WHY ETHICAL GUARD RAILS ARE VITAL IN A WORLD RUN ON BIG DATA

By Lim Sun Sun

For The Straits Times

As the public hearings for the Select Committee on Deliberate Online Falsehoods draw to a close, the realm of views that have been considered has stretched far and wide.

Experts from a host of countries in areas as diverse as communication, computer science, foreign policy, international relations, law, and religious studies have shared their insights on the risks and dangers of online falsehoods, as well as the measures needed to tackle them.

Yet this is a discussion that extends well beyond online falsehoods and speaks to the growing role of big data and the technologies undergirding it. Crucially, it shines the spotlight on the ethical principles that must be engineered into the technologies that pervade our digitally-connected lives.

That the Cambridge Analytica scandal should erupt during the Select Committee hearings has seen the issue of big data being conflated with the ills of fake news. Yet this connection is by no means a trivial one.

The ferocious proliferation of fake news that has triggered legislative responses around the world has been largely powered by big data.

The algorithms that determine which stories appear in our social media newsfeeds, which photographs garner more shares, or which videos are more often suggested, are built on the backs of data that technology companies have been stealthily collecting about us and our behaviour, both online and offline.

THE HAZARDS

Such algorithms are also designed to enhance the profit generation of the companies that convey or host such content, explicitly calibrated to favour whatever is sensationalist, alarming, or extreme, so as to drive up views and boost advertising revenue.

In her recent op-ed, academic Zynep Tufecki shared her disturbing experience on YouTube where after watching a video on a particular topic, the platform proceeded to suggest related videos that veered towards conspiracy theories and hoaxes. She surmised that YouTube actively promotes such outrageous and incendiary content because it lures viewers to stay on the platform for longer periods of time, thus exposing them to more commercials.

Hence, in a media landscape that is principally motivated to fatten corporate coffers, fake news stories that raise eyebrows and draw gasps are destined to gain traction at a rapid pace.

Beyond profit motivations, actors with nefarious intentions have been quick to exploit the virality of fake news, seizing on inherent societal fissures to produce and disseminate online falsehoods that can drive a greater wedge between different groups.

Hyper-connected countries such as Britain, France, Germany, the United States, and Singapore are thus grappling with the wide-ranging effects of fake news and are seeking nimble and practicable legislative or regulatory solutions.

And yet fake news is but the canary in the coalmine, signalling the grave dangers of big data if misused, mismanaged, and misgoverned.

BIG DATA AND BIG TECH

To return to the Cambridge Analytica scandal, it highlights the massive scale to which big technology companies have been allowed gain insights into individual consumers, thereby possessing overwhelming influence over our everyday lives.

Through innocuous tools such as personality quizzes, location check-ins and online cookies, technology companies such as Facebook, Google and Amazon have been able to collect vast amounts of data on individual users' online behaviour, demographic and psychographic profiles, ideological leanings, and consumption preferences, just to name a few.

These companies then use such data to offer us a customised online experience that we may often find beneficial. When I look up a holiday to Poland, for example, advertisements for hotels in Warsaw will magically appear, along with recommendations for day tours and key landmarks. But what is the price we consumers pay for these fleeting conveniences? In this instance, the very data that is used to offer us discount hotels could be parlayed into computing overpriced travel insurance packages.

As the Cambridge Analytica scandal has sharply revealed, consumers are largely disempowered in determining how the data that is collected about us is categorized, utilized, and monetized, for good or for ill.

Rightfully therefore, concerns are growing about the avid use of big data in sectors as varied as retail, healthcare, education, and finance.

There are utopian visions surrounding big data as the panacea for ill-formed policies based on subjective discretions and systemic inefficiencies.

Yet critics assert that if the data used to generate technologically-grounded solutions reflect society's inherent biases, these distortions will simply be reproduced and exacerbated by the resulting algorithms.

As mathematician Cathy O'Neil revealed in her book Weapons of Math Destruction, seemingly neutral and dispassionate algorithms can cause severe problems such as unfair dismissals, discriminatory product pricing, and inequitable social policies.

Nevertheless, let us not throw out the baby with the bath water. Big data is certainly of value but the question arises as to how it remains within the wrestle-hold of a handful of big technology companies.

How can we allow data on so many, to be controlled by so few? To enhance the benefits that big data can confer on society, the ethical principles that underpin the collection and use of big data must be clearly identified, enunciated and practised.

CODES OF CONDUCT

As big data becomes heavily incorporated into other innovations such as the Internet of Things, machine learning, and robotics, the imperative for such principles to be clarified becomes ever more pressing. These ethical principles must further be incorporated into industry codes of conduct and a broader regulatory framework that involves greater scrutiny and oversight.

Beyond just mobilising the regulatory playbook, we can wire ethics into the digital infrastructures we create by instilling these positive values in our next generation of technologists. At the Singapore University of Technology and Design, 22 per cent of our engineering and architecture students' curricular exposure is dedicated to humanities and social sciences courses.

Through understanding diverse perspectives from disciplines such as anthropology, history, literature, philosophy, psychology, and sociology, our students develop a firm grasp of how technology can have significant impact on society as it is adopted, incorporated, and socially shaped.

They are also imbued with the principles of design thinking that marry creativity with empathy, and innovation with responsibility. They are thus sensitized to the ethical dilemmas and opportunities that software and hardware engineers and architects face today.

For example, a hardware engineer who designs a fitness tracker has to assess the extent of geolocation tracking the device enables so that consumer privacy is preserved, while the software engineer who develops the accompanying app must consider the degree to which users' health data can be shared with third parties for diagnostic or marketing purposes.

As irksome as the problem of fake news has been, it has greatly illuminated the issues surrounding big data and its growing presence in society, as well as our urgent need to forge an ethical world by design.

Lim Sun Sun is professor of media and communication and head of humanities, arts and social sciences at the Singapore University of Technology and Design.