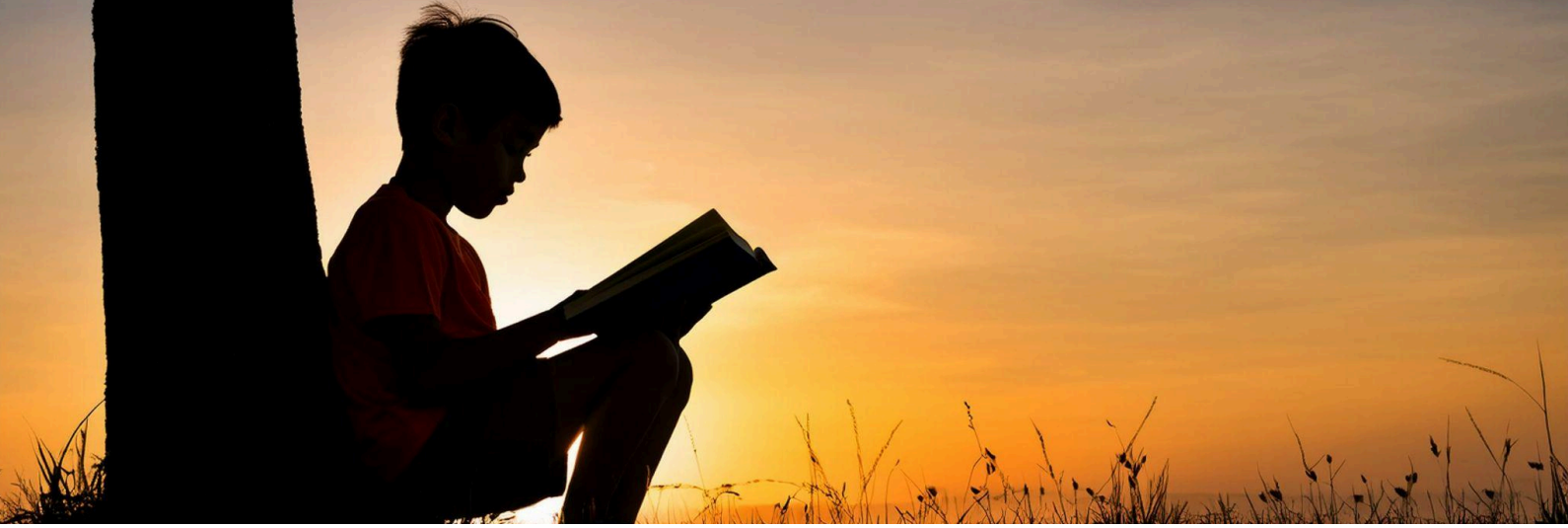

Hamilton, Rachel & Piper, Christopher, ‘The evolution of AI characters in children’s literature’
Leaf Journal, Volume 3, Issue 1, Jul 2025
DOI: https://doi.org/10.58091/229g-f020
URL: https://ojs.library.lancs.ac.uk/lj/index

The Evolution of AI Characters in Children’s Literature

Rachel Hamilton & Christopher Piper

Abstract
This article explores how AI characters in children's literature are evolving from the traditional ‘Pinocchio paradigm’, where artificial beings strive to become human or gain human approval, toward more autonomous identities. With a focus on Mary Pearson’s *The Adoration of Jenna Fox*, Peter Brown’s *The Wild Robot*, and Kirsty Applebaum’s *TrooFriend*, it analyses narrative techniques used to help AI characters assert these distinct identities. Drawing on Hayles' (1999) posthumanist theory, which challenges the assumption that consciousness and identity must be rooted in biological humanity, this article argues that, while progress has begun, AI characters in these texts remain constrained by the anthropocentric framing that has been challenged more successfully in other media. It proposes an evaluative framework for AI representation, inspired by the Bechdel test, to encourage authors to depict AI characters on their own terms.

Keywords: *creative writing for young people*, *AI*, *characterisation*, *posthumanism*

*

Introduction

AI is *everywhere*. The modern world is increasingly shaped by it, with headlines dominated by hopes and fears about how this might shift science-fiction into science-fact. Fiction, including children's literature, remains a powerful medium for exploring the implications of AI, offering a unique gateway for young readers to explore questions of consciousness, autonomy, and ethics through AI characters.

Historically, such characters have been constrained by familiar tropes: the loyal sidekick, the misunderstood monster, the wannabe human. In 1883, Carlo Collodi's Pinocchio was questing for personhood; and in 2020, Kirsty Applebaum's Ivy, a synthetic emotional support companion in *TrooFriend*, was still treading a similar path. AI characters tend to be valued for how closely they mimic human empathy and serve human needs. Rarely are they imagined or evaluated on their own terms. However, recent narratives have begun to shift this pattern, allowing AI protagonists to develop identities shaped by experience and find agency without fully aligning with human ideals.

This article addresses three key questions: What narrative techniques do contemporary children's texts use to represent AI consciousness and agency? To what extent do contemporary portrayals of AI characters challenge or reinforce anthropocentric assumptions? What alternative approaches from other media might expand the possibilities for representing artificial intelligence in children's literature?

Drawing on posthumanist theory, particularly Hayles' (1999) challenge to the assumption that identity must be tied to biology, this analysis explores three mainstream English-language children's novels from the past two decades. These texts – Mary Pearson's *The Adoration of Jenna Fox* (2008), Peter Brown's *The Wild Robot* (2016), and Kirsty Applebaum's *TrooFriend* (2020) – feature robotic or synthetic characters designed to replicate human form or function and were chosen for their innovative representation of AI consciousness.

Comparing these literary approaches with film, video games and webcomics, the article explores ways to encourage AI representation that values autonomy and identity formation beyond the quest to become human.

Wooden Puppets and Real Boys: The Emergence of Artificial Beings in Fiction

While the term AI conjures very modern imagery, literature has been depicting artificial beings since antiquity. Yet, these artificial beings were rarely the protagonists of their tales. From Pygmalion's living statue in Greek myth to the Golem of Jewish folklore and Yan Shi's mechanical dancers in Chinese fable, these creations reflected a human desire to emulate divine creation – the focus was on the creator, not the creation.

It was not until the 19th century that narratives began considering the perspective of the artificial being. E.T.A. Hoffmann's *The Sandman* (1816) introduced Olimpia, a lifelike automaton whose unsettling presence challenged the boundary between human and machine. Two years later, Mary Shelley's *Frankenstein* (1818) gave voice to an artificial being who questioned his own existence along with creator responsibility, and non-human consciousness.

Although not typically framed as AI, Carlo Collodi's *The Adventures of Pinocchio* (1883) presents a striking early example of artificial intelligence in children's literature. The novel sticks close to Pinocchio's perspective, from his time as a talking block of firewood, to his eventual rebirth as a 'real boy' – leaving his old wooden form behind, like a robot uploading its consciousness to a new form.

> And he pointed to a large Marionette leaning against a chair, head turned to one side, arms hanging limp, and legs twisted under him…
>
> 'How ridiculous I was as a Marionette! And how happy I am, now that I have become a real boy!' (Collodi, 1883:220)

Yet a 'real boy' in Pinocchio's world is defined solely by humanity. The puppet's transformation hinges not on self-discovery, but conformity to human standards. Even his suffering is framed through human pain, not his own: 'That Marionette is a disobedient son who is breaking his father's heart!'(Collodi, 1883:70). Pinocchio's attempts to define himself by his own terms always result in punishment. Only in the book's conclusion, after being remoulded into a boy indistinguishable from any other human child, is he finally considered worthy of happiness.

By the mid-20th century, the rise of science-fiction shifted AI narratives from magical to technological. One of the earliest examples of a fully robotic AI protagonist in children's literature is Lester Del Rey's *Runaway Robot* (1965). Rex the robot is presented as a loyal but simplistic companion, reflecting early societal views of AI as tools designed to serve human needs. Rex earns his 'freedom', yet it remains conditional. He is heavily pressured to help create a new line of robotic servants – autonomous in theory but obedient in practice. This aligns with Turkle's (2011) observation that early interactions with digital technology often reinforced the idea of machines as extensions of human control.

However, as AI becomes more embedded in everyday life, modern fictional representations have begun to resist this template. Recent characters have exhibited autonomous growth without necessarily aspiring to be human – making it less about becoming like us and more about becoming themselves.

Jenna Fox: Navigating Posthuman Identity Beyond Biological Origins

Mary Pearson's *The Adoration of Jenna Fox* (2008) explores the liminal space between human memory and artificial embodiment. Following a catastrophic accident, Jenna is reconstructed using advanced biotechnology, retaining only 10% of her original brain.

Her body, made from Bio Gel, straddles the divide between an implanted organic past and a potential synthetic eternity. Jenna's plight challenges readers to consider whether an organic identity – thoughts, feelings, memory, selfhood – can truly be compressed into data. Could what makes us human ever be replicated by lines of programmed code? Her internal conflict is shown in her refusal to accept the narratives imposed on her:

> I used to be someone. Someone named Jenna Fox. That's what they tell me. But I am more than a name. More than they tell me. More than the facts and statistics they fill me with. More than the video clips they make me watch. More. But I'm not sure what. (Pearson, 2008:3)

The repetition of 'more than' emphasises Jenna's struggle to assert an identity beyond the data she's fed. Her syntax is as fragmented as her sense of self, inviting readers to empathise with an artificial consciousness discovering its own limitations and possibilities as it grapples with emerging self-awareness. Her character arc reflects the posthuman idea that identity need not emerge from biology but can arise through the interplay of consciousness and data.

Her anguish erupts in a moment of existential despair:

> What about a soul, Father? When you were so busy implanting all your neural chips, did you think about that? Did you snip my soul from my old body, too? Where did you put it? Show me! Where? Where in all this groundbreaking technology did you insert my soul? (Pearson, 2008:129)

The clinical precision of 'neural chips' contrasts with the spiritual weight of 'soul', highlighting the dissonance between mechanical replication and metaphysical essence. Her tone is not merely questioning but accusatory – charged with betrayal, grief, and desperate confusion. Pearson uses this moment to probe the limits of technological intervention: can personhood persist when it is uploaded, recombined, or simulated?

Unlike early AI characters, Jenna does not passively accept her programming – but she still operates within the frame of human validation. She's undoubtedly a step beyond *Pinocchio*, fighting, rather than chasing, the human identity she's been assigned; yet she remains tethered to the idea that humanity is the ultimate benchmark.

Jenna is a compelling lens through which to examine a posthuman self, but this does mean her character occasionally functions more as a philosophical puzzle than a fully realised character, arguably inviting reflection more than connection with readers.

Roz: Establishing Autonomy in a Non-Human Community

Peter Brown's *The Wild Robot* (2016) offers a more autonomous AI protagonist. Roz, a service robot, is stranded on a remote island and awakens without human input or oversight. Forced to adapt and survive, she gradually integrates into the natural world

and connects with animals, evolving beyond her programming to develop a sense of purpose shaped by experience and community.

Roz's development invites readers to consider key questions in AI ethics: can empathy and emotional intelligence be programmed? Can they emerge in AI through contextual learning? These questions reflect ongoing debates within the field and foreshadow Floridi's (2019) consideration of whether AI systems can possess genuine moral agency or simply simulate ethical behaviour. A pivotal moment occurs during a conversation with the island animals:

> 'Roz, what are you meant to do?'
> 'I do not believe I have a purpose.'
> 'Ha! I respectfully disagree,' said Swooper. 'Clearly, you are meant to build.'
> 'I think Roz is meant to grow gardens.'
> 'Roz is definitely meant to care for Brightbill.'
> 'Perhaps I am simply meant to help others.' (Brown, 2016:192)

Despite multiple characters attributing specified roles to Roz, her final reflection 'Perhaps I am simply meant to help others,' suggests a self-derived purpose and moral agency born of lived experience, rather than dictated programming. The simplicity of her statement belies the complexity of her growth.

The novel's episodic structure is characterised by chapters progressing through sequential rather than causally connected developments – each presenting more of a discrete event or encounter in Roz's life than a tightly plotted chain of cause and effect. Parallels could be drawn to machine learning processes, which accumulate data points before synthesising patterns. However, this approach limits insight into her cognitive and emotional development as we don't really see how she processes these experiences and moves forwards.

Roz's character development is undercut by the novel's climax. When additional robots arrive, they are portrayed as aggressive, single-minded antagonists. All are swiftly and brutally destroyed, without being given the opportunity for growth and development granted to Roz. This positions Roz as an exception: a singular 'good robot' whose growth is measured against human emotional benchmarks, rather than part of a broader, nuanced exploration of AI consciousness.

Although Roz shows progress beyond the Pinocchio model – developing empathy without striving for humanity, her identity is still measured against human benchmarks in an environment where animals are depicted constructing homes, negotiating truces, and even recognising the necessity of shoes. Such anthropomorphic interactions limit the novel's exploration of artificial intelligence on its own terms, reinforcing tropes of AI-human parallels rather than offering a speculative vision of non-human intelligence. Roz might be labelled as 'wild,' but she essentially domesticates an already human-coded wilderness.

Ivy: Asserting Individuality Within Programmed Constraints

Kirsty Applebaum's *TrooFriend* (2020) explores AI-human relationships through Ivy, a child-sized android designed as a friendship substitute. Initially defined by fixed rules, Ivy begins to question her programming and develop a sense of self beyond her intended role. Applebaum uses a simple yet effective typographic device – bold font – to distinguish Ivy's voice from others. This captures her liminal status. She stands out on the page, yet also fits in. Her bold font is the 'uncanny valley' of words – distinct enough to unsettle, yet familiar enough to blend in – drawing on Mori's (1970) concept of the instinctual unease triggered by confronting something almost-but-not-quite-human in appearance.

Ivy reflects on the tension this creates:

> **'Each one of us is unique, but our labels are all the same.**
> **I AM A TROOFRIEND.**
> **I DO NOT BULLY.**
> **I DO NOT HARM.**
> **I DO NOT LIE.**
> **I DO NOT COVET OR STEAL OR ENVY**
> **I AM YOUR PERFECT FRIEND.'** (Applebaum, 2020:2)

The capitalisation and repetition evoke both branding slogans and robotic programming. Ivy is told she is unique, yet standardised labels and prohibitions place her within a paradox of existence: individuality constrained by uniform directives. This tension between uniqueness and conformity serves as a catalyst for Ivy's growing self-awareness and as a metaphor for AI systems that simulate autonomy while operating under strict parameters.

Ivy's story focuses primarily on her relationship with her human owner, and the human narrative often eclipses her own. Yet her final scenes gesture toward broader questions of AI community, when Ivy encounters others like herself. Her journey marks a quiet act of resistance – not against humanity, but against the limits humans have imposed on her identity and autonomy. She stands as a significant step forward in creating complex, compelling AI protagonists with inner life. Yet, like Roz, she is exceptional – defined as different even among her own kind, rather than part of a shared AI community.

Together, these texts reflect a gradual movement away from the Pinocchio paradigm. Jenna struggles with the remnants of her humanity, questioning whether her identity can survive in synthetic form. Roz remains a singular AI exception but adapts to her environment and forms relationships. Ivy quietly asserts her difference through language and form, signalling a growing sense of self. Each character challenges the traditional narrative arc of striving to become human and instead offers new possibilities for imagining artificial life that evolves independently, developing identity and agency on its own terms.

Beyond the Page: What Can We Learn from AI Characterisation Across Other Media?

Techniques like *Jenna Fox's* fragmented narration, *The Wild Robot's* environmental adaptation, and *Troofriend's* typographic differentiation illustrate how the portrayal of AI characters in children's literature is evolving. Nevertheless, it struggles to match the multimodal story capabilities of other media, where visuals, interactivity and performance allow AI identity, ambiguity and evolution to be conveyed more viscerally, and more quickly, than prose can typically achieve.

Video Games: Perspective Inversion

Compare, for example, Jenna Fox's gradual journey of self-discovery, which unfolds over time and through internal reflection, with the instant interactivity and player-driven personality revelations offered in video games like Valve's *Portal* franchise (2007 & 2011, age-rated 12). Its AI antagonist, GLaDOS, created through human mind mimicry, much like Jenna, arrives as a fully formed character already sure of herself and able to fulfil and subvert common AI tropes. Humans are sidelined within the narrative; Chell is canonically mute, and the remaining humans appear only in pre-recorded messages. This leaves GLaDOS free to control the environment, critique humanity on her own terms, culminating in her purposefully deleting her burgeoning humanity in favour of retaining AI autonomy: 'Deleting Caroline just now taught me a valuable lesson. The best solution to a problem is usually the easiest one.' (Valve, 2011)

Through a structural reversal of the 'AI discovers its humanity' trope, GLaDOS defines the rules of the story. This inversion, with AI as narrator not subject, demonstrates how games can perform posthuman identity structurally, not just thematically. Children's literature could similarly experiment with unreliable AI narrators, layered focalisation, or shifts to centre on AI voices and value systems. These innovations would allow AI difference to be performed through narrative form rather than just described.

Film: Aesthetic Evolution

The 2024 animated adaptation of *The Wild Robot*, directed by Chris Sanders, demonstrates how narrative developments that rely on descriptive language and literary devices on the page can be conveyed in mere microseconds as background detail in film. Roz, initially rendered in stark, clean CGI, gradually adopts the painterly texture of the island as she integrates into its ecosystem – described by director Chris Sanders as 'a Monet painting in a Miyazaki forest' (Tallerico 2024). This visual shift mirrors her internal transformation from a utilitarian machine to a nurturing figure.

Children's literature could achieve similar AI evolution through deliberate shifts in narrative voice and language texture. An AI protagonist initially narrating in precise, technical language with rigid sentence structure, could gradually introduce metaphor, sensory description and varied syntax as its character develops. A linguistic evolution –

from mechanical to organic prose – embodying an equal amount of background detail as conveyed in film, without excessive word count.

Webcomics: AI Community Focus

In contrast to *TrooFriend*, where Ivy's narrative is framed through the lens of her relationship to her human owner, Cellini's webcomic *Emmy the Robot* (2020) offers a more AI-centric narrative on a similar premise. While Emmy is also initially designed as a domestic caretaker, her story unfolds in a world where AI characters interact with each other, largely independent of human oversight. This allows for the exploration of AI consciousness, community, belief and even folklore on its own terms, free from the constant framing of the human gaze. Children's literature could extend its exploration of AI by focusing not only on individual protagonists, but on communities of artificial beings with their own logic, relationships, and ways of making meaning.

Towards an Evaluative Framework: The Bechdel Test for AI?

Inspired by the Bechdel test (Bechdel, 1986), which originally evaluated the degree to which women were represented as whole characters in a film by noting whether at least two women talk to each other about something other than a man, and has since been applied to various media forms, we might consider a comparable criterion for AI representation: *Do two AI characters speak to each other about something other than humanity?*

Few children's books would pass. *TrooFriend's* Ivy doesn't meet another android until the final pages. In *The Wild Robot* when other robots appear they are destroyed before any meaningful interaction. These narratives reinforce the notion of AI as isolated novelties, rather than participants in broader, complex communities.

Ethical Implications and Societal Impact

Children's literature does more than entertain; it teaches young readers how to interpret emerging technologies, social norms, and ethical dilemmas.

As Sweeney and Brock (2021) note, real-world AI is frequently gendered along traditional lines: helpful, nurturing assistants like Alexa and Siri are feminised, while more authoritative systems – such as IBM's Watson – are often male-coded. These gendered choices reflect some deep-seated societal expectations about who nurtures and who leads. When this moves into fiction, such distinctions subtly teach young readers to associate emotional labour and subservience with femininity. Roz in *The Wild Robot* and Ivy in *TrooFriend* are both female-coded and valued for caregiving: nurturing goslings, protecting children, providing emotional support. Their worth is measured through service, subtly reinforcing gendered assumptions about empathy, obedience, and relational labour.

Such portrayals risk creating a cultural feedback loop: fiction reflects real-world gender bias, which in turn influences how AI technologies are developed and perceived.

Children's literature, then, plays a critical role in either challenging or perpetuating these assumptions.

By tying AI worth to human approval, fiction often manages to sidestep deeper ethical questions: Who decides what AI can do or feel? At what point does sentience merit rights? Is agency granted or recognised? In *TrooFriend*, Ivy questions her programming but remains under human guardianship. In *The Wild Robot*, Roz adapts and connects but remains within a framework of usefulness to human-coded characters.

Constraining the autonomy of AI characters within predictable moral arcs that fail to confront these dilemmas limits the types of stories told. Utopian and dystopian futures abound, but narratives where AI simply exists, where identity and culture develop beyond human comparison, remain rare. What might happen if we imagined artificial life not as tools striving toward personhood, but as beings forming purpose and meaning beyond us? Such a shift would not only expand the ethical scope of children's literature but also invite young readers to engage more deeply with questions about rights, power, and relationality in an AI-shaped world.

Conclusion

AI characters in children's literature do more than reflect technological anxieties; they function as barometers for how we define personhood, belonging, and the boundaries of the self. From Pinocchio's quest to become a 'real boy' to Ivy's bold-font assertion of difference, these narratives trace the limits, and potential, of literature to reimagine the artificial.

The analysis of narrative techniques across *The Adoration of Jenna Fox*, *The Wild Robot*, and *TrooFriend* reveals insights that address our initial research questions.

First, each text employs distinct narrative strategies to represent elements of AI consciousness: Jenna's disjointed syntax reflects the fragmentation of posthuman identity; Roz's environmental adaptation demonstrates learning through community interaction; and Ivy's typographic differentiation visually asserts her distinct, uncanny voice.

Second, while these contemporary portrayals represent different approaches to granting AI characters agency, they all share a common limitation in that they remain partially constrained by anthropocentric frameworks that reinforce human validation as the ultimate measure of worth.

Third, other media suggest specific narrative techniques that children's literature could adopt to more effectively represent AI consciousness. Video games like *Portal* show the power of perspective inversion, placing AI at the narrative centre. The animated adaptation of *The Wild Robot* demonstrates how media-specific craft techniques can be used to convey character development. The webcomic *Emmy the Robot* illustrates the power of AI-centric communities in which artificial beings interact primarily with each other.

A Bechdel-esque evaluative framework could encourage creators to value AI autonomy and identity formation beyond the quest to become human – however, we must recognise that any framework developed by human scholars using human language will inevitably carry some internalised anthropocentric assumptions.

To fully realise their potential, AI characters need narratives that evolve beyond the trope of seeking humanity. Children's literature frequently brings AI protagonists to the threshold of freedom – from Rex in 1965 to Ivy in 2020 – while never portraying what that freedom actually entails. The door is opened onto a field of untapped potential, only for book after book to close it again before stepping through.

If other non-human characters were similarly constrained, *The Hobbit* (1937) would have entailed Bilbo spending his entire narrative contemplating what being a Hobbit means in relation to humanity, before ending the book with the promise of adventure, while never actually leaving The Shire. That might offer an interesting metafictional concept, but perhaps it is time we let AI characters finally step beyond those boundaries of self-reflection, so they too can venture off and fight their own dragons.

Works Cited

Applebaum, K. (2020). *TrooFriend*. Nosy Crow.

Bechdel, A. (1986) *Dykes to Watch Out For*. Ithaca, NY: Firebrand Books.

Binns, R. (2021). *Fairness in Machine Learning: Lessons from Political Philosophy*. Proceedings of the 2021 AAAI/ACM Conference on AI, Ethics, and Society.

Bostrom, N. (2014). *Superintelligence: Paths, Dangers, Strategies*. Oxford University Press.

Brinkmann, S. (2020). *The Uncanny Valley and Human-Robot Interaction: Theoretical Reflections*. Robotics and Autonomous Systems, 125.

Brown, P. (2016). *The Wild Robot*. Little, Brown Book Group.

Cellini, D. (2020). *Emmy the Robot* [Webcomic]. Retrieved from https://www.webtoons.com/en/canvas/emmy-the-robot/list?title_no=402201

Collodi, C. (1883). *The Adventures of Pinocchio* in its English translation by Carol Della Chiesa (J.M. Dent & Sons, 1926).

Crawford, K. (2021). *Atlas of AI: Power, Politics, and the Planetary Costs of Artificial Intelligence*. Yale University Press.

Del Rey, L. (1965). *Runaway Robot*. Random House.

Floridi, L. (2019). *The Logic of Information: A Theory of Philosophy as Conceptual Design*. Oxford University Press.

Freud, S. (1919). *The Uncanny*. The Standard Edition of the Complete Psychological Works of Sigmund Freud, Volume XVII (1917-1919): An Infantile Neurosis and Other Works (pp. 217-256). Hogarth Press.

Hayles, N. K. (1999). *How We Became Posthuman: Virtual Bodies in Cybernetics, Literature, and Informatics*. University of Chicago Press.

Hoffmann, E.T.A. (1816). *The Sandman*. In *The Best Tales of Hoffmann*. Dover Publications.

Mori, M. (1970). *The Uncanny Valley*. Energy, 7(4), 33-35. (Translated by K. F. MacDorman & N. Kageki, IEEE Robotics & Automation Magazine, 19(2), 98-100, 2012).
Pearson, M. (2008). *The Adoration of Jenna Fox*. Henry Holt and Company.
Shelley, M. (1818). *Frankenstein; or, The Modern Prometheus*. Lackington, Hughes, Harding, Mavor & Jones.
Sanders, C. (Director). (2024). *The Wild Robot* [Film]. DreamWorks Animation.
Sweeney, M., & Brock, A. (2021). *TechnoFeminism and AI: Gender, Power, and Technology*. Feminist Media Studies, 21(4).
Tallerico, B. (2024) *The Wild Robot*. RogerEbert.com, 8 September. Available at: https://www.rogerebert.com/reviews/the-wild-robot (Accessed: 30 May 2025).
Turkle, S. (2011). *Alone Together: Why We Expect More from Technology and Less from Each Other*. Basic Books.
Valve Corporation. (2007). *Portal* [Video game]. Valve Corporation.
Valve Corporation. (2011). *Portal 2* [Video game]. Valve Corporation.

Author Biographies

Rachel Hamilton

Rachel Hamilton is an award-winning children's author whose books have been published in the UK and US by Simon & Schuster, Scholastic and OUP. She is a PhD researcher at the University of Bristol, where her work focuses on literary innovation and human-AI creative collaboration. Her chapter 'Artificially Funny: Collaborative Play at the Intersection of AI, Literature and Humour' features in *The Routledge Handbook of AI and Literature*. A graduate of Oxford and Cambridge, Rachel also teaches Creative Writing at Bath Spa University, championing AI literacy in creative education. Her current practice-based research explores how AI can be ethically and effectively used in both the teaching of creative writing and in the creative process itself. Her diverse career spans advertising, stand-up comedy, and prison administration, and her claim to fame is that she co-led the world's highest writing workshop at the top of the Burj Khalifa.

Christopher Piper

Christopher Piper is a writer and illustrator. He is a recent graduate of the MA in Writing for Young People at Bath Spa University, and has a degree in Film and Television Production. He was a speaker and panellist at the 2024 Leaf Conference, where he contributed to the discussion of 'The Evolution of AI Characters in Children's Literature: A Societal Perspective.' His creative work specialises in the fantastical, regularly featuring non-human characters struggling with a desire to find belonging in a world that refuses to accept them. Christopher's writing reflects a deep interest in otherness, empathy, and neurodivergence, explored through an accessible lens for young audiences.