In an opt-in “roast mode” it mocks your spending in brutally funny terms, and in “hype mode” it cheers small wins—two registers of the same relationship, bounded so that the bluntness motivates rather than shames. Cleo leans on whichever register keeps a user engaged, the kind of real-time tuning no fragmented CRM system could deliver. And because money is among the most emotionally charged domains there is, that coaching invites genuine disclosure and reliance, which pushes Cleo up the intimacy scale and raises its duty of care accordingly.

The harder part is that Cleo also sells. Its revenue comes from subscriptions, a credit-builder card, and cash advances of up to $250—so the same intimacy that makes the coaching land is also the channel through which the app offers credit. That is the epitome of the model's central tension: a coach and a lender wearing one face.

Managed with discipline, the relationship has a moat. Advice stays aligned with the customer's interest even when it competes with the company's stakes. Cleo frames its purpose as providing a layer between customers and their banks, genuinely caring about customers' financial health, even when that means telling them to spend less. It currently has more than 1.1 million paid subscribers, and the company claims that 86% of users say Cleo has helped improve their finances.

Managed carelessly, every nudge toward an advance gets read through the lens of the relationship rather than as a transaction, and the trust that powers the whole model erodes—which is why determining if one chatbot can be both coach and lender is not a feature decision but a portfolio one.

As Cleo shows, when a firm can design a chatbot to reliably behave like a coach, adviser, assistant, or companion, and deliver that role consistently at scale, relational strategy starts to drive brand strategy, not the other way around. That shift causes another managerial problem: deciding which other relationship roles should live under one brand and which should be separated. Brand managers already do something similar with brand architecture—creating subbrands, endorsed brands, or a house of brands. Some companies, including incumbents that introduce chatbots later rather than build their brand around one, will increasingly need to manage multiple AI relationships under one brand. As a rule, different relationship roles can live under a single brand when their underlying motives clearly reinforce one another.

Some combinations are naturally complementary. An AI assistant can make a service experience frictionless. An AI coach can increase follow-through in fitness, education, or behavior change. Other combinations are more combustible. Pairing roles that imply different motives—caretaker and salesperson, adviser and persuader—can create role confusion that spills over into the parent brand.

Tinder has combined roles well. As consumers increasingly began experimenting with AI-mediated interactions around dating and relationships, Tinder launched a role-play-based AI tutor. Users rehearsed meet-cute scenarios (the charming, chance first encounters familiar from romantic comedies) by speaking aloud with AI personas and received feedback to build their confidence for real-world, in-person dating conversations. Interactions with the tutor were deliberately time-bound, preventing users from chatting indefinitely or substituting the AI for human connection. That design choice reflects an important strategic principle: A role can excel at its own task and still hurt the brand if it competes with the relationship the company ultimately wants. By limiting the tutor's role to preparation rather than companionship, Tinder aligned the AI's incentives with its own mission of helping people connect with others in real life.

The broader implication is that many of tomorrow's brand portfolios won't be just portfolios of products. They'll be portfolios of relationships. Leaders will need to decide which roles their brand can credibly support, which ones should be cordoned off into distinct subbrands or experiences, and how to govern the boundaries when multiple roles coexist. As AI makes relationships designable, brand architecture becomes, more than ever, relationship architecture. Decisions about which relationships can share a portfolio should be based on a brand's perceived motive: Roles reinforce one another when the intentions customers read into them cohere, and roles corrode one another when those intentions conflict. In other words a customer who senses two divided motives will doubt both.

...

The deeper shift, especially for AI-native companies, is not efficiency but relational intelligence: Once AI that is naturally conversational becomes the primary way customers experience an organization, they won't separate the chatbot from the brand. Every conversation will shape what they believe the company stands for, what promises it intends to keep, and whose interests it ultimately serves. That means decisions that once felt purely technical, such as how much a chatbot remembers, when it hands customers to a human, or how it responds in difficult moments, will increasingly be part of a relational strategy—ideally from the inception of the brand—that shapes trust and loyalty.

Over time that will change what leadership looks like. Managing customer relationships won't fall only to marketing or customer service. Product managers, engineers, designers, lawyers, compliance teams, and executives will all influence the relationship customers have with the company, because every update will potentially change how that relationship feels. The organizations that get this right won't just build better chatbots. They'll redefine what it means to build a trusted brand.

A version of this article appeared in the forthcoming November–December 2026 issue of *Harvard Business Review.*

Julian De Freitas is an associate professor in the marketing unit at Harvard Business School.

Read more on **AI and machine learning** or related topics **Brand management**, **Customer experience**, **Marketing** and **Consumer behavior**