

AI Narratives: A History of Imaginative Thinking about Intelligent Machines

Review Number: 2433

Publish date: Friday, 11 December, 2020

Editor: Stephen Cave
Kanta Dihal
Sarah Dillon

ISBN: 9780198846666

Date of Publication: 2020

Price: £55.00

Pages: 448pp.

Publisher: Oxford University Press

Publisher url: <https://global.oup.com/academic/product/ai-narratives-9780198846666?cc=us&lang=en&>

Place of Publication: Oxford

Reviewer: Leah Henrickson

‘Artificial intelligence (AI)’ is a loaded term, rife with connotative contradiction that inspires debate, disagreement, and disillusion. But what *is* AI, really? How have our expectations of computational capability, and even a robot Armageddon, come to be? Why does it matter how we talk about increasingly sophisticated technology, not just in expository prose, but also in fiction? How might visual and cinematic representations of AI technologies reflect and contribute to particular understandings and potentialities? *AI Narratives: A History of Imaginative Thinking about Intelligent Machines* is an ambitious effort to prompt widespread critical consideration not of AI itself, but of the stories—particularly fictional stories—surrounding this pervasive but still largely misunderstood technology. Comprising 16 chapters by scholars from various disciplines, *AI Narratives* offers a panoramic snapshot of a land still largely unexplored: AI humanities.

Given the relative newness of AI as a term, though, the editors of *AI Narratives* have opted for the more inclusive ‘intelligent machines’. Recognising the range of terms used to refer to such machines (automaton, android, robot, and cyborg), they explain that ‘[t]he chapters in this book engage with AI in its broadest sense, one that encompasses all of the aforementioned terms: that is, any machine that is imagined as intelligent’ (p. 4). The editors further specify an Anglophone Western scope, and divide the collection into two parts: ‘Antiquity to Modernity’, with each chapter focusing on an era; and ‘Modern and Contemporary’, with each chapter focusing on a theme. Both parts question, commend, and criticise portrayals and explanations of AI. ‘The work of this book contributes to establishing the importance of narratives as constituent parts of any sociotechnical imaginary,’ the editors write. ‘Narratives of intelligent machines matter because they form the backdrop against which AI systems are being developed, and against which these developments are interpreted and assessed’ (p. 7). Despite widespread societal concerns about AI technologies, the editors are optimistic that, through conscious consideration and adjustment of AI narratives, we will progress towards affirmative integration of AI technologies into everyday life.

It is this optimism that drives the University of Cambridge’s Leverhulme Centre for the Future of

Intelligence, with which two of the editors are affiliated. Indeed, this collection has emerged from the AI Narratives project workshops jointly led by the Centre and the UK's Royal Society. These workshops also resulted in a pamphlet about *Portrayals and perceptions of AI and why they matter*, published by The Royal Society in 2018.⁽¹⁾ Those seeking further analysis of representations of AI may benefit from reading other publications stemming from the Centre.

AI Narratives begins with antiquity, as Genevieve Liveley and Sam Thomas examine the intelligent machines in the Homeric epics (Chapter 1). Many of these examples are weak, both in the sense that they could be categorised as weak AI and that their intelligence is largely implied. However, Liveley and Thomas' analysis of the Phaeacians' self-sailing ships (*Odyssey*) in particular illustrates the hierarchies of cognitive ability that continue to distinguish human and machine. The ships, after all, are 'programmed by the knowledge and intelligence found in human minds (*noos/noemata*). [...] The ships' "independent" thinking and "independent" movement is limited by their human programmers and users' (p. 35). The use of 'programmed' here is admittedly anachronistic, but such language reinforces the relevance of Homeric epics in framing AIs as tools under human control. Even in Homer's strongest 'embodied AI'—identified as the *amphipoloi* of the *Iliad*—the machines are portrayed as servants of sorts, existing solely to assist their masters. 'Already in antiquity, then, automata and intelligent machines were conceived as artificial slaves, occupying a special social status no less than ontological category: above that of an inanimate tool—an oar, say; but below that of a human,' Liveley and Thomas aver (p. 41).

E.R. Truitt elaborates upon the slave argument in her chapter about AI in works of the Latin Middle Ages (Chapter 2), focusing on the representation of early AIs as means for maintaining existing social hierarchies. 'The desire to use artificial intelligence to spy, to police boundaries and populations, and to act as slaves in the service of power appears equally in the medieval period and in the now,' Truitt observes. '[S]o what does that tell us about now?' (p. 66) The understanding of AIs as slaves – or, when considered more positively, powerful or even superhuman servants – pervades the history of AI presented in this collection. In her chapter (Chapter 8), Kanta Dihal makes direct connections between narratives of AI and narratives of slavery, delving into the grey area between personhood and thinghood to argue that AI may justifiably resist enslavement. Ultimately, through Dihal's chapter, 'we see that the AI uprising narrative is not about the technology itself, and certainly not about the current state of the technology. Instead, these narratives reflect much older accounts of intelligent people being used as tools' (p. 208). Other chapters elaborate upon the varying potential levels of control that AIs may exert over their human makers (Chapter 9), and even ponder whether distinctions between humans and machines actually matter (Chapter 10). Kate Devlin and Oliva Belton assert in their chapter (Chapter 15), though, that sometimes these distinctions are not so clear. Through an analysis of gynoids and sex robots, Devlin and Belton show how fictional representations of 'female' robots and real-world technological development have perpetuated harmfully narrow expectations of womanhood.

Fictional representations of AI transcend genres, but in her chapter (Chapter 12) Anna McFarlane looks to cyberpunk 'no longer simply as a genre of science fiction, but as a mode that provides an avenue for representing and interrogating the role of algorithms and the power of data over our lives' (p. 285). Indeed, as McFarlane observes, '[r]ather than aiming to predict the future, cyberpunk diagnoses the state of technology in the present, arguing that we are already cyborgs, components in a system that we cannot completely understand' (p. 291). To repurpose the words of Julie Park (Chapter 5), AI has long seeped through 'the porous boundary between popular and intellectual culture' (p. 136), being the subject of overt public spectacle, as well more implicit deception. Minsoo Kang and Ben Halliburton's chapter about a device reputedly created by medieval philosopher Albertus Magnus (Chapter 3) further shows how AI may be between mechanics and magic. Those 'components in a system that we cannot completely understand' may be justified through comparison to more familiar—or familial—circumstances (Chapter 11), or they may simply be accepted as technological feats. These feats, however, do not necessarily change the proverbial game. AI is, Sarah Dillon and Michael Dillon argue (Chapter 14), simply another *player* in social games, governance, and life in general.

AI Narratives introduces readers to numerous fictionalised accounts of seemingly intelligent machines, and offers rich analyses of these accounts as they may pertain to nonfictional technological and social developments. The collection reviews the very real fictional histories of AI, drawing attention to the ways in which the expository and aesthetic may come together to instil different understandings of mechanical capacity in different publics. All of the book's chapters address the use of AI to establish and maintain power, whether human or robotic. When considered together, the chapters constitute a pointed argument for a necessarily historical and literary study of technology to meaningfully question and counteract hegemonic norms.

From Aristotle to Alexa, intelligent machines have long been the subject of human imagination, and this collection provides much-needed historical insight into representations of AI that have informed modern perceptions. Yet imagined applications and implications of intelligent machines actually appear to have changed little over time, and *AI Narratives* may have been enhanced with more inter-chapter citation to stress the transhistorical nature of themes addressed and avoid repetition of particular examples (e.g. *Blade Runner*). The collection may too have been enhanced with greater attention to nonfiction narratives. In his chapter (Chapter 13), Stephen Cave examines the narrative devices of works of speculative nonfiction, arguing that these elements have helped the concept of 'mind uploading' gain public traction and support. In contrast, Cave observes, works of fiction have served as spaces of critical imagination 'through narratives that are more nuanced and multisided than those of the proselytizers' (p. 329). While the focus of *AI Narratives* is described in the Introduction as being on the fictional, 'imaginative thinking' is by no means limited to the realms of fiction, and is it hardly contentious to recognise the bilateral conference between the real and imagined. The book's narrowed scope prevents digression, but it also inspires future research into the gaps that it adroitly leaps over.

Another such gap is the possible distinction between human and machine intelligence. Although some chapters allude to the unique potential of AI's cognitive and physical exploits, AIs are always measured against human counterparts. In his comparison of Renaissance AIs to those of our modern day (Chapter 4), for example, Kevin LaGrandeur concludes that '[u]ltimately, superhuman slaves—and our worries about them—reflect an archetypal need to supersede our natural limits by whatever ingenious means, even if by developing super-powerful artificial servants we threaten the very dominion we seek' (p. 114). Yet as Paul March-Russell concludes in his discussion about robot consciousness (Chapter 7), '[m]achine consciousness as something separate and distinct from the human gaze (an often white, male gaze) remains, like the figure of Hadaly [from Villiers de l'Isle-Adam's *Tomorrow's Eve*], mysterious and unexplored' (p. 184). Megan Ward (Chapter 6) gingerly attempts such an exploration. Reviewing arguments for and against computational creativity, she writes that '[o]n both sides, machine creativity is being defined against human creativity, whether through imitation, evaluation, or even attempts at machine independence altogether. [...] Victorian fictions of originality offer a form of creativity not directly traceable either to machine or human but to their symbiosis' (p. 157). Ward suggests that computational creativity is not—and need not be—equivalent to human creativity, but might instead be distinct, and might further contribute to new processes and outputs that could be deemed creative. Ward's chapter ends with the declaration that '[i]n order to become creative, machines need not necessarily be more like humans; rather, humans could recognize the creative capacities inherent in mechanicity' (p. 159). For this to be would require adjustments to modern conceptions of creativity, but Ward's insight into Victorian conceptions shows the malleability of the term, as well as the possibility of a sort of mechanical autonomy. However, nowhere in *AI Narratives* is intellectual autonomy of machines sufficiently addressed. In this way, while *AI Narratives* presents original research, its analyses largely restate what is already widely accepted: intelligent machines may be superhuman, slaves, or superhuman slaves, but their existence and value remains dependent upon human control. For a book about imaginative thinking, these analyses—however insightful—are perhaps not as imaginative as one would expect.

In the book's final chapter (Chapter 16), Gabriel Recchia applies digital humanities techniques to identify key characteristics of current AI narratives in screen media from the Anglophone West. Recchia emphasises

the prevalence of the word ‘control’ in references to AI, confirming that many uses correspond to fears of both humans and AIs losing control. But, Recchia notes, *how* AI agents may lose control in reality differs from lost control on screen. The former sees AI optimising its programmed instructions by, for example, lobbing a pancake to keep it in the air as long as possible, while the latter sees circuits gone haywire and robotic rebellion. ‘While the “evil overlord” theme may strike some as a legitimate real-world concern as the purposes that corporations and states put AI to come under greater scrutiny, it somehow seems a shame that real failure modes of AI systems are not better represented,’ Recchia concludes. ‘By inflating implausible fears [...] and missing genuine risks [...], many popular depictions of AI miss opportunities for creative storylines that are more in line with reality, and which will increasingly resonate with audiences as AI systems become increasingly integrated into everyday life’ (pp. 404-405). And so *AI Narratives* concludes: not with a roboapocalypse, but with a hopeful plea for recognition of the almost unbelievable current state of the art.

AI Narratives triumphantly paves the way for future work in AI humanities. Individual chapters—all balancing historical context with sharp analysis—would make valuable additions to relevant module syllabi, and the volume would be of certain advantage to any reader seeking a fresh and substantiated approach to AI scholarship. This *History of Imaginative Thinking about Intelligent Machines* is only a first glance into this kaleidoscopic field of study, but it positions future researchers well for imaginative thinking about their own perceptions.

Notes

1. Portrayals and perceptions of AI and why they matter (London, 2018). <<https://royalsociety.org/topics-policy/projects/ai-narratives>> [2][Back to \(1\)](#)

Source URL: <https://reviews.history.ac.uk/review/2433>

Links

[1] <https://reviews.history.ac.uk/item/328335>

[2] <https://royalsociety.org/topics-policy/projects/ai-narratives>>