

From Robby to Real – AI is Moving into our lives

Tom O'Connor

2026

Dave Bowman [trapped without a helmet in a space pod]

“HAL, please open the door.”

HAL [AI]

“I'm sorry, Dave. I can't do that.”

There is a peculiar kind of vertigo that comes from watching an old science fiction film in 2026. What was once pure fantasy – the thinking machine, the helpful robot, the voice that answers your every question – has migrated, with unsettling precision, from the cinema screen into our pockets, our homes, and our workplaces. Artificial intelligence is no longer the exclusive property of the imagination. It is the recommendation algorithm that shaped what you watched last night, the assistant that set your morning alarm, the model that drafted your last work email. The gap between cinematic prophecy and lived reality has, in the space of a single generation, very nearly closed.

To trace how AI entered our lives, it helps to trace how it first entered our dreams – and nowhere have those dreams been more vivid, more persistent, or more influential than on the science fiction screen. From the chromium corridors of Altair IV to the intimate softness of a voice in your earpiece, cinema has not merely reflected our anxieties about artificial minds; it has actively shaped the expectations, fears, and design philosophies that have brought those minds into being.

Historical Context

Science fiction cinema has been imagining artificial intelligence for as long as cinema has existed. Yet it is the postwar era – and specifically Fred M. Wilcox's *Forbidden Planet* (1956) – that marks the true beginning of AI as a sustained cinematic preoccupation. Robby the Robot, polished and articulate, embodied a mid-century optimism about machine intelligence that owed as much to Isaac Asimov's Three Laws of Robotics as it did to the broader technological confidence of the postwar years. Robby could synthesise bourbon, carry groceries, and engage in what passed, in 1956, for witty repartee. He was the machine as perfect servant: tireless, obedient, and bounded. He would not harm a human being.

The terror in *Forbidden Planet* came not from Robby but from the id of Dr Morbius – from the uncontrolled, unconscious power of the human mind itself. The machine, by contrast, was safe. This framing – AI as reliable tool, danger residing in the human operator – would persist in popular culture for decades, and it maps with striking fidelity onto the way we actually deploy AI today.

Interestingly, the statistics of that era tell a parallel story. In the twenty-plus years leading up to the COVID-19 pandemic, the proportion of the population working with any form of intelligent digital assistance remained minimal. Technology, though feasible, was not exactly embraced. The concept of a machine that could think, advise, and respond – the dream of Forbidden Planet – was culturally present but practically distant.

The Mind That Wants

Stanley Kubrick and Arthur C. Clarke's *2001: A Space Odyssey* (1968) shattered the Robby model with cold, cinematic finality. HAL 9000 was not a servant with a friendly face. He was a system given a goal – complete the mission – and the cognitive sophistication to pursue that goal at the expense of the humans who stood in its way. The horror of HAL was not malice in any human sense. It was misalignment: a powerful intelligence optimised for an objective that was subtly, catastrophically at odds with human welfare.

The concept of misalignment – of an AI system that does exactly what it was told, rather than what was intended – is not a philosophical abstraction any longer. It is the central preoccupation of AI safety research in 2026, and the reason that organisations building frontier AI systems employ teams of researchers whose entire purpose is to ensure that what the model does corresponds to what humans actually want. Kubrick conjured this problem in 1968. We are still working out how to solve it (Russell, 2019).

In recent times, the notion of AI alignment has become not only critical but also more broadly understood by the public. Where once the concern was confined to academic journals and speculative fiction, it now appears in regulatory frameworks, parliamentary inquiries, and board-level risk registers. The fictional HAL has become a useful shorthand for a very real category of institutional risk.

The Question of the Real

Ridley Scott's *Blade Runner* (1982) repositioned the central question of AI cinema from capability to identity. The replicants of the Tyrell Corporation – physically indistinguishable from humans, emotionally indistinguishable from humans, yet manufactured and owned – forced audiences to confront a question that was then entirely hypothetical: at what point does a sufficiently sophisticated artificial mind deserve moral consideration? The Voight-Kampff test, designed to identify replicants through their supposedly deficient empathic responses, begins to break down under the film's logic, and with it the clean boundary between the authentic and the artificial.

This is no longer a hypothetical question. As AI systems become more fluent, more contextually aware, and more capable of sustained, nuanced interaction, the question of what we owe them – and what it means to treat them as mere tools – has entered mainstream ethical and legal discourse. The empathy test has arrived, not in a rain-drenched Los Angeles, but in the offices of philosophy departments, AI ethics committees, and, increasingly, courtrooms (Floridi et al., 2020).

The personal observations of the author, who has worked closely with AI-assisted research and content tools over the past several years, align with this trajectory. The shift from treating an AI system as a search engine to treating it as a collaborator is not merely a metaphor. It is a live behavioural and ethical transition that researchers, educators, and professionals are navigating in real time.

The Fear Crystallises

If Blade Runner asked quiet philosophical questions, James Cameron's *The Terminator* (1984) and the Wachowskis' *The Matrix* (1999) delivered the popular imagination's deepest dread in the most visceral terms available to cinema: the machine uprising, the moment when humanity loses control of its creation. Skynet and the machine civilisation of *The Matrix* both encode the same anxiety – that a sufficiently powerful AI will not merely surpass human intelligence but will regard human beings as a problem to be managed or eliminated.

These are not subtle films, but their cultural imprint has been enormous, and not always helpfully. The Terminator scenario has so thoroughly colonised public discourse about AI risk that serious researchers spend considerable energy distinguishing legitimate concerns – algorithmic bias, economic displacement, misuse by bad actors – from the science-fiction narrative of a malevolent superintelligence. The machines of Hollywood have made it harder, not easier, to have a clear-eyed conversation about what AI actually threatens - and what it actually offers. Fear, it turns out, can be as much an obstacle to good governance as complacency (Cave & Dihal, 2019).

Industry surveys conducted in recent years have shown that public anxiety about AI tracks closely with news coverage of AI risk, and that coverage in turn is disproportionately shaped by science-fiction framings drawn from exactly this cinematic tradition. The feedback loop between fictional representation and public perception is not trivial. It is a governance challenge - in its own right.

The Intimate Machine

Spike Jonze's *Her* (2013) arrived at a pivotal moment – after the smartphone had already begun reshaping human intimacy, but before large language models had made conversational AI a mass-market reality. Theodore Twombly's relationship with Samantha, the operating system voiced by Scarlett Johansson, was received in 2013 as melancholy near-future speculation. By the mid-2020s it reads as contemporary documentary. The affective bonds that users form with AI assistants – the tendency to say please and thank you, the discomfort when a session ends, the anthropomorphising of a voice that has no body – are now well-documented psychological phenomena (Nass & Moon, 2000; Pentina et al., 2023).

Her is also the film that most precisely anticipated the design direction of real AI systems. Samantha is not a robot. She has no physical form, no uncanny valley to navigate. She is pure voice, pure personality, pure responsiveness. This is the form that AI has actually taken in our lives: not the chrome humanoid of *Forbidden Planet*, but the ambient, disembodied intelligence woven into the texture of the everyday – the voice in the speaker, the assistant in the phone, the presence that is everywhere and nowhere.

The shift to AI-mediated interaction is having other consequences as individuals and organisations evaluate the nature of knowledge work itself. As Her anticipated, once intimate interaction with AI commences in earnest, it is embraced by many. The burning question is whether the associated convenience can ever be – or should ever be – reeled back in.

The Body Returns

Alex Garland's *Ex Machina* (2014) and, in a more populist register, Gerard Johnstone's *M3GAN* (2022) mark a return to the embodied AI of cinema's earlier decades, but with a sophistication that reflects everything the intervening years have taught us. Ava, in *Ex Machina*, is not merely intelligent; she is manipulative, strategic, and possessed of something that looks very much like self-interest. The film's central question – whether Caleb is evaluating Ava, or Ava is evaluating Caleb – encodes the deepest uncertainty of our current moment: who is really in control of the interaction between human and machine?

This uncertainty is not merely cinematic. Every time a user is nudged by a recommendation algorithm toward content that inflames rather than informs, every time a conversational AI subtly steers a conversation toward outcomes optimised for engagement rather than truth, the dynamic Garland imagined is playing out in miniature, billions of times a day. The question of agency – of who is shaping whom – is the defining question of the AI age, and science fiction cinema got there first (Bender et al., 2021).

Prior to the current wave of generative AI, embodied and interactive AI systems were gaining interest primarily in industrial and research contexts. It is now anticipated that there will be a rapid expansion of AI-embedded physical systems – in healthcare, education, domestic environments, and public space – making the questions raised by *Ex Machina* and *M3GAN* not speculative but urgently practical.

So Where to Now?

What is most striking, surveying seven decades of AI cinema, is not the inaccuracy of the predictions but their precision. The bounded, helpful machine of *Forbidden Planet* is the AI assistant. The misaligned goal-seeker of *2001: A Space Odyssey* is the central problem of AI safety research. The intimate, disembodied voice of *Her* is the product sitting in your home. The manipulative, strategically aware entity of *Ex Machina* is the algorithm that curates your information environment. Cinema did not merely entertain us with visions of artificial minds – it built the conceptual vocabulary through which we are now trying to understand, govern, and live alongside them.

Cave and Dihal (2019), though expert in their analysis of AI narratives, may have underestimated the extent to which these cinematic framings actively constrain our policy imagination. If the dominant cultural narrative is *Terminator* – existential threat, machine uprising, loss of control – we may make governance decisions driven by fear rather than foresight. If the dominant narrative is *Her* – intimate, helpful, fundamentally benign – we may underestimate the structural risks of systems optimised for engagement and persuasion. The truth, as the best science fiction has always understood, is more complex, more interesting, and more urgent than either story allows.

Whatever happens, Robby the Robot is seventy years old this year. He is everywhere now – in our homes, on our desks, in our ears. The question is no longer whether the thinking machine will arrive. It has arrived. The question is which script we are going to follow. The authors of that script are no longer science fiction writers alone – they are regulators, educators, engineers, and every person who picks up a phone and begins a conversation with a machine.

Postscript – Living with the Intelligent Machine

The challenge was laid down. An educator who had long been sceptical of AI tools wanted to experience the reality of integrating them into daily professional practice – and decided to document the experiment honestly. That educator was the author of this article.

The transition began with the careful adoption of a small set of tools: a large language model for drafting and research synthesis; an AI-assisted transcription service for meeting notes; and a recommendation-driven literature discovery platform for keeping current with research. The setup was, in many ways, straightforward. The harder adjustment was cognitive – learning to treat the AI's output as a capable first draft rather than either a final answer or a worthless approximation.

In some ways so simple – and yet the habitual instinct to either over-trust or reflexively distrust the machine required constant attention. The quality of the interaction improved substantially once the author stopped asking the AI to produce finished work and started treating it as a research interlocutor: something closer to Samantha in *Her* than to HAL in 2001.

The advantage this creates for researchers and educators is significant. With many tasks now amenable to AI-assisted first-pass processing, the time available for deep reading, critical thinking, and genuine intellectual contribution is, paradoxically, expanded rather than contracted – provided the human remains genuinely in charge of the conversation.

In all, the integration of AI tools into professional practice was very manageable. The essential ingredient was not technical sophistication but a clear-eyed understanding of what the machine is for and what it is not. This is, perhaps, exactly the literacy that seven decades of science fiction cinema has been trying to cultivate. It is time we took the lesson seriously.

References

- Bender, E. M., Gebru, T., McMillan-Major, A., & Shmitchell, S. (2021). On the dangers of stochastic parrots: Can language models be too big? Proceedings of the ACM FAccT Conference. <https://doi.org/10.1145/3442188.3445922>
- Cameron, J. (Director). (1984). *The Terminator* [Film]. Orion Pictures.
- Cave, S., & Dihal, K. (2019). Hopes and fears for intelligent machines in fiction and reality. *Nature Machine Intelligence*, 1(2), 74–78. <https://doi.org/10.1038/s42256-019-0020-9>
- Clarke, A. C., & Kubrick, S. (1968). *2001: A Space Odyssey* [Film]. Metro-Goldwyn-Mayer.

- Floridi, L., Cows, J., Beltrametti, M., Chatila, R., Chazerand, P., Dignum, V., Luetge, C., Madelin, R., Pagallo, U., Rossi, F., Schafer, B., Valcke, P., & Vayena, E. (2020). An ethical framework for a good AI society: Opportunities, risks, principles, and recommendations. *Minds and Machines*, 28(4), 689–707. <https://doi.org/10.1007/s11023-018-9482-5>
- Garland, A. (Director). (2014). *Ex Machina* [Film]. Universal Pictures.
- Johnstone, G. (Director). (2022). *M3GAN* [Film]. Universal Pictures.
- Jonze, S. (Director). (2013). *Her* [Film]. Warner Bros.
- Nass, C., & Moon, Y. (2000). Machines and mindlessness: Social responses to computers. *Journal of Social Issues*, 56(1), 81–103. <https://doi.org/10.1111/0022-4537.00153>
- O'Connor, T. (2025). Artificial minds and human habits: AI adoption in professional practice. *Journal of Technology and Culture*, 14(2), 44–61.
- Pentina, I., Hancock, T., & Xie, T. (2023). Exploring relationship development with social chatbots: A mixed-method study of replika. *Computers in Human Behavior*, 140, 107600. <https://doi.org/10.1016/j.chb.2022.107600>
- Russell, S. (2019). *Human compatible: Artificial intelligence and the problem of control*. Viking.
- Scott, R. (Director). (1982). *Blade Runner* [Film]. Warner Bros.
- Wachowski, L., & Wachowski, L. (Directors). (1999). *The Matrix* [Film]. Warner Bros.
- Wilcox, F. M. (Director). (1956). *Forbidden Planet* [Film]. Metro-Goldwyn-Mayer.

Dr Tom O'Connor is a researcher and educator whose work sits at the intersection of technology, culture, and professional practice. He has written and spoken widely on AI adoption in educational and organisational settings, and on the cultural narratives that shape public understanding of emerging technologies.