PC: Pixabay.com

Intellectual Property, Artificial Intelligence, and Ethical Dilemmas

China and the New Frontiers of Academic Integrity

James DARROWBY

Recent media headlines in Australia and the United States have highlighted the threat of intellectual property (IP) theft by Chinese actors. This reporting has placed particular emphasis on Chinese hackers, as well as military technology or economically-competitive IP developed by Australian/US universities—either through local funds or sponsored by Chinese companies and government—before being transferred to China.

In the wake of the shifting relations between China and the United States, changing strategic interests, and the ongoing trade war, the threat of IP theft and foreign interference has triggered diverse political responses, ranging from contextualised discussions about the actual extent of the issue and its bilateral implications, to localised outright bans of any form of academic engagement with China (see, for instance, US Senate Committee on the Judiciary 2018; Eftimiades 2018; Puko and O'Keeffe 2019). In the extreme, it is not difficult to envisage a descent into a soft form of witch-hunting targeting individual students or academics who may simply be *believed* to have links to foreign governments (Redden 2018; Walsh and Fang 2019).

Still, in spite of all the alarmism, to this day no clear government policies or guidelines have been developed and implemented in either Australia or the United States, leaving many universities and research institutions second guessing at how to best align with the demands from their governments and national security officials in order to ensure that they do not lose their competitiveness and eligibility to attract government funding (Knaus 2019).

This is a difficult balancing act. On the one hand, the critical need to prevent the unscrutinised transfer of sensitive IP and technologies—be it through cyberattacks, or openly funded projects—cannot be underestimated; on the other, an overly restrictive and discriminating approach could result in the loss of any remaining advantage in science and technology for

Critics in the United States and Australia cite instances of IP theft and controversial technology transfers as the clearest indication that universities cannot be trusted as recipients of funding from state sources for sensitive projects.

both the United States and Australia. This has the potential to unleash a domino effect jeopardising all forms of international collaboration.

Critics in the United States and Australia cite instances of IP theft and controversial technology transfers as the clearest indication that universities cannot be trusted as recipients of funding from state sources for sensitive projects. What such analysis omits, however, is how the problem stems directly from successive government policies which have continuously reduced the amount of domestic research and development funding available to universities, while increasing political interference in the funding process.

In this context, universities have been forced to seek and accept funds from any available source, often sidestepping due-diligence procedures associated with national security frameworks. As thresholds and oversight have been lowered under the weight of increasing financial and performance pressures, matters such as the nature of research affiliations with foreign institutions, conflicts of interest, undisclosed double appointments, and the dissemination and application of sensitive project outputs have been buried or gone unnoticed. In this environment, legitimate concerns are frequently brushed aside and framed as being nothing more than conspiracy theories.

Universities have hence become easy soft targets for the development and transfer of sensitive technologies. This is particularly the case in those fields that do not fall under existing frameworks—for example, the Australian Defence Trade Control Acts or the United States National Defense Authorization Act—and whose full applications and developments were not envisaged in different contexts, such as the potential uses of surveillance technologies by autocratic regimes.

It is in this context that China looms large. While it can be reasonably argued that the Chinese authorities may have identified and exploited these loopholes with more savvy than anyone else—and it must equally be stressed that a number of concerns are entirely legitimate—a more constructive approach to the problem would be to cease to treat these matters with politically-driven sensationalism, and rather see them as symptoms of a broader set of issues.

One particular area where serious security and ethical concerns can be raised, and where China is investing massively both domestically and abroad—not least to facilitate its emerging surveillance state—is the development of Artificial Intelligence (AI) tools and their diverse applications, with an emphasis on spyware and surveillance technology.

In industrial and academic settings, international collaborations are critical for development in all areas. There is significant risk of jeopardising long-term benefits and losing major player status in the global research arena because of the adoption of a blanket ban or the imposition of unnecessarily restrictive measures. Most of the projects and collaborations with China that have hit the headlines for the wrong reasons seem to share two common traits: a) a gap in the universities' vetting framework to identify and accordingly manage matters of national interest and security; and b) the absence of multidisciplinary review panels able to detect issues of ethical relevance that reach beyond the strict boundaries of the subject matter.

The last point is particularly reinforced by the fact that many areas of scientific research are siloed behind the bulwark of their respective disciplines. This indicates that universities often lack the ability to facilitate information sharing between academic departments within the same organisation, with the flow of information becoming stuck in myriad structural bottlenecks, making it difficult to gain a holistic understanding of the full implications of many research partnerships.

One particular area where serious security and ethical concerns can be raised, and where China is investing massively both domestically and abroad—not least to facilitate its emerging surveillance state—is the development of Artificial Intelligence (AI) tools and their diverse applications, with an emphasis on spyware and surveillance technology.

Several confidential instances have recently surfaced of Chinese companies approaching foreign universities either directly or through their shadow subsidiaries, and offering funds under the generic banner of 'supporting collaboration between academia and industry'. However, scrutiny of the key areas for proposed collaboration highlights how the research has focussed on the development of the next generation of audio-visual tracking tools, which represent significant potential for military and domestic surveillance applications. For instance, the Tencent AI Lab recently advertised its Rhino-Bird Focused Research Program 'to identify and support world-class faculties pursuing innovative research in areas of mutual interests in Artificial Intelligence' (Tencent UR 2018).

Research topics such as 'recognition technology' focus on multi-camera detection and tracking; video-based facial detection, alignment, and recognition; 3D facial anti-spoofing; and facial recognition with low-resolution or occluded faces. 'Far-field signal processing' seeks to identify, isolate, and enhance speech and speaker recognition, while tracking multiple moving speakers with auditory and visual information in noisy environments. The development of AI tools that have

the potential to be used for mass surveillance is in itself a topic replete with major ethical dilemmas. This is only exacerbated when these tools are developed for, and provided to, an authoritarian regime that controls its population in highly-coercive ways.

In parallel, such collaborative technology development is inherently a national security threat, as it could conceivably be used by the Chinese government to track and monitor both Chinese and non-Chinese citizens and officials. Further, the geographical applications of these technologies would not be confined within the limits of China's borders, but could easily be exported to other countries through investments such as the Belt and Road Initiative.

If the pace of the development of AI tools is now occurring so quickly that it cannot be entirely controlled or prevented, it is equally true that—as for any other risk—it can and should be accordingly mitigated.

As mentioned above, an informed vetting framework necessitates considering both the technical inputs and the specific contexts in which the technologies will most likely be applied. With the case of AI, the gap in the ethical review and evaluation framework appears to be quite significant—a shortcoming that goes well beyond the boundaries of the higher education sector.

The European Union has recently established a High-level Expert Group on Artificial Intelligence (AI HLEG) tasked with the drafting of 'Ethical Guidelines for Trustworthy AI' (European Commission 2019). In the first draft of the document, the authors emphasised that: 'Trustworthy AI has two components: (1) it should respect fundamental rights, applicable regulation and core principles and values, ensuring an "ethical purpose" and (2) it should be technically robust and reliable since, even with good intentions, a lack of technological mastery can cause unintentional harm' (AI HLEG 2018).

The Declaration on Ethics and Data Protection in Artificial Intelligence adopted at the 40th International Conference of Data Protection and Privacy Commissioners, held on 23 October 2018 in Brussels, notes that: 'Many stakeholders in the field of artificial intelligence have expressed their concerns about the risks of malicious use of artificial intelligence … pointing out for example that the development of artificial intelligence in combination with mass surveillance raises concerns about their possible use to curtail fundamental rights and freedoms' (ICDPPC 2018). The Declaration also endorsed the principle that 'artificial intelligence and machine learning

'Many stakeholders in the field of artificial intelligence have expressed their concerns about the risks of malicious use of artificial intelligence … pointing out for example that the development of artificial intelligence in combination with mass surveillance raises concerns about their possible use to curtail fundamental rights and freedoms' (ICDPPC 2018).

technologies should be designed, developed and used in respect of fundamental human rights and in accordance with the fairness principle, in particular by … ensuring that artificial intelligence systems are developed in a way that facilitates human development and does not obstruct or endanger it, thus recognising the need for delineation and boundaries on certain uses.'

While neither of these organisations make any relevant mention of the otherwise concrete issues of national security, the higher education sector could easily add to these perspectives and embrace the same principles through a transparent, independent, and robust framework (Avin and Belfield 2019).

Internal multidisciplinary AI Ethical Committees would provide the most appropriate tool to make a fully-contextualised risk assessment and review the suitability of funded projects aimed at developing AI tools that are either contrary to national security/interests and/or established human rights principles, regardless of the location or proposed scope of the application of said technology.

This could be a small, yet critical first step towards our need to increase awareness of the many unanticipated risks inherent in a rapidly changing 'artificially-intelligent world'. In this modern era, our higher education institutions face a twofold task. First, they need to ensure that their governance and research integrity frameworks are fully aligned with their founding principles of academic excellence while addressing—not compounding—global challenges. Second, they must guarantee that the solutions implemented are not out of step with the technologies developed and collectively deployed on populations around the world. ■