



# SERIES 2

Artificial Intelligence & Robotics

Chapter 4: AI and ethical considerations

powered by





# WHY IS THERE DISCUSSION ABOUT ETHICS IN THE A.I. SPACE?

There's a lot of anxiety associated with all the unknowns around AI – we don't know where it will take us. Visions abound of some dystopian future in which we are not *using* the tools of technology but have become the tools *of* technology.

Specifically, there's general unease afoot about what's known as Artificial General Intelligence (AGI), AI which is capable of learning tasks in the same way that humans do. Insiders say that while the disruptive effects of AGI cannot be overstated, we remain decades away from the reality. This [McKinsey piece on AGI](#) is worth reading.

HBR expresses the core question well: Do the machines actually “know” things about us, or are they only making informed guesses? And, if they're making an inference, just as any human might do, is there anything wrong with them being so astute?

The answer is yes, potentially. Predictive analytics is about the generation of new data, based on joining dots to find patterns that carry meaning. As a crude exercise, type “school girl” into Google Images. The page is filled with (mainly) hyper-sexualised images. When you type “school boy” there are none. That's because predictive analytics expects one thing of browsers using one search, and something else from the other search – and the Google machines learned that by adding 2+2 across the universe of browsers. Many would argue there's real harm in this.



## SERIES 2

AI & ROBOTICS:  
CH 4: ETHICAL CONSIDERATIONS



# WHAT DOES IT LOOK LIKE WHEN IT GOES WRONG?

Less crudely, the “dots” we put out out via social media and other digital platforms feed into a personal algorithm; all those personalised 2+2s can add up to a solid 4 in predictive analytics. Equally, however, they can add up to 5, and the prediction about you can be used against you. For instance, there was a time when Hewlett-Packard was predictively mapping its more than 300 000 workers against a “Flight Risk” score, i.e. the probability of how long they’d stay with the firm. This was shared with managers, unbeknownst to employees who never necessarily knew whether they had been labelled a flight risk. This cannot have been good for many careers.

Machine learning in the form of facial recognition technology has also famously raised red flags around ethnic repression in China, and privacy & surveillance issues in other countries.



**SERIES 2**

AI & ROBOTICS:  
CH 4: ETHICAL CONSIDERATIONS



# AI AND FAKE NEWS

Fake news was thrust into the spotlight following the 2016 US presidential election, when investigations found co-ordinated efforts by a Russian “troll farm”, the Internet Research Agency, to manipulate the result. The intervening years in which US interest groups used social media to drive unsubstantiated and often deliberately misleading information culminated this month in the extraordinary spectacle of disaffected Trump supporters- all of whom bought into the thoroughly disproven idea that the US election had been rigged – storming the Senate.

Closer to home was the infamous Bell Pottinger fake news campaign bought by the Gupta family, which dramatically escalated racial tensions in South Africa and ultimately led to the PR firm’s “death by embarrassment”. That campaign was built on Twitter bots and attack websites pushing out a toxic narrative. This wasn’t new technology: for the PR firm as early as 2011 it had been boasting about search engine manipulation to people it thought were potential clients. Since then, technology has become ever more sophisticated in terms of generating agenda-driven campaigns, including so-called “deep fake” manipulated videos for commercial or political gain.

Contrarily, AI has also been driving the ability to identify fake news. It’s a tension which is long way off resolution



**SERIES 2**

AI & ROBOTICS:  
CH 4: ETHICAL CONSIDERATIONS



# WANT TO EXPLORE FURTHER?

HBR wrote an excellent piece on the ethics of AI: [find it here](#).  
McKinsey's AI hub touches on many of the issues: [find it here](#).  
Unesco has produced a [Preliminary Study on the Ethics of Artificial Intelligence](#).

powered by



RCS

BNP PARIBAS GROUP

LevelUp  
Business development support programme

**NEXT WEEK: GENERATIONAL**