

SUPERINTELLIGENCE by NICK BOSTROM

So the author is a philosopher from University of Oxford. He also created infamous paperclip analogy. According to this analogy, we can tell an AI to create paperclips. If we don't have control mechanism in place, this AI can first use entire earth, then solar system and then entire galaxy to create paperclips. If you are into gaming, there is a game called mass effect which revolves around this topic as well.

By its nature, this is quite different from most other books due its philosophical nature. The topics of discussion are not particularly mainstream. The books begins with a little introduction to the field (it was published in 2014), and then describe possible paths to superintelligence. He discusses forms of superintelligence and dynamics of it.

After this point, the philosophical part starts dominating the books, in my opinion. He discusses possible AI takeover scenarios, the will of the superintelligence and also outcome of the AI. He proposes certain control mechanisms to the AI problem. Lastly, he focuses on the ethics and values of the AI.

Although I already read the book once, I'll definitely read it again. In my opinion, this is not just a book you read and set aside. It is beneficial to do some brainstorming as well. And I found some discussion about ethics quite interesting. For example, assuming we created workers with AI to do our bidding, do we enslave them, sentient creatures.

Despite its different tone than most of AI books, it is an interesting read.