



Reith Lectures: AI and why people should be scared

**A preview discussed with
Prof Stuart Russell**

<https://www.bbc.com/news/technology-59326684>

November, 2021

Discussion Extracts

Do you have a formula, a definition, for what artificial intelligence is?

Yes, it's machines that perceive and act and hopefully choose actions that will achieve their objectives.

But could a dishwasher not fit into that definition?

It's a continuum.

- Thermostats perceive and act and, in a sense, they have one little rule that says: "If the temperature is below this, turn on the heat."
- All the way up the other end - you have the self-driving cars, where the decision-making is much more complicated, where a lot of learning was involved in achieving that quality of decision-making.

But there's no hard-and-fast line.

We can't say anything below this doesn't count as AI and anything above this does count.

So what is the destination for AI and where are we in terms of that journey?

Clearly, there are some major breakthroughs that still have to happen. And I think the biggest one is around complex decision-making.

- So if you think about the example of building a school - how do we start from the goal that we want a school, and then all the conversations happen, and then all the construction happens, how do humans do that?
- Well, humans have an ability to think at multiple scales of abstraction.



SAINTS PERSPECTIVES

Things Scientific and Technical

- So we might say: "OK, well the first thing we need to figure out is where we're going to put it. And how big should it be?"
- We don't start thinking about should I move my left finger first or my right foot first, we focus on the high-level decisions that need to be made.

Related Articles:

The Economist: June 30, 2021:

[An understanding of AI's limitations is starting to sink in](#)