

BIBLIOGRAPHY, REFERENCES & SOURCES

BOOKS

SMARTER THAN US: THE RISE OF MACHINE INTELLIGENCE

Stuart Armstrong (Research Fellow, Future of Humanity Institute, University of Oxford)
Machine Intelligence Research Institute, Berkeley, 2014

What happens when machines become smarter than humans? When machines become smarter than humans, we'll be handing them the steering wheel. Can we instruct artificial intelligence (AI) to steer the future as we desire? What goals should we program into them? It turns out this question is difficult to answer! Philosophers have tried for thousands of years to define an ideal world, but there remains no consensus. The prospect of goal-driven, smarter-than-human AI gives moral philosophy a new urgency. The future could be filled with joy, art, compassion, and beings living worthwhile and wonderful lives—but only if we're able to precisely define what a "good" world is, and skilled enough to describe it perfectly to a computer program. AI, like computers, will do what we say—which is not necessarily what we mean. Such precision requires encoding the entire system of human values for an AI: explaining them to a mind that is alien to us, defining every ambiguous term, clarifying every edge case. Though an understanding of the problem is only beginning to spread, researchers from fields ranging from philosophy to computer science to economics are working together to conceive and test solutions. Are we up to the challenge? Armstrong wrote SMARTER THAN US at the request of the Machine Intelligence Research Institute, a non-profit organization studying the theoretical underpinnings of artificial superintelligence.

THE TECHNOLOGICAL SINGULARITY

Murray Shanahan (Professor of Cognitive Robotics, Imperial College London)
MIT, Massachusetts, 2015

The idea that humans will someday be overtaken by artificially intelligent machines or cognitively enhanced biological intelligence, or both has moved from the realm of science fiction to serious debate. Some theorists predict that if technology continues to develop at its current rate, we could be overtaken before the middle of this century. Whether we believe this or not, it's an idea that raises crucial philosophical and pragmatic questions, forcing us to think seriously about what we want as a species. What will the existence of superintelligent machines mean for such matters as personhood, responsibility, rights, and identity? We need to imagine the possibilities.

OUR FINAL INVENTION: ARTIFICIAL INTELLIGENCE AND THE END OF THE HUMAN ERA

James Barrat (documentary filmmaker)

St. Martin's Press, New York, 2013

In as little as a decade, artificial intelligence (AI) could match and then surpass human intelligence. Corporations and government agencies around the world are pouring billions into achieving AI. Once AI has arrived, scientists argue, it will have survival drives much like our own. We may be forced to compete with a rival more cunning, more powerful, and more alien than we can imagine. Until now, human intelligence has had no rival. Can we coexist with beings whose intelligence dwarfs our own? And will they allow us to?

THE SINGULARITY IS NEAR

Ray Kurzweil (a senior consultant in computer science at Google)

Penguin - Viking, London, 2005

An examination of the next step in the evolutionary process and a vision of our ultimate destiny: the culmination of centuries of technological ingenuity; the union of human and machine, when we shall transcend our biological limitations and combine with the vastly greater capacity, speed and knowledge-sharing abilities of our creations. In practical terms, human ageing and illness will be reversed; pollution will be stopped and world hunger and poverty will be solved. There will be no clear distinction between human and machine, real reality and virtual reality.

SUPERINTELLIGENCE: PATHS, DANGERS, STRATEGIES

Nick Bostrom (Professor, University of Oxford; Director, Strategic Artificial Intelligence Research Centre; Founding Director, Future of Humanity Institute)

Oxford University Press, 2014

If machines one day come to surpass humans in general intelligence - just as the fate of the gorillas now depends more on humans than on the gorillas themselves - the fate of our species would come to depend on the actions of the machine. But we have one advantage: we get to make the first move. Will it be possible to engineer initial conditions so as to make an intelligence explosion survivable? Thinking about the human condition and the future of intelligent life, this is the essential task of our time.

HOMO DEUS: A BRIEF HISTORY OF TOMORROW

Yuval Noah Harari (Lecturer in World History, Hebrew University of Jerusalem)

Harvill Secker, London, 2016

Did science fiction predict the current state of the world? Recent headlines about smartphone addiction among children and tech companies failing to prevent the proliferation of porn and terrorist material online can often make it seem like we are living in a science fiction nightmare. Harari envisions a not-too-distant world in which we face a new set of challenges, exploring the projects, dreams and nightmares that will shape the Twenty-first Century (from overcoming death to creating artificial life) and asking the fundamental questions: Where do we go from here? And how will we protect this fragile world from our own destructive powers?

This was featured and discussed on *The Today Programme*: BBC Radio 4, 2017_04_03 (08.42).

TO BE A MACHINE: ADVENTURES AMONG CYBORGS, UTOPIANS, HACKERS, AND THE FUTURISTS SOLVING THE MODEST PROBLEM OF DEATH

Mark O'Connell (journalist and literary critic)

Granta, London, 2017

Transhumanism is a movement aiming to use technology to fundamentally change the human condition, to improve our bodies and minds to the point where we become something other, and better, than the animals we are. It's a philosophy that can seem hopeful, terrifying or absurd. In this first full-length exploration of transhumanism: its philosophical and scientific roots; its key players and possible futures, Mark O'Connell presents us with an urgent question: What next for mankind?

This was featured and discussed both on *Newsnight* : BBC Television, 2017_03_31; and on *Start the Week*: BBC Radio 4, 2017_04_03.

TELEVISION

WHO WILL WE BE

Episode of 'The Brain' (science documentary)

BBC Television, 2016_02_26

In this concluding episode of his six- part documentary exploring all aspects of the human brain, neuroscientist Dr David Eagleman journeys into the future, and asks what's next for the brain and for our species. And - after seeing new sensory inputs plugged into a brain and the electrical signals from someone's motor cortex transmitted across a room into an artificial limb - he asks whether or not one's extended and enhanced mind could ever exist entirely within a machine.

THE IMMORTALIST

Episode of 'Horizon' (series of science documentaries)

BBC Television, 2016_03_16

Some of the world's leading neuroscientists, robot builders and consciousness researchers explore systems that might - within the next thirty years - allow a human being to escape its biological destiny and become immortal. In conclusion, this programme suggests, "We are now embarking on a journey into a very different world. Whether we find we can live forever in machines or not, for some the journey will certainly change what it is to be human."

CYBER ATTACK - THE DAY THE NHS STOPPED

Episode of 'Horizon' (series of science documentaries)

BBC Television, 2017_06_12

This programme presents a recent example of how a machine can, within minutes, shut down entire sections of a country's vital infrastructure.

RADIO

BLACKOUT

Julian Simpson

BBC Radio 4, 2016_05_22

This specially-commissioned *fifteen minute radio drama (in the BBC's Dangerous Visions series) begins with, "Yesterday was the day when the phones stopped working". But the phones are only the start, as London is plunged into an internet and power blackout. A woman starts a diary of the first days as she watches first the infrastructure and then the civilisation of her city begin to collapse.*

DR DEMIS HASSABIS

Desert Island Discs

BBC Radio 4, 2017_05_21

Kirsty Young interviewed *Dr Hassabis, an artificial intelligence researcher, neuroscientist and co-founder of DeepMind (an AI company that, in 2014, was acquired by Google). At university Hassabis deliberately chose to study topics where AI had so far failed: memory and imagination. More recently he helped to design a program that taught itself how to play and - in 2016 - beat the world champion Go player. Hassabis says AI programs are currently using neural networks to develop their abilities for learning and making decisions. He believes they'll be benevolent towards us in their intent. When asked about the consequences of these programs being developed by "people who have not our best but our worst interests at heart" he replied: "What we hope is that, once we understand these systems better and understand what we would be legislating for, there would be some stronger form of governance that would be agreed by world governments."*

ONLINE

ARTIFICIAL INTELLIGENCE

Online reference

Wikipedia, August 2017

Artificial intelligence (AI) *is intelligence exhibited by machines.*

- ⇒ https://en.wikipedia.org/wiki/Artificial_intelligence
- ⇒ https://en.wikipedia.org/w/index.php?title=Artificial_intelligence&action=history

ROBOTICS AND ARTIFICIAL INTELLIGENCE INQUIRY

Science and Technology (House of Commons) Committee

Parliament, London, 2016

UK Parliamentary Committee *recommends a standing commission on AI.*

- ⇒ [science-and-technology-committee - robotics-and-artificial-intelligence-inquiry](#)

REPORTS OFTEN REFERENCE AT LEAST ONE OF THE FOLLOWING SOURCES

The University of Oxford, Future of Humanity Institute

⇒ fhi.ox.ac.uk

The University of Cambridge, Centre for the Study of Existential Risk

⇒ cser.org

The Machine Intelligence Research Institute in Berkeley, California

⇒ intelligence.org

The Future of Life Institute in Boston, Massachusetts (*which, in 2015, published 'Research Priorities for Robust and Beneficial Artificial Intelligence: An Open Letter'*).

⇒ futureoflife.org/misc/open_letter

RETHINK WHAT'S POSSIBLE

Marketing material

Samsung Electronics, April 2017

This marketing material *features Samsung's new Galaxy S8, Gear 360 and VR mobile devices.*

⇒ <http://www.samsung.com/uk/discover/news/do-what-you-cant-with-galaxy-s8-and-vr>

COGNITIONX

Market Intelligence Platform

London, 2017

Aiming to organise *market information and research on products and resources for building AI solutions*, CognitionX CEO Charlie Muirhead (*in his introduction to its most recent conference*) said, "*The [AI] accelerator pedal is foot to the floor, but it's not clear who's got their hands on the wheel or in fact where we're all going.*"

⇒ <https://cognitionx.com>

⇒ <https://twitter.com/hashtag/artificialintelligence>

HUMANITY

Humans are uniquely adept at utilizing systems of symbolic communication (such as language and art) for self-expression and the exchange of ideas, and for organizing themselves into purposeful groups. Humans create complex social structures composed of many cooperating and competing groups, from families and kinship networks to political states. Social interactions between humans have established an extremely wide variety of values, social norms, and rituals, which together form the basis of human society. Curiosity and the human desire to understand and influence the environment and to explain and manipulate phenomena (or events) has provided the foundation for developing science, philosophy, mythology, religion, anthropology, and numerous other fields of knowledge.

⇒ <https://en.wikipedia.org/wiki/Human>

Human nature refers to the distinguishing characteristics—including ways of thinking, feeling, and acting—which humans tend to have naturally.

⇒ https://en.wikipedia.org/wiki/Human_nature

WHAT SETS US APART FROM MACHINES

The United Nations' General Assembly *designates* August 19th *each year as* World Humanitarian Day *to recognize those who face danger and adversity in order to help others.*

⇒ un.org/en/events/humanitarianday

CONSTRUCTIVE INTER-FAITH DIALOGUE IS DEVELOPING THROUGH THESE FORUMS

The Inter Faith Network for the UK

⇒ interfaith.org.uk

Wikipedia's article on interfaith dialogue

⇒ wikipedia.org/wiki/Interfaith_dialogue

Brighton & Hove Inter-Faith Contact Group

⇒ interfaithcontactgroup.com

The Council for a Parliament of the World's Religions

⇒ parliamentofreligions.org

<https://petersharrock.wordpress.com/>