**Ditchley Conference - 7-9th June**

**‘How do we make the internet safer, without destroying its vitality?’**

Pre-conference reading submitted by Jeffrey Thomas.

**“You have to run faster just to stay in the same place” ~ The Red Queen’s Race, Alice in Wonderland**

We have crossed the Rubicon of convenience and dependency offered by Cyberspace and its associated tools like Machine Learning and Artificial Intelligence (AI). There is no going back. However, with every step into this *Brave New World* we are increasingly aware of the threats and vulnerabilities that we face when confronted by actors with maleficent intent. We are now fully committed to the *Red Queen’s Race*!

This thought piece will explore how Artificial Intelligence (AI) and Cyber Security are likely to evolve and what humanity must do in order to avoid this robo-apocalypse.

**We are all algorithms**

Whether silicon or carbon, synthetic or biological, Yuval Noah Harari in his book *Homo Deus* proclaims that we are all algorithms that can be examined and manipulated as predictable data processing units. The recent spate of democratic manipulation allegations would assert that this idea has some truth. Our growing reliance on cyberspace for all aspects of our modern lives results in a myriad of vulnerabilities which may be exploited by cyber-criminals using AI.

**Running Faster**

The real power of AI is the speed with which it can process repetitive and predictable tasks. AI inspired cyber-attacks are unlikely to look remarkably different from our current experiences - they will simply take place at a speed which is impossible for humans to deal with. This situation is further complicated by cyber criminals using existing AI tools not for their intended usages but with malicious intent. Good examples are the plethora of social media AI tools that highlight likely associations and recommended searches that are used in developing, targeting and executing mass spear finishing attacks.

In *Platform Capitalism,* Nick Srnicek examines the rise of the platform as the means to generate value and how the FANG[[1]](#footnote-1) corporates capitalise on this. These platforms make extensive use of AI to streamline a user experience and maximise consumer convenience but they, too, are experiencing a “*Crisis in Trust*” as the latest Cambridge Analytica investigations suggests. The inevitable bias within their algorithms is comparable to Gresham’s Law. Gresham’s Law is a fiscal phenomena where “*bad money tends to drive out good money”* in a controlled system. The same phenomena are being exploited by criminals who are inducing unanticipated behaviour or forcing system shutdown via “bad data driving out good data”.

**Who Watches the Watchers**

However, it is not just the cyber criminals who need to be watched. Any discussion on cyber security must also consider citizen privacy and we must analyse the use of the tools rather than the tools themselves. In *Weapons of Math Destruction*, Dr Cathy O’Neil provides a convincing suite of societal examples where citizens are being subjugated by algorithms and our humanity is debased to a series of questionable metrics. From US school teachers being sacked despite working miracles in deprived areas to feedback loops driving up stop-and-search figures within social subgroups in the UK - the point here is that we need to keep an “eye on AI” as well as the people who are using it.

**So What?**

The only way that we can deal with these challenges is through a multifaceted approach that involves citizens, corporations, governments and digital leaders.

For the Citizen, the principle activities will be to self-educate throughout their lifetime on how to use the technology, exercise their democratic right to shape digital policy making and report incidents or concerns of AI misuse or unintended consequences. The “crisis in trust” will need to be addressed at citizen / consumer level too. Nick Shapiro, AirBnB’s CRO, recently acknowledged this challenge for AirBnB when he said: “[Trust] doesn't magically take place, it takes work. You have to facilitate trust. You have to earn trust. You have to build trust. And technology plays a huge role in that."

This is a good segway into the role of the Corporate. The Platform Capitalists need to adopt an ethical stance in how they host and publish their content. Within their platforms, every effort must be taken to overcome bias; unintended or otherwise. For Boards, the art of dealing with AI super-charged cybercrime will be to make it an economic and risk-based challenge by making it economically prohibitive for the would-be cybercriminal. This can be achieved through a number of means but comes down to the enforcement of the necessary controls and thresholds.

For Government, the challenge will be to shape the ecosystem to deal with a challenge that is, for all intents and purposes, invisible. The timelines for a cyber breach tends to be too long and the ROI for good cyber hygiene is very difficult to measure for Boards. The secret to dealing with them is to adopt a realistic regulation framework that is policed and enforced. In the longer term, it is necessary to implement liberal paternalism style mechanisms to nudge good cyber behaviour across the entire ecosystem.

The final challenge in keeping up with the *Red Queen’s Race* will be in dealing with Harari’s claim that we are carbon algorithms – it will be critical to keep a “human in the loop” at every level in order to identify false positives, maintain humanity’s unpredictability and creative essence. For this challenge we will need digital leaders who are capable, critical and creative.

1. Facebook – Apple – Netflix - Google [↑](#footnote-ref-1)