

# Superintelligence Summary

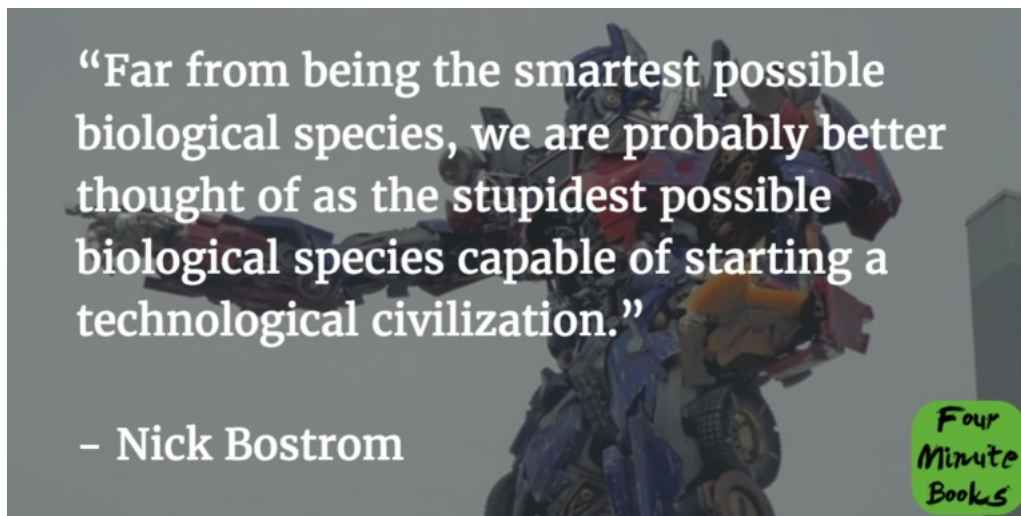
---

 [fourminutebooks.com/superintelligence-summary](https://fourminutebooks.com/superintelligence-summary)

**1-Sentence-Summary:** *Superintelligence asks what will happen once we manage to build computers that are smarter than us, including what we need to do, how it's going to work, and why it has to be done the exact right way to make sure the human race doesn't go extinct.*

**Read in:** 4 minutes

**Favorite quote from the author:**



The Matrix, Terminator, Iron Man – the list of movies in which some form of artificial intelligence suddenly goes crazy and tries to take over the world is long. The idea has been around since the 50s, but it seems we're getting closer to a world where computers are as smart as humans or even smarter.

The question is: what will that look like? Will the machines really be able to rebel against us? Will we put AI into humanoid robots? Will the internet start to think?

Nick Bostrom calls this phenomenon superintelligence. And while some books try to predict what will happen after we've created it, *Superintelligence* is pre-occupied with determining what the path towards it must be like in order for it to work out in our favor.

Here are 3 lessons about the state of artificial intelligence to show you it's up to all of us to make our future a good one:

1. AI used to be limited by hardware, but now it's mostly a problem of data.
2. There are two different ways to design superintelligent computers.
3. Superintelligence must be the result of global collaboration, not some secret government program, or we're screwed.

Forget the dusty crystal ball in your attic, this book will give you a much clearer glimpse into the future!

## Lesson 1: Initially AI was limited by hardware, now it's mostly a matter of feeding computers enough data.

---

After Alan Turing's "Turing Machine" was invented as the first device to systematically follow and execute instructions in an automated way (watch *The Imitation Game* for more details, superb movie), the first real "digital" computer was completed in 1946.

Ever since that moment, computer scientists have been wondering how we can get these machines to actually think like us. The Dartmouth Summer Research Project on Artificial Intelligence was one of the first proper workshops in this area in 1956, and even though the next few years showed some results, like machines solving math problems or writing music, AI soon hit its limit – **the hardware simply didn't suffice to process all the necessary information for really complex tasks.**

It took until the 80s for hardware to slowly catch up, but then the development of expert systems gave rise to the first, proper AI, which, for example, could diagnose cars like a mechanic would. **Soon information was the limiting factor again**, because with enough hardware to store, but not enough information to access, even the best expert systems could still not beat humans (it took Deep Blue over ten years of development to beat world champion Garry Kasparov, for example).

Since the 90s we've gotten smarter in how we build AI though, now modeling a lot more after neural systems in the brain and human genetics – by now AI has made its way pretty far into our daily lives, with smartphones and , for example.

What we're still missing though, is an AI that can, for example, beat not just the best guy in chess, but *also* the best person in Jeopardy and Scrabble – we usually custom-build AI for a specific purpose.

Bostrom and other experts expect computers to be as smart as humans by 2075. Give it another 30 years until 2105 and we'll have true superintelligence.

## Lesson 2: Superintelligence could either imitate or simulate humans, building on biology or technology.

---

What we're currently doing with AI is mostly **teaching computers to imitate human thinking**. Computers use logic to navigate a wealth of information, calculate probabilities and then take shortcuts humans can't come up with to imitate their behavior – just faster.

As described above, this requires access to a lot of information in real time and that's a problem. **An alternative would be to get computers to simulate the human brain, not just imitate it.** This is called WBE – whole brain emulation – and would result in a computer that's like a child: equipped with basic information about the world and the ability to learn the rest on its own. To achieve this we don't even need to decode the entire human brain, we just need to be able to copy it.

However, this would require us to take actual human brains, get the data out of them, and upload it somehow. Sounds like Minority Report? Well, that's also about how far it's away

### Lesson 3: If some secret government program comes up with superintelligence first, we're probably screwed – we all have to work together.

---

Just like there are two ways to technologically make superintelligence a reality, there are also two socially different ways it can be developed.

One is again very similar to what you see in a lot of movies: **some secret government unit or program toils away behind closed curtains for decades, until it emerges with a new piece of highly superior technology** – you know, something like the A-bomb.

In this scenario, a small group would produce *one, single*, superintelligent machine. This would give that country a strategic advantage over all others – but it'd also be a problem. Because if just one unit exists, it only takes one set of evil hands to wipe out our entire species. And if something goes wrong, there aren't enough people who know how to fix it either.

The only way it can really work is the second scenario: **a global collaboration to gradually develop superintelligence, based on humankind working together as one.**

Such a team effort would make sure all steps taken are the safest ones, because many parties and the public control the project, developing safety regulations along the way. It might not be as fast, but it's sure as hell safer.

## My personal take-aways

---

A lot of books about these topics paint a bright future of infinite life, zero work and perfect health. But we have to get there first. And there's a very real chance that if we do it the wrong way, we'll get none of that, and instead be at the mercy of something we built, but have no control over. I like the fantasies, but I try to stay in the present too. Out of all the books on AI, *Superintelligence* is the one you should read first.

[Read full summary on Blinkist >>](#)

[Free Preview >>](#)

Learn more about the author >>

## What else can you learn from the blinks?

---

- Why we're much closer to superintelligence than history would indicate
- Which project should be the model case for developing superintelligence
- How we can keep superintelligence from turning against us
- Why superintelligence would make all human work obsolete
- How we'd keep ourselves busy in such a utopian world
- What sparrows can do to live with an owl and how that relates to our approach to AI

## Who would I recommend the Superintelligence summary to?

---

The 17 year old computer geek, who loves watching sci-fi movies, the 35 year old government employee who smells something funny is up in the office down the corridor, and anyone who's worried about a Matrix-like future.