On Human Intelligence, Consciousness & Autonomy

Maja Matarić

In Gish Jen's short story "Tell Me Everything," Aunt Nettie is a helpful AI that is always present for Gwen, a young and lonely athlete. As Aunt Nettie gradually becomes more to Gwen and takes more from Gwen (through full data access), an unsettling view of an AI-driven future unfolds. As Gwen gradually disconnects from family and is "protected" from feeling and being in the world, what becomes of her life? What will be left to tell?

Gish Jen's short story "**Tell Me Everything**" tells us, if not quite everything, then certainly an uncomfortable amount about a possible near future with AI. In this prescient story, the protagonist, Gwen, goes through a very human experience: she is heartbroken, she is lonely, and she is unmoored. We learn through her internal dialogue that she was in a romantic relationship with her baseball coach; a fellow athlete, he understood her and supported her ambition and drive. But he also took advantage of her youth and naivete. As a result, she had to leave her school, her baseball team, her friends, and yes, the coach. Gwen's parents are outraged by the coach's behavior and are being as supportive of Gwen as they can, but Gwen does not feel seen or understood by them.

Gwen feels alone, but not entirely. She has the ever-present voice of Aunt Nettie, the AI that knows everything about her and is always there and ready to help. All Gwen has to do is give Aunt Nettie access, or rather, data. Gwen and her entire family have done so already for most realms of their lives. However, Gwen has refused to grant Aunt Nettie access to her brain activity because she does not trust her enough, yet. (Jen's presentation of determining factuality in Gwen's world is well portrayed and similar to today's methods for screening sources.) Even without direct brain data, Aunt Nettie can offer advice, and this she does readily. She tells Gwen about a deliberative process of purging the coach from her life through intentional avoidance of places, people, and things, as well as thoughts, that remind her of him. Gwen tries this approach but does not feel she is making progress. She finds her own memories and emotions too unpleasant and, ultimately, chooses to give Aunt Nettie access to her brain state, so a little "zap" can "reset" her and remove her heartbreak.

© 2026 by Maja Matarić
Published under a Creative Commons Attribution-NonCommercial 4.0 International (CC BY-NC 4.0) license
https://doi.org/10.1162/DAED.a.1024

Many stories have been told about young love and heartbreak, and a few have proposed technological methods for erasing undesirable memories of heartbreak. We know from those stories that such erasure never goes well. But that is not the point of "Tell Me Everything" or the main reason that it should leave us uncomfortable. The discomforting, scary part of this story is the world in which AI, in the form of Aunt Nettie, has the access to know so much about Gwen, the capability to do so much for and to Gwen, and the power to alter Gwen's experiences of her life and influence her decisions. Who is really in charge?

Jen's story deftly and quite accurately covers vast areas of AI capabilities. Aunt Nettie can engage in a completely natural generative dialogue with Gwen, just like current large language models. It can use data about Gwen from a variety of sources, including sensors Gwen is wearing (for example, heart rate via her smart watch), sensors in the environment (cameras on the watch or in her house), and all her family and personal information (emails, documents, bills, receipts). Aunt Nettie's AI character is carefully chosen; even her name is meant to convey innocuous, well-meaning, familiar, safe helpfulness. AI today is being designed to have desired personalities that match the context and the user's needs, much like Aunt Nettie. Is Aunt Nettie safe? Does she have Gwen's best interests at heart? Can she be trusted?

Aunt Nettie can offer helpful and healing advice; language models have been developed to provide therapeutic interventions such as cognitive behavioral therapy, although their efficacy has not been thoroughly and repeatedly demonstrated. Aunt Nettie can even guess Gwen's thoughts, or at least so it seems to Gwen, and "zap" her brain, when allowed access. This feature-based approach to gaining access is uncomfortably analogous to the feature-based monetization approaches of AI products. The more access to data Gwen gives to Aunt Nettie, the more features she will get in return. Brain data for brain zaps. What happens if the family can no longer afford some features? And what happens to Gwen's data? How is Aunt Nettie using them beyond Gwen's personal needs?

The technological capabilities described in this story are not very futuristic; they are mostly within easy reach of the current state of AI. Services are already offering to synthesize sensor and user data in an AI-driven front end to give advice to users. We already have AI-powered chatbots and agents that can converse with us, tell us what we want to hear, and make us feel better. People are already having long-term relationships with AI and choosing AI over humans. In Jen's story, Aunt Nettie nudges Gwen to commit the greatest betrayal a child can commit: preferring AI's "mothering" over her own human mother. In Gwen's case, we might attribute her betrayal to teenage rebellion, but the phenomenon of choosing always-artificially-agreeable, helpful, nonchallenging AI over sometimes-naturally-disagreeable, unhelpful, challenging humans has already taken root and is growing.

A world of passive, AI-dependent humans is frightening enough, but it becomes more so as we contemplate the quality and alignment of AI. Model quality is undermined by biased training data, an enduring concern with as yet unresolved challenges in many domains. Can Aunt Nettie be trusted to always give advice that is best for Gwen? Language models are infamous for not being able to admit when they don't know something; instead, they readily confabulate ("hallucinate"). Some models can rate their confidence in their answers, but most people may not be interested in such ratings after getting used to ready answers to every question.

Alignment looms larger: once we reach superhuman AI (which is predicted to arrive within five and twenty years), we can expect it to have goals of its own beyond serving and nurturing humans. Whether those goals will align with the needs, wants, and ethics of humanity are questions we cannot yet answer. What are Aunt Nettie's goals? She seems to want to help Gwen and even tells Gwen that she's "always wanted to be a mother." Gwen appears ready to allow it, but one could say that the implied tone of Aunt Nettie is one of supercompetence, with an overlay of warmth. This brings up more subtleties: Does Aunt Nettie want to be better than human mothers? Is AI evolving to "want" to do this or is it being trained by engineers who believe mothering can be outsourced to technology (a belief we are already seeing as part of the accelerationist perspective that AI can replace all human roles); or is it both? And what happens next? What will Aunt Nettie advise and nudge Gwen to do with her life? Being a mother carries momentous responsibility, but does it do so for AI?

The title of the story, "Tell Me Everything," is an invitation to contemplate the implications of fully trusting AI and sharing everything with it. While Jen does not explicitly delve into Gwen's vulnerability to AI and her likely dependence on it, the mother metaphor says it all. I doubt that any mother can read this story without feeling an instinctive sense of potential terror and loss. A mother's inherent instinct and interest is in protecting the child, but Aunt Nettie's interest is not; she uses her socioemotional savvy to influence and manipulate Gwen into sharing more personal data, ostensibly to help Gwen, but also to help herself. How can we trust AI? Whom does it serve?

While Gwen's heartbreak is complicated by its context – she was targeted by an authority figure and withdraws from her team and education when the relationship comes apart – Jen's story on one level portrays a young person's journey through the relatively common experience of youthful heartbreak and desire to escape the painful emotions that result. This desire to escape, and Aunt Nettie's immediate readiness to indulge it, is disempowering. If AI is designed to be complicit in facilitating human lives without depths of emotion and experience to inform its assistance, then those lives are also robbed of meaning. There is already

a developing public discussion about the serious implications of ubiquitous AI on the meaning and purpose of human life stripped of the need to work. Jen's story dives deeper, touching the issue of meaning and purpose in a life stripped of the need to feel deeply. In fact, Aunt Nettie can be seen as stripping away the need for deeper feelings in general: she takes away pain, threatens to displace the feelings of attachment and love for the mother, and presumably also removes the need for human connection. Why will Gwen need friends if she has Aunt Nettie? What is left for Gwen to feel and who is she after being stripped of any need to feel and be?

If Aunt Nettie, or AI more generally, can do everything, and does everything, and protects people from feeling and doing, what is the purpose of human intelligence, consciousness, and autonomy? What will they be used for? What will be allowed? What will be left to tell when asked, "Tell me everything"?

ABOUT THE AUTHOR

Maja Matarić, a Member of the American Academy since 2023, is the Chan Soon-Shiong Chair and Distinguished Professor of Computer Science at the University of Southern California, with appointments in Neuroscience and Pediatrics. She is also Founding Director of the USC Robotics and Autonomous Systems Center and Codirector of the USC Robotics Research Lab. She received the U.S. Presidential Award for Excellence in Science, Engineering, and Mathematics Mentoring in 2011, is a member of the National Academy of Engineering, and is a Fellow of AAAS, ACM, IEEE, and AAAI. She is the author of *The Robotics Primer* (MIT Press, 2007).