Publication: The Business Times, pg 22
Date: 25 March 2020
Headline: Nudging students to rebut privacy violations

Nudging students to rebut privacy violations

To build a viable ecosystem, educators should effectively engage the youth not only with regard to the opportunities of "good" AI but also its risks. BY THOMAS MENKHOFF

ASK university students here what keeps them awake at night, and chances are many of them might say "I worry about the economic impact of the coronavirus" or "hopefully I will be able to find a job". I bet most undergraduates won't voice concerns about social media safety, unregulated artificial intelligence (AI) or the surveillance capabilities of those who know how to master machine intelligence, be it commercial entities, governments or hackers, including all the tech giants on which many of us rely when it comes to using social media.

The plan to build a viable AI ecosystem in Singapore necessitates that "we educators" effectively engage our youth not only with regard to the opportunities of "good" AI (think customised financial advice based on a person's accounts; AI-enabled improved detection of breast cancer from mammograms; or the deployment of predictive AI to fight poachers) but also its risks. Examples include the potential of Big Data sets producing biased algorithms; encrypted data on smartphones being leaked to unauthorised parties without users knowing it; or sleepwalking into George Orwell's *1984* world characterised by invisible privacy violations, non-stop location-tracking by apps (and the death of democratic socialism as Orwell would have stressed).

The rapidly progressing commodification of human behaviour along with the digitisation of everything creates not only great social media connectivity (enabling binge watching or skyping loved ones overseas) and new business opportunities but also fertile ground for digital surveillance. Related incidences include the case of Julian Assange's Wikileaks, Edward Snowden's National Security Agency whistleblowing, the Siri Spying incident (caused by an external contractor) or the recent activities of Clearview AI – a facial recognition tech startup that grabbed publicly available photos from social media accounts for law enforcement safety purposes. What do we want others to know about us?

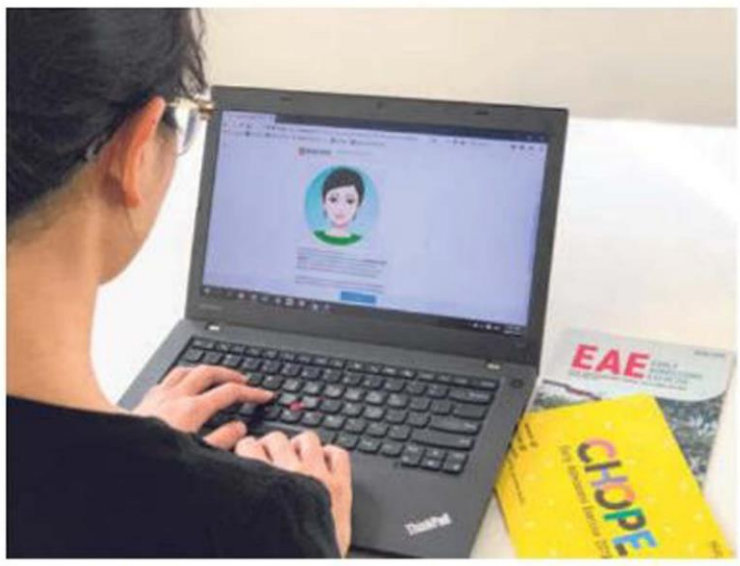
A lot has been achieved in the areas of regulating data privacy protection and AI governance since data breaches and cases of mishandling of personal data surfaced. Local examples include the Personal Data Protection Act 2012, the Centre for AI & Data Governance (CAIDG) established in 2018 in the Singapore Management University School of Law, and the second edition of the "Model AI Governance Framework" released by the Personal Data Protection Commission (PDPC) in 2020.

HUMAN INVOLVEMENT

The governance framework is supplemented by a compendium of use cases and an implementation guide for organisations developed by PDPC and the Infocomm Media Development Authority in collaboration with the World Economic Forum's Centre for the Fourth Industrial Revolution. The framework provides guidance to organisations so that they can deploy AI in a responsible manner with regard to internal governance structures and measures, human involvement in AI-augmented decision-making, stakeholder interaction and communication, and operations management.

One interesting use case features the "Early Admissions Exercise Virtual Assistant" (EVA) piloted by Ngee Ann Polytechnic aimed at making admissions selection easier with responsible AI. Its AI-powered platform and chatbot function is considered to be human-centric – that is, even though EVA might reject a candidate, a "human-in-the-loop" will ensure that unselected applications are reviewed and shortlisted in case they are suitable (in line with the principles of responsible AI).

While such governance efforts are critical and certainly laudable, more needs to be done to bring about high trust in AI technology. It's one thing to claim that third parties are not allowed to make use of user data for surveillance purposes as part of corporate data and privacy policies, but another to see contractors violating "strict con?dentiality obligations".



One interesting use case features the "Early Admissions Exercise Virtual Assistant" (EVA) piloted by Ngee Ann Polytechnic aimed at making admissions selection easier with responsible AI. PHOTO: NGEE ANN POLYTECHNIC

A well-known case in point is the 2018 Cambridge Analytica scandal. The UK-based data analytics and political consulting firm (that worked on Donald Trump's 2016 presidential campaign) had obtained data from millions of Facebook users worldwide without permission. Mark Zuckerberg seemingly had a challenging time relating to the persistent queries by US Congresswoman Alexandria Ocasio-Cortez during the 2019 Congressional hearing. What followed was a series of visits by Facebook's CEO to overseas markets to assure "stakeholders" that it takes user privacy seriously and that the company will do the right thing in future.

In Germany, Facebook partnered the Technical University of Munich (TUM) to create an Institute of Ethics in Artificial Intelligence (the social media enterprise has committed US$7.5 million to the institute over five years).

In the meantime, Facebook has published its own white paper on regulating content – arguably in anticipation of increasing legal platform liabilities. Whether this will enable Facebook to make better ethical decisions in future remains to be seen.

Maintaining Internet civil liberties in the digital world is a complex challenge. While China's government (which regards security as a necessity for surveillance) is testing a new (national reputation system) plan to urge its citizens "to do more good", the US Digital Rights Group The Electronic Frontier Foundation (EFF) defends Internet civil liberties on the basis of *pro bono* litigation work and annual transparency reports. EFF scrutinises and ranks the big Internet companies in areas such as "transparency in reporting government takedown requests based on legal requests" or "providing meaningful notice to users of content takedowns and account suspensions".

According to EFF, many firms do not enact best practices around transparency or don't find it important to stand up for user privacy. Regulatory (and investment-oriented) concerns about the ethical development of AI, privacy, accuracy, safety and fairness are key drivers behind the European Union's new "White Paper on Artificial Intelligence – A European Approach to Excellence and Trust". How the European Commission will manage the balancing act of promoting the deployment of "trustworthy" AI while simultaneously addressing all risks will be interesting to observe.

A related educational challenge is to convince undergraduates that invasive digital stalking is immoral and to enable "stalkees" to be aware of what big tech companies can (and sometimes actually) do with the kinds of insights they have gained about them. Nudging both parties to enact corporate digital responsibilities ethically and firmly rebutting ubiquitous surveillance is easier said than done. Many digital natives arguably do take things for granted, having been brought up in a political culture of government responsibility. Not knowing how AI actually works "under the hood" and how the various types of cyber attacks such as "Man-in-the Middle attacks" unfold in reality make things more complicated.

LOOKING BACK IN HISTORY

One first responder approach which non-tech educators can use to turn students' apathy into deep(er) concerns about privacy-invading tech is to invite them to look back in history and to consult some of the great classical scholars (besides Orwell) in order to make sense of what is really happening in society.

Potential candidates include German sociologist Max Weber's (1886-1920) theory of power and domination (for example, to assess both the pervasive influence and socio-digital responsibilities of big tech companies who are pushing the AI agenda forward); French sociologist Emile Durkheim's (1858-1917) functionalist view of capitalist society with its importance of social integration and commonly held norms-cum-values aimed at avoiding anomie from arising (for example, to acknowledge the merits of social policies such as SkillsFuture); or English historian EP Thompson's (1924–1993) notion of a moral economy based on "goodness, fairness, and justice" contrary to one where uncontrolled market forces commodify everything (for example, to create empathy for public policymakers, AI regulators, critical moralists and digital tax collectors).

Turning to the wisdom of modern management gurus such as Babson College's Tom Davenport and his new book *The AI Advantage*, there is no doubt that knowing "how to put the AI revolution to work" is an asset indeed.

But one cannot ignore the opposite aspect of tech might such as the possibility of unexpected harm caused by an adversarial AI system and discriminatory algorithms or the generally insufficient participation of AI-conversant consumers and employees in developing AI governance regulations, a process which is dominated in many countries by industry rather than ethicists and an informed public. The long-term consequences of many new AI technologies enabling faster diagnoses in healthcare, connected home devices or algorithmic journalism are simply unforeseeable.

The writer is professor of Organisational Behaviour and Human Resources (Education) at Lee Kong Chian School of Business, Singapore Management University.