How will knowing, or not knowing, whether an agent is a human or a machine influence our interactions?

BY DAVID HAREL AND ASSAF MARRON

The Human-or-Machine Issue: Turing-Inspired Reflections on an Everyday Matter

ALAN TURING'S 1950 paper[37] introduced the famed "imitation game" as a means of determining whether a computer can be considered intelligent, thus informing the definition of machine intelligence. Over the years, the Turing test has been the subject of analysis and discussion, resulting in several variants, and has been reflected upon in retrospective reviews (see, for example, French[10]). Similar tests have been proposed in quite different areas, including automotive, games, urban and industrial planning, biological and biochemical modeling, and odor reproduction. The purposes of such variant tests range from offering practical techniques to discern an agent's identity to serving as a norm, or yardstick, for assessing the quality and fidelity of a model or reproduction process in mirroring the original's properties (see, for example, Harel[11]).

Here, we completely sidestep the issue of defining or measuring intelligence, as well as the practical question of whether a machine can be built to replace, or mimic, a person in the performance of some specific task.[33]

Instead, we look more broadly at a concept that we term the *human-or-machine issue (H-or-M issue)*. In a future world where, in some interactions, machines will be able to impressively mimic humans, new social, psychological, functional, and technical issues are bound to become relevant. For example:

▸ Will humans care whether the agent they interact with is a human or a machine, and if yes, why?

▸ How will a person's behavior or emotional state differ between interactions with another human and interactions with a machine whose behavior is indistinguishable from a human's?

▸ How will the answer to the question of an agent's human-or-machine identity (hereafter, the *H-or-M question*) be elicited?

▸ Will human language and social practices change when machines can adequately mimic humans?

▸ Will machine-machine interactions change when the behavior of one or both of the participants is very close to a human's?

» **key insights**

▪ **In a world where machines mimic humans, several research questions about interactions arise that are relevant to system engineering and psychology. These questions concern everyday real-world situations and are quite different from the Turing test's focus on defining machine intelligence based on the ability to pass a controlled test. They include:**

▪ **How will a person's behavior and emotions differ when interacting with a human-like machine through text, voice, or video?**

▪ **How should the design of human-computer interaction accommodate such differences?**

▪ **Will people care about whether they are interacting with a human or with a machine, and will they try to discover the agent's identity? How will human agents react to such attempts?**

▸ Will machines indeed be indistinguishable from humans, or will this be a non-issue because openly taking advantage of machine capabilities will be prioritized over manifesting human-like behavior?

In examining these questions, we discuss research, opinions, and predictions about differences between humans and machines, and differences between human-human and human-machine interactions.

Of special concern here is the importance of the H-or-M question itself, which represents the interest, or curiosity, of a person who is engaged in an everyday interaction with an agent and is wondering whether the agent is a human or a machine.

Currently, machines are unable to disguise as humans; however, we predict that in the future this will change dramatically. Key factors will include: (i) pervasive automation of human service and office functions, as in service centers with automated chatbots, healthcare conversational agents, and service robots in stores,[9,22,28,31] especially in view of recent advances in language processing through large language models (LLMs); and (ii) the prevalence of interactions that hide an agent's identity, as in text-only or voice-only interactions, or when one is unable to determine whether a vehicle or a device in the public domain is autonomous or is operated and controlled by a human.[5]

Even when the agent with whom we are interacting appears as clearly a human (or clearly a machine), we may wonder whether each and every step we see is actually controlled by that agent or whether there is another machine (or another human) dictating the text and actions of the agent's interaction.

We are inspired by the Turing test insofar as our focus is on confined human-agent interactions, rather than on the broader issues of the role of new human-like machines in the world, or on forensic issues like whether a non-interactive artifact, such as a picture or a document, was created by a person or a machine. Indeed, the original Turing test and its variants can be viewed as a special case in which answering the H-or-M question is the sole purpose of the interaction, and

the interaction occurs in a highly controlled and rigid setting.

Specifically, the current discussion diverges from discussions of the Turing test as follows:

▸ We concentrate on everyday interactions, rather than on a controlled lab setup. In our setting, the interrogation cannot stray from the intended subject matter of the human-agent interaction, whereas in a Turing test setting the interrogator is free to guide the conversation as he or she sees fit. Note that both in Turing's paper and here, there is no proposal for a concrete interrogation protocol.

▸ We focus on how the agent's H-or-M identity affects the current interaction in possible future encounters too, rather than on whether the answer can or cannot be elicited, whether one can conclude that, in general, machines can mimic humans well, or whether a particular machine can be labeled "intelligent." Our main interest is in learning about human behavior, not in assessing a machine's capabilities.

▸ We are interested in *patterns* of such effects across many interactions between humans and agents, where the agents can mimic humans well, as compared with interactions of such humans with *human* agents. In discussions of the classical Turing test, patterns in the interaction itself across multiple tests are not an issue.

Parts of our discussion are presented as questions, some of which may justify separate, focused research.

The H-or-M issue is presented here as binary. Clearly, there may be mixed modes. For example:

▸ The apparent agent is a human physician, who, while consulting a human patient, relies extensively on online search for information or is informed, openly or discreetly, by an automated agent listening in on the conversation.

▸ The apparent agent is a vehicle, but while in many aspects it acts autonomously, it is also remotely supervised and occasionally even controlled by humans (and we are interested in this vehicle's interactions with other road users).

▸ Setups like the above two, but where the mixed-mode agent depends on more than one machine and/or more than one human.

▸ Extending the above one-to-one human-agent interactions to group interactions, and potentially even without a clear delineation of "the agent."

Such mixed-mode agents may be treated by default in the same way as a single pure machine or a pure human. We defer to future work discussion of cases in which the mixed mode is substantially different from the binary one.

Are We Different When We Interact with Machines?

One kind of relevance the H-or-M question might have is on the way in which knowing the answer could affect human behavior during a particular interaction.

Taking a broader perspective, the relationship between actual humans and machines that present themselves as almost human has been explored in a variety of ways in the arts, science, and philosophy. Consider, for example, movies like *The Matrix*, *Blade Runner*, *The Terminator* series, and *Her*, and books like *Machines Like Me* and *I, Robot*.[a] Scientists have also researched human-machine relations (see, for example, Chaturvedi et al.[2] and Reinkemeier and Gnewuch[25]), covering aspects such as gaze, facial expressions, and clothing, and have proposed that the field of sociology should study AI-related issues.[17]

As stated earlier, however, our focus here is on how wondering about and then knowing a particular agent's H-or-M identity will affect the interaction at hand, and whether, in general, it will shape future interactions of humans with agents. Such effects may span many aspects of discourse analysis, including, among others, the actual text, discourse structure, questions vs. monologues, speech acts, vocabulary, discourse length, expression of emotional states, theory of mind, and, for speech interactions, prosody (see, for example, Coulthard and Condlin[4] and Magashoa[19]).

There are many studies of human interactions with chatbots—text-based

---

a   Excluded from this article is a comprehensive summary of how these movies and books present the relations and interactions between humans and human-like machine agents, which we were able to readily obtain with a few queries to OpenAI's ChatGPT.

conversational agents (see, for example, literature reviews in Chiturvedi et al.,[2] Mariani et al.,[18,21,24] and references therein). Research themes include: the analysis of chatbot functionality and its relationship to certain success factors, such as the ability to affect user actions; aspects of the interactions, such as the language used or length of conversation; and human-chatbot relations, such as acceptance and trust. Studies that focus specifically on the differences between humans and human-like machines in normal kinds of interaction (for example, a service robot in a store) are also emerging.[9] In most of that work, the H-or-M question itself is not at the center of the research. In many studies, the fact that the agent is a machine is disclosed up front; in others, the researchers were interested in whether the human users ascribed humanness, or human-like behavior, to a machine agent.

Partly motivated by published research on human-robot and human-computer interaction (HRI and HCI), we provide below some examples of possible differences between everyday interactions among humans, and interactions on the same subjects with machine agents that mimic humans well. One should note, however, that while validating or refuting each such candidate effect on the user's actions or emotional state is an intriguing issue, the examples appear here only to support the main claim of the article: that the H-or-M issue will quickly become relevant in many everyday situations.

Language. Some languages require distinguishing humans from nonhuman agents and, in the case of a human agent, often also identifying their gender. A person conducting a text exchange with a service center may be inclined to use different pronouns or verbs for humans and for machines, both when addressing the representative and when discussing what another representative may have communicated in a prior exchange. Furthermore, special linguistic patterns may evolve for cases where such a determination remains unknown.

Structure and style. As summarized in Rapp et al.,[24] some research on human-chatbot interactions suggests that, when interacting with a machine

> The current emphasis on prompt-writing skills for interacting with LLMs suggests that we will make a stronger effort to explain ourselves knowing that a machine is expected to be more limited than a human in understanding our intentions and needs.

as compared with a human, the human may be briefer and less polite and more inclined to abruptly stop or divert a conversation, or even to employ profanity.[14]

One may wonder if we will be more accepting of a machine agent's formal, dry, or even rude attitude knowing that machines will not normally be considerate and use a more restricted subset of natural language (see, for example, Mu and Sarkar[20]). Similarly, will people be more patient with "stupid" or repeated answers, or with inconsiderate actions, such as when driving behind an overly cautious and slow autonomous vehicle (AV), knowing that machines are limited and their behavior cannot be readily changed? (See, for example, Hidalgo et al.,[13] who write, "[P]eople may expect machines to be rational and people to be human.") We expect that people will be less patient when experiencing delayed responses, expecting response times common to most computer applications.[40]

Theory of mind. When interacting with a new environment, humans often build a mental model of the logic and causalities in that environment in order to plan their interactions.[23] We expect that humans will actively seek such mental models—that is, patterns in the behavior of the conversing agent when the agent is known to be a machine rather than a human—and make more of an effort to relate to those models. This may occur in real time during an interaction or offline, when looking for information about behavior patterns in certain classes of machine agents. See, for example, the great efforts in explainable AI[23] or the pervasiveness of "tips and tricks" for using various software applications, such as how to search for flight tickets without triggering program-driven price hikes.

The current emphasis on prompt-writing and prompt-engineering skills for interacting with LLMs suggests that we will make a stronger effort to explain ourselves knowing that a machine is expected to be more limited than a human in understanding our intentions and needs.[38] Also, will people report a machine's undesired behavior to the agent itself or to its owner or manufacturer, expecting a professional response like that which

follows a bug report from a user, in contrast to, say, directly criticizing another human's driving, which may cause severe repercussions?

Will we learn from or override a machine agent's behavior? Consider, for example, observing autonomous vehicles (AVs) negotiating a certain class of complicated driving scenario differently from the way in which we would have dealt with it. Will we be inclined to mimic the AVs, assuming that much thought and serious design and testing had been carried out to yield such behavior—"following the crowd," as often happens in human-crowd interactions[7]—or will we prefer to make our own decisions, thinking ourselves to be more knowledgeable and experienced than a typical machine.[6]

Emotions and feelings. The issues of trust building, willingness to disclose personal information, and developing a personal relationship with and feeling empathy toward machine agents have all been discussed in the literature.[1,18,24] Some research shows difficulties in these areas, which may be partly related to the agents being perceived as uncanny. Other research has shown a much warmer attitude from users; clearly these effects may evolve with the technology.

What will be the effect of an incorrect determination? For example, will a human agent be offended when they realize that the person they are interacting with thinks they are a machine? How will that person feel when they realize their mistake? How embarrassed or angry will a person become when they realize that the agent (perhaps even a coworker[28]) whom they thought was human, and with whom they have developed a relationship, is actually a machine?

Neutrality toward H-or-M. It would also be interesting to identify areas in which having the answer to the H-or-M question does not noticeably affect human-agent interactions. Would we still be curious about the answer, and if so, why? Will the question arise subconsciously, like the inevitable tendency to try to incorporate gender perception into our first impressions?[34] Or will indifference to an agent's H-or-M identity in some cases affect human-human interaction in other ways?

How embarrassed or angry will a person become when they realize that the agent (perhaps even a coworker) whom they thought was human, and with whom they have developed a relationship, is actually a machine?

Should H-or-M Be Easily Resolvable?

There are numerous studies of the effects that disclosure of information about participants has on the content, manner, and results of interactions. In particular, the issue of anonymity—and conversely, disclosure of information about the agent—is of great interest in a variety of circumstances for human-human (both direct and mediated by machines) and human-machine interactions (see, for example, Lapidot-Lefler and Barak[16]).

Given the relevance of the H-or-M identity of an agent, when and how should this information be made readily available? And should such information be provided once, explicitly, in advance, as is the case with some service chatbots, or perhaps constantly and automatically, as is done with "recording in progress" indicators in phone calls and teleconferences?

Currently, most chatbots disclose the fact that they are machines. Should autonomous vehicles be clearly marked as such? Should autonomous drones be marked differently from remotely controlled ones?[36] Should a human-like receptionist robot be clearly marked as such, in order to not be mistaken for a human? And should interactions with human agents be labeled as such, or should this be the default?

Should there be standards for communicating this H-or-M identity, using, say, text, icons, or spoken words? Should this information also be provided through programming interfaces?

When should the H-or-M question be left for the interested person to answer for themselves, without a dedicated, explicit interface? One context in which this is likely to be the case is when the agent's behavior is clearly a mixed-mode, collaborative operation, partly human and partly machine. The exact division of subtasks may be interesting to humans but may not be readily available.

Excluding cases of deception or oppression, where much more than the H-or-M identity is fraudulently presented, we ask whether there are ethical circumstances in which people will actually *want* the agent's H-or-M identity to be well hidden. Here are a few candidate scenarios:

▸ When the agent's role is to help

train a human user in interacting with other humans, complete with their errors and misunderstandings, as in training aircraft pilots, therapists, dancers, or athletes (see, for example, Sackl et al.,[27] Scassellati et al.,[29] and Taylor et al.[35]).

▸ In human-in-the-loop machine learning, where it may be desirable for the human to not know whether they are training a machine or another human (see, for example, Wu et al.[39]).

▸ In a variety of research situations focused on studying the behaviors of humans and machines (see, for example, Scassellati et al.[29]).

Some Inherent Differences Between Humans and Machines

One cannot delve into the H-or-M issue without considering the essential differences between the behavior of human agents and that of machine agents, in general and specific contexts. Turing himself dedicated a section to such a discussion in his 1950 paper,[37] though clearly some distinctions have changed dramatically over time; for example, in the capability to learn and to adapt to changing conditions.

Interest in this issue continues, with arguments discussing the differences or absence thereof in areas such as intelligence, common sense, memory and learning, cognition, creativity, emotions, social and conversational interaction skills, computational complexity, machine/neurological complexity, dynamical systems and modeling, programmability, ethics and morality, and theology.

Such differences between humans and machines are sometimes phrased as tantalizing goals in achieving artificial intelligence in perception, cognition, and reasoning (see, for example, Sifakis,[32] Russell and Norvig,[26] and Langrebe and Smith[15]) and in achieving a sense of humanness when interacting with machine agents.[1,24,30] Gaining insights into these inherent differences can help in studying their effects on interactions and in designing interrogation strategies.

Deviating from science-driven psychological, biological, and philosophical discussions, below we list some such tentative differences between machines and humans, as they may be identified by typical people in the context of everyday interactions [b] When, during an interaction, challenges associated with such differences arise, they may become indications as to the agent's H-or-M identity (if not already known), and may cause an obvious shift in the flow of the interaction:

▸ *Free will:* Machines are completely preprogrammed, whereas humans have free will.

▸ *Emotions*: Humans have emotions and feel compassion, pain, and more, whereas machines do not.

▸ *Context awareness*: Humans are sensitive to context and to innumerable explicit and tacit inputs, to which a typical machine is blind.

▸ *Common sense and worldly familiarity*: A human has more common sense and knowledge with regard to relations between entities and cause-and-effect patterns in the world than any single average machine.

▸ *Narrow specialties*: We expect a human's expertise to be focused in only a few domains; a machine's knowledge can span vast areas.

▸ *Learning and adaptivity*: Turing claimed that humans retain both long- and short-term memory and learn from them, and machines often do not. These days, however, the opposite might be the case. Machines can be equipped with vast memories and can access voluminous repositories of data, to which they can then apply powerful machine learning algorithms, whereas humans' capabilities are more limited. Still, one may say that humans can adapt to new conditions and demands and learn to perform new tasks faster than machines.

▸ *Collaboration*: Machines may demonstrate more efficient and more consistent collaboration than humans. For example, car-to-car coordination on a highway is probably easier to implement technologically than establishing such coordination among human road users. The use of the idiom "like a well-oiled machine" to describe the operation of

---

b We follow examples of such distinctions between scientific thoughts and person-on-the-street opinions on current pressing issues like the environment or vaccinations; in the absence of readily available surveys, we derive this view from early depictions in literature and film as well as published discussions of AI that precede the development or invention of deep learning and large language models.

an efficient human organization hints at our intuition in this regard.

▸ *Mistakes:* Humans make more mistakes than machines.

▸ *Diversity:* Human behavior involves more randomness and arbitrary actions and is less predictable than that of machines. Different humans working on the same task therefore exhibit more diversity than different machines of the same model working on the same task. Similarly, the performance of a human repeating a given task is more diverse than that of a machine repeating the task.

Besides disclosing the nature of the agent or redirecting the interaction, finer understanding of these differences may help in bridging them, by endowing machines with certain desired human capabilities and, to a lesser extent, vice versa.

Some Aspects of H-or-M Interrogation

As stated in the introduction, we do not seek here a strategy or protocol for eliciting the answer to the H-or-M question in everyday situations. Still, it is worthwhile to briefly discuss some relevant issues, and hint at some tentative ingredients of a potential strategy. Such issues, including participant identity, context, structure and protocol, atmosphere, duration, verbal vs. nonverbal cues, and levels of participation, have been discussed in studies of diverse types of interaction such as small talk, job interviews, research interviews, law enforcement interrogations, and HCI (see, for example, DiCicco-Bloom and Crabtree[8]).

Participation and roles. The classical Turing test is a true interrogation: Only the interrogator is proactive and in control of the interaction; the agent is expected to merely react to the inquiries and statements coming its way. General H-or-M inquiries will have to accommodate different positioning.

Verbal vs. nonverbal. Some variants of the Turing test are nonverbal in nature (see, for example, Ciardo et al.[3]). The interrogator challenges the agent to act in certain ways, and then analyzes the resulting behavior, including seeking patterns therein. However, in this case too, the entire exchange is orchestrated as an interrogation.

Related interrogation techniques

can be found in captcha challenges that are built around a cognitive task and in human-driven interrogations in contexts ranging from psychiatric therapy[12] to reasoning about drone behavior.[36]

Contents and nature of interaction. When a person is interacting with a service center, the conversation is expected to be focused on the service issue at hand, rather than on unmasking the agent's H-or-M identity. If the person is interested in this information, and the agent does not directly disclose it, the person can derive it only from the agent's communications on the service issue. Similarly, a human driver who observes the nonverbal behavior of a nearby vehicle and is interested in determining whether it is autonomous has to suffice with passive observation and ordinary road behavior, such as passing the vehicle in question.

This leads to another aspect of interrogation: What is the medium or channel of interacting with the agent? Clearly, even just seeing the agent in action may provide some relevant clues. Hearing is another important channel. The classical Turing test is constrained to typewritten textual interaction. However, while this limitation seems appropriate for achieving fairness—since it masks gender differences between human speakers and overcomes technological constraints in speech synthesis—it robs the interrogation of the emotional elements found in speech prosody. This could be appropriate for testing intelligence with less of a focus on emotions, but it may be inappropriate if we are interested in the H-or-M question in interactions that normally involve speech. The same may also apply to interactions where agent actions could involve touch, smell, and possibly even taste.

What about other kinds of physical interactions? Can an interrogator ask for the results of a blood test from an agent? We leave such "limitations of imitation" to a future discussion.

Will H-or-M interrogation practices disappear or become routine? With advances in machine capabilities and use, we expect the importance of the H-or-M issue to increase over time and that techniques will evolve for eliciting this information from matter-of-fact verbal or nonverbal interactions. These techniques may be crafted from the knowledge about distinctions between humans and machines or may evolve naturally or subconsciously, leading to further understanding of such distinctions. Development of such techniques may be supported by sharing historical information about interactions and interrogation results.

The ability to discern humans from machines may even become an algorithmic/computational thinking skill, perhaps even a "required" social skill. Furthermore, if the techniques can be formalized, we may see automated tools that assist in such delicate interrogations. And, if such interrogations become routine, will humans and machines eventually learn to detect them? Such detection could trigger direct responses in order to save time and effort, or perhaps drive redoubled efforts to conceal the answer. Would a human agent be offended if they noticed that the person they were interacting with is not sure that they are indeed human? Will people use such interrogation to tease agents, or perhaps to hint that an agent's behavior is too rigid?

H-or-M interrogation and society. Finally, it is possible that while the H-or-M issue will become highly relevant, no specific effective interrogation protocols will evolve in the foreseeable future. In fact, social norms or judicial regulations may result in a practice of routinely disclosing an agent's H-or-M identity. Moreover, in some contexts, people may just learn to live with not knowing and not asking, as is the case when the gender of one's counterpart in a text-only interaction is unknown (although gender is known to be a primary component of first impressions; see, for example, Signorella[34]).

Technological deficiencies in mimicking humans may render the entire issue moot; conversely, technological superiority over human performance in key aspects of the interaction may cause developers to forego the effort to mimic humans in secondary aspects. Human agents in roles that are also fulfilled by machines may limit their own behaviors to the purely professional and bureaucratic ones, thus mimicking machines and reducing the advantages (or the significance of the differences) of interacting with a human. Or humans in such roles may emphasize behaviors that disclose their being human. Finally, it is possible that while interrogation pro-

tocols will be developed, both humans and machines will learn to detect them and avoid playing along, rendering the protocols useless.

Discussion

While the issues and questions we have raised regarding the human-or-machine issue may pique one's curiosity, we may still ask: Why are they interesting now? Why do we want to know now what people will do with answers to the H-or-M question in common interactions? Can't we just wait and see what people will do, for example, when they find out that the agent they thought was a human was really a machine, or vice versa?

Better understanding of these issues can advance science and technology in many ways. Here are some examples.

First, current HCI design involves a delicate balance between the value of friendly, intuitive, human-oriented behavior (say, by using natural language) and the value of succinctness and predictability (say, using templates and menu-based selections). Understanding how human behavior and expectations differ when interacting with humans and with machines may improve productivity and quality in the development of agents and business processes. For example, if it turns out that people use a certain subset of natural language when interacting with machines, then training agents on that subset may become more efficient than training them on general natural language.

Second, a major factor in rich interactions is trust. Understanding the differences between how trust-building emerges in human-human interaction as compared with human-machine interaction may allow us to better understand this elusive concept and create protocols for enhancing and accelerating trust-building more broadly.

Third, we are all familiar with cartoons depicting people grumbling or getting angry with their computers. For our own well-being, knowing that we are interacting with a machine rather than with a human may require us to channel our own natural emotions differently. System developers are already well aware that certain system behaviors may evoke anger, frustration, and other emotions. Translating such knowledge into design decisions

will become even more complicated when designing agents that mimic humans. While there is a body of research about various aspects of human emotions when interacting with chatbots, the challenge here may be broader, due to the wide variety of types of agents and the fact that a growing portion of one's interactions may eventually be carried out with machines. Research and therapy methods related to this area are already emerging.[2,12]

Fourth, in a world with many disparate autonomous agents, insights into how humans build mental models of a machine's underlying logic may enable enhancements to certain machine-to-machine protocols for the discovery of available interfaces, agents' goals, and collaboration opportunities.

Carrying out research on human interaction with agents who mimic human behavior with high fidelity in common, real-world situations may not be easy at all. Will researchers be able to create the everyday nature of such interactions in a controlled environment? Will lab experiments with a limited number of kinds of machine agents be representative? And, conversely, when collecting data from real-world interactions, will enough ground truth information be available with regard to whether the agents are humans or machines?

In summary, we do not know if intelligent machines in everyday roles will come to be treated as conventional objects, like computers or ATMs, or as different kinds of living species. In the long run and in particular cases, they may even become indistinguishable from human professionals.

However that may turn out, we are convinced that determining whether one is interacting with a machine or with another human is likely to become a central question. The insights to be gained from studying the question and its ramifications may have surprised even Turing.

Acknowledgments

References

1. Cai, D., Li, H., and Law, R. Anthropomorphism and OTA chatbot adoption: A mixed methods study. *J. Travel & Tourism Marketing 39*, 2 (2022), 228–255.
2. Chaturvedi, R., Verma, S., Das, R., and Dwivedi, Y.K. Social companionship with artificial intelligence: Recent trends and future avenues. *Technological Forecasting and Social Change 193*, (2023), 122634.
3. Ciardo, F., De Tommaso, D., and Wykowska, A. Human-like behavioral variability blurs the distinction between a human and a machine in a nonverbal Turing test. *Science Robotics 7*, 68 (2022), eabo1241.
4. Coulthard, M. and Condlin, C.N. *An Introduction to Discourse Analysis.* Routledge, 2014.
5. Cusack, J. How driverless cars will change our world. BBC (Nov. 29, 2021); https://bbc.in/3VbPid5.
6. de Véricourt, F. and Gurkan, H. Is your machine better than you? You may never know. *Management Science*, 2023.
7. Dekel, I. and Shayo, M. Follow the crowd: But who follows, who counteracts, and which crowd? *SSRN*, 2023.
8. DiCicco-Bloom, B. and Crabtree, B.F. The qualitative research interview. *Medical Education 40*, 4 (2006), 314–321.
9. Frank, D.-A. and Otterbring, T. Being seen...by human or machine? Acknowledgment effects on customer responses differ between human and robotic service workers. *Technological Forecasting and Social Change 189*, (2023), 122345.
10. French, R.M. The Turing Test: The first 50 years. *Trends in Cognitive Sciences 4*, 3 (2000), 115–122.
11. Harel, D. A Turing-like test for biological modeling. *Nature Biotechnology 23*, 4 (2005), 495–496.
12. Heiser, J.F., Colby, K.M., Faught, W.S., and Parkison, R.C. Can psychiatrists distinguish a computer simulation of paranoia from the real thing?: The limitations of Turing-like tests as measures of the adequacy of simulations. *J. Psychiatric Research 15*, 3 (1979), 149–162.
13. Hidalgo, C.A. et al. *How Humans Judge Machines.* MIT Press, 2021.
14. Hill, J., Ford, W.R., and Farreras, I.G. Real conversations with artificial intelligence: A comparison between human–human online conversations and human–chatbot conversations. *Computers in Human Behavior 49*, (2015), 245–250.
15. Landgrebe, J. and Smith, B. *Why Machines Will Never Rule the World: Artificial Intelligence without Fear.* Taylor & Francis, 2022.
16. Lapidot-Lefler, N. and Barak, A. Effects of anonymity, invisibility, and lack of eye contact on toxic online disinhibition. *Computers in Human Behavior 28*, 2 (2012), 434–443.
17. Liu, Z. Sociological perspectives on artificial intelligence: A typological reading. *Sociology Compass 15*, 3 (2021), e12851.
18. Mariani, M.M., Hashemi, N., and Wirtz, J. Artificial intelligence empowered conversational agents: A systematic literature review and research agenda. *J. Business Research 161*, (2023), 113838.
19. Mogashoa, T. Understanding critical discourse analysis in qualitative research. *Intern. J. Humanities Social Sciences and Education 1*, 7 (2014), 104–113.
20. Mu, J. and Sarkar, A. Do we need natural language? Exploring restricted language interfaces for complex domains. In *Extended Abstracts of the 2019 CHI Conf. On Human Factors in Computing Systems*, 1–6.
21. Nicolescu, L. and Tudorache, M.T. Human-Computer interaction in customer services: The experience with AI chatbots—A systematic literature review. *Electronics 11*, 10 (2022), 1579.
22. Parmar, P. Health-focused conversational agents in person-centered care: a review of apps. *NPJ Digital Medicine 5*, 1 (2022), 21.
23. Qing, Y., Liu, S., Song, J., and Song, M. A survey on explainable reinforcement learning: Concepts, algorithms, challenges. *Arxiv Preprint Arxiv:2211.06665*, 2022.
24. Rapp, A., Curti, L., and Boldi, A. The human side of human-chatbot interaction: A systematic literature review of ten years of research on text-based chatbots. *Intern. J. Human-Computer Studies 151*, (2021), 102630.
25. Reinkemeier, F. and Gnewuch, U. Match or mismatch? How matching personality and gender between voice assistants and users affects trust in voice commerce. In *Proceedings of the 55th Hawaii Intern. Conf. on System Sciences*, 2022.
26. Russell, S. and Norvig, P. *Prentice Hall Series in Artificial Intelligence.* Prentice Hall, Upper Saddle River, NJ, 2002.
27. Sackl, A. et al. Social robots as coaches: How human-robot interaction posit ively impacts motivation in sports training sessions. In *2022 31st IEEE Intern. Conf. on Robot and Human Interactive Communication.* IEEE, 141–148.
28. Sadeghian, S. and Hassenzahl, M. The "artificial" colleague: Evaluation of work satisfaction in collaboration with non-human coworkers. In *27th Intern. Conf. on Intelligent User Interfaces.* ACM, 2022, 27–35.
29. Scassellati, B., Admoni, H., and Matarić, M. Robots for use in autism research. *Annual Rev. of Biomedical Engineering 14*, (2012), 275–294.
30. Setlur, V. and Tory, M. How do you converse with an analytical chatbot? Revisiting Gricean maxims for designing analytical conversational behavior. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems.* 1–17.
31. Sheth, A. et al. Cognitive services and intelligent chatbots: current perspectives and special issue introduction. *IEEE Internet Computing 23*, 2 (2019), 6–12.
32. Sifakis, J. *Understanding and Changing the World: From Information to Knowledge and Intelligence.* Springer Nature, 2022.
33. Sifakis, J. Testing system intelligence. *Arxiv Preprint Arxiv:2305.11472*, 2023.
34. Signorella, M.L. Remembering gender-related information. *Sex Roles 27*, (1992), 143–156.
35. Taylor, J.L. et al. The effects of information load and speech rate on younger and older aircraft pilots' ability to execute simulated air-traffic controller instructions. *J. Gerontology 49*, 5 (1994), 191–200.
36. Traboulsi, A. and Barbeau, M. A reverse Turing-like test for quad-copters. In *2021 17th Intern. Conf. on Distributed Computing in Sensor Systems (DCOSS).* IEEE, 351–358.
37. Turing, A. Computing machinery and intelligence. *Mind 59*, 236 (1950), 433.
38. White, J. et al. A prompt pattern catalog to enhance prompt engineering with ChatGPT. *Arxiv Preprint Arxiv:2302.11382*, 2023.
39. Wu, X. et al. A survey of human-in-the-loop for machine learning. *Future Generation Computer Systems 135*, (2022), 364–381.
40. Yu, M. et al. Unravelling the relationship between response time and user experience in mobile applications. *Internet Research 30*, 5 (2020), 1353–1382.

David Harel is a professor of computer science at the Weizmann Institute of Science and president of the Israel Academy of Sciences and Humanities. He works on logic and computability, software and systems engineering, systems biology, odor communication, and the analysis of prosody. He invented statecharts and authored the book *Algorithmics: The Spirit of Computing*.

Assaf Marron is a researcher at the Weizmann Institute of Science in the Department of Computer Science and Applied Mathematics working on machine learning, software engineering, scenario-based programming, and biological modeling. Prior to joining the Weizmann Institute, he worked in senior management and technical positions at leading technology companies.

Watch the authors discuss this work in the exclusive *Communications* video. https://cacm.acm.org/videos/human-or-machine-matter