Proposal for Edited Collection:
Minding the Future: Contemporary Issues in Artificial Intelligence

Ed. Barry Dainton, Will Slocombe, and Attila Tanyi

PROJECT OVERVIEW

This collection examines the ways in which Artificial Intelligence is understood within contemporary culture, and the extent to which it challenges the *status quo* or offers solutions to existing socio-cultural problems, specifically in relation to how these issues are treated in speculative and science fictions. *Minding the Future* explores the metaphysical, ethical (moral), cultural, and societal implications of Artificial Intelligence through philosophy, literature and connected fields. Issues to be addressed include:

- The issues of defining "Artificial Intelligence" and the ramifications of such definitions; considerations of the different ways in which AI might "think", "reason" or "be aware" and the implications of these differences.
- The different forms Artificial Intelligences might take: algorithms v. AIs; programmed v. self-taught computers; drones & robots; AIs that are wholly non-conscious and AIs that are fully conscious.
- The potential benefits and potential dangers of AIs, and the ways in which these have been represented: an end to drudgery and poverty (or the creation of mass unemployment), saving us from environmental catastrophes (or exterminating us), the creation of heavenly (or hellish) virtual worlds for us to inhabit.
- The issues of ethical, moral, and legal responsibility towards, and of, Artificial Intelligences: what sort of AI could have the moral significance of a human? What sort of AI could *be* moral??
- The extent to which Artificial Intelligence challenges human exceptionalism and/or existing moral and ethical frameworks; how will our conception of ourselves be affected if (as seems increasingly likely) non-conscious machines can possess supra-human levels of intelligence?
- The Singularity: the *very* rapid expansion of AI, soon leading to hyperintelligence, and the potential to take over the world (or universe) – how real is the threat?
- Robots: how we'll interact with intelligent machines that (a) don't outwardly resemble human beings, (b) which do outwardly resemble human beings, and which (c) may or may not be conscious.
- The metaphysical and existential challenges posed by AI-generated virtual and augmented realities.
- The tropes of "Artificial Intelligence" and the ways it is epistemologically- and socially-constructed, alongside the effects that such tropes have in influencing our perceptions and understanding of AIs.
- The mutually enriching interaction between science fiction and the human and social

sciences: the way science fiction incorporates philosophical, social and cultural ideas and at the same times influences philosophical, social and cultural thinking.

PRIMARY FIELD & AUDIENCE

Bringing together literary scholars, media commentators, cultural theorists, ethicists, philosophers of mind, and those from closely affiliated disciplines, this collection offers important and timely insights into the pasts, presents, and possible futures of Artificial Intelligence. The volume does not aim merely to gather together contributions from different fields: the ideas that emerge in relation to speculations about the future as they appear in speculative and science fiction literature will serve as "springboards" for contributors to bring disciplinary knowledges and paradigms to bear on those issues, to elucidate both the issues and the variety of approaches that current research can bring to bear upon them. Moreover, contributions will be required to connect to other fields via these issues, and contributors will be asked to comment and reflect on each other's work during the writing and editing process.

In short, the volume's guiding methodological "mission" is to ask contributors to reflect on, discuss, analyse and engage with (science and speculative) fictional work(s), topics and issues that figure Artificial Intelligence prominently. In this way, we will produce a volume that has a clear literary focus while bringing in all sorts of disciplines with their own ways of looking at problems and analysing them. Because of the nature of contributions, a genuinely interdisciplinary dialogue on the issues that Artificial Intelligence raises will emerge that will be of interest to an array of fields and disciplines.

The volume will produce tangible knowledge for its readership, and we envisage that the following are the most important ways readers will benefit from the volume:

- Readers will gain an in-depth understanding of the wide range of pressing issues to which Artificial Intelligence gives rise, and the various ways in which science fiction narratives have been used to represent them.
- Readers will be able to relate theoretical positions on Artificial Intelligence to discussions in fictional works (movies as well as novels and stories).
- Readers will be able to reflect on the diverse aspects of Artificial Intelligence—its philosophical, social, legal, economic, scientific and cultural ramifications—through the various disciplinary perspectives of the volume.

PROPOSED STRUCTURE

The complete collection will come to approx. 100,000 words, in total (including index, table of contents, notes, etc.), and comprise of twelve essays and an introduction by the editors. To ensure that the proposed collection has both the requisite depth as well as breadth, the twelve essays will be written by a diverse range of authors who come from different disciplines and apply different methods of investigation to their subject matter. However, in line with our methodological remarks above, the focus will throughout be on fictional representations of

Artificial Intelligence and the issues it raises. Although an exhaustive list of this topic is impossible, given the proliferation of AI even in English-language media, indicative texts that contributors might utilise include fictional texts—by authors such as Peter Watt, Isaac Asimov, Daniel H. Wilson, Greg Egan, William Gibson, Ann Leckie, Iain M. Banks, Stanislas Lem, Robert Charles Wilson—alongside films and franchises—for instance, *The Matrix*, *Tron*, *Total Recall*, *Terminator*, *Transcendence*, *Wall-E*, *Westworld*, *Battlestar Galactica*, *Blade Runner*. If contributors wish, we will also extend the remit of the collection to computer games such as *Deus Ex*, *Fallout*, and *Mass Effect* series.

These four sections (of three chapters each), which each inflect the titular topic of "minding the future", are as follows:

1. MINDS OF THE FUTURE: *Agency, Consciousness, and Identity*. This section examines the problem of consciousness in relation to Artificial Intelligence and autonomous systems, and issues that emerge from the various ways in which they have been defined, represented, and understood. It will consider areas such as uploading and digital consciousness, virtual and simulated worlds, philosophy of mind, philosophy of intelligence, robotics.
2. FUTURE MINDS IN (INTER)ACTION: *Human-Computer Interactions*. This section involves an exploration of the ways in which artificial systems and automata (might) interact with humanity on both an individual level, such as between person and robot, as well as the broader social dimensions within which these interactions will operate (such as the question of the Singularity). Using the various disciplinary frameworks that contributors will bring to the topic in part defines the areas that will be analysed, but care will be taken to ensure that interpersonal and societal areas are covered.
3. ETHICAL MINDS? *Ethics, Politics, and Law in the Future*. This section interrogates current issues in ethics, politics and law in relation to Artificial Intelligence, and considers the ways in which understandings of these systems are impacted by the creation of Artificial Intelligence and autonomous systems. This involves both the ways in which current ethical, legal, political approaches and systems might be applied to Artificial Intelligences as well as the ways in which the relevant frameworks will need to be re-conceptualised or re-framed to respond to the emergence of Artificial Intelligence.
4. BEING MINDFUL OF THE FUTURE: *Speculations*. This section explores the ways in which speculations about the future, such as science fiction texts, are models based upon a particular "mindset" or mental image about the future. The section thus considers the very act of being "future-minded", albeit particularly in relation to Artificial Intelligence, and considers the perspectives and politics through which Artificial Intelligences have been represented and why they hold such fascination for us today.

EDITORS

Prof Barry Dainton is Professor of Philosophy at the University of Liverpool, UK. He specializes in philosophy of metaphysics and philosophy of mind. His previous publications include *Stream of Consciousness* (2nd edition, 2006), *The Phenomenal*

Self (2008), *Time and Space* (2nd edition, 2010) and *Self* (Penguin, 2014). He is also the co-editor (with Howard Robinson) of *The Bloomsbury Companion to Analytic Philosophy* (Bloomsbury 2014).

Dr Will Slocombe is a lecturer in the Department of English at the University of Liverpool. Aside from publishing in the field of sf writing, specifically on depictions of Artificial Intelligence, and editing special issues of journals, including a *Genre* issue on "Narratives and New Media," he has published numerous articles on contemporary literature and literary theory. He is currently working on his second monograph, *Emergent Patterns: Artificial Intelligence and the Structural Imagination* (intended for Peter Lang, 2019). He is a member of the Science Fiction Foundation Committee and the theme lead for "Designing Futures" for the University of Liverpool's Centre for Humanities and Social Sciences of Health, Medicine and Technology.

Dr Attila Tanyi is currently Associate Professor in the Institute of Philosophy and First Semester Studies at the University of Tromsø: The Arctic University of Norway. He specializes in moral and political philosophy but his work stretches over disciplinary boundaries, and he regularly collaborates with philosophers whose specializations are very different from his own, as well as with non-philosophers with an interest in philosophical problems. He has particular interest in experimental philosophy, practical rationality, moral demands, institutional consequentialism as well as immortality, and recently led the German Science Foundation (DFG) funded research group *Consequentialism and Its Demands.* His works have appeared in several journals (*Philosophical Studies, Utilitas, Theoria, Ethical Theory and Moral Practice*, *Religious Studies, Criminal Law and Philosophy, Philosophia,* among others) and is presently working on two manuscripts: one on reasons and another on moral demands.

INTENDED READERSHIP

This is intended to be a *scholarly* volume primarily: for academic researchers of the fields involved to use in their own research. However, we are confident that that the volume will be useful for higher-level – advanced undergraduate and postgraduate level – *teaching* purposes. This will be in the form of secondary or recommended reading on the increasing number of courses that intersect with the topic areas of the collection. Since the volume will be strongly interdisciplinary, it can be used for teaching and research in several, sometimes rather different fields (such as literature, media studies and philosophy).

THE MARKET – WHAT GAP TO FILL?

Since this will be a scholarly volume, its primary competitors will be similar research monographs and edited volumes, and not the burgeoning popular science literature on AI and related topics (singularity, robotics, mind uploading etc.). Within the scholarly literature, there are no clear direct competitors. Some volumes are too old to be up-to-date (e.g. Bolden: *The Philosophy of Artificial Intelligence*); too broad in scope (e.g. the 2014 and 2016 edited volumes by Vincent C. Müller, with Springer, or More and Vita-More edited volume *The Transhumanist Reader*, with Wiley); and none of them figure literature and science fiction in them (with the exception of the Wiley produced volume, edited by Russel and Broderick, *Intelligence Unbound,* but this covers not so well-known sci-fi writers and it focuses mostly

on mind-uploading and little on AI).

MIT Press appears to be particularly active in this area but their volumes also bear one or more of the above marks or have other shortcomings (from our point of view). Three appear to be particularly relevant:

- Franchi, Stefano and Güven Güzeldere, eds. *Mechanical Bodies, Computational Minds: Artificial Intelligence from Automata to Cyborgs*
- Shanahan, M., *The Technological Singularity*
- Lin *et. al.* (ed.), *Robot Ethics*

These volumes either focus only on technical/technological aspects of AI (Shanahan), only certain aspects (epistemology, philosophy of mind) of the philosophy of AI (Franchi & Güzeldere), or do not discuss AI in much detail at all (Lin). None of these volumes, moreover, bring in fields other than philosophy and science: no literature (let alone science fiction) figures in them, to mention one.

So far as Cambridge University Press is concerned, the following volumes have been published recently (according to the online catalogue):

- Frankish, K. and Ramsey, W.M eds., *The Cambridge Handbook of Artificial Intelligence*
- Hernández-Orallo, J., *The Measure of All Minds: Evaluating Natural and Artificial Intelligence* (forthcoming)
- Poole, D.L. and Mackworth, A.K. eds., *Artificial Intelligence: Foundations of Computational Agents*

However, as with the MIT press volumes, the last two books focus entirely on technical/technological aspects of AI (the Poole/Mackworth volume is a textbook, moreover and the Hernández-Orallo forthcoming monograph is only interested in the specific problem of measurement). The Frankish/Ramsey volume, on the other hand, is too general within the research on AI, even if only philosophy is concerned.

Finally, although there are several science fiction and philosophy books on the market, many of these are too old (e.g. Michael Phillip's *Philosophy and Science Fiction*, from 1984). Of the newer ones Nichols, Smith and Miller (eds.) *Philosophy through Science Fiction* (2008) only devotes 4 pages out of 400 to AI. Although Susan Schneider's *Science Fiction and Philosophy* (2009) does include a substantial 100-page section entitled "Mind: Natural, Artificial, Hybrid and Superintelligent", it largely consists of standard philosophy essays by Block, Clark and Dennett, and the relevant connections with science fiction are not analyzed or discussed in any detail – the 2016 second edition is essentially the same in these respects. Moreover, none of these books focus on AI and they are primarily intended to be textbooks for undergraduate level teaching.

Hence it seems to us that our proposed volume would clearly fill a niche in the market.