I've got technology… get me out of here! : The privacy risks of an artificially intelligent exit to lockdown

Ali Vaziri, Managing Associate at Lewis Silkin LLP, looks at the privacy and data protection risks associated with the use of surveillance technologies, such as facial recognition and location and contact tracing systems, in the context of the current health crisis

The English Divisional Court, in a landmark case from 2019 on the police's controversial use of facial recognition technology (FRT), opened with this observation: "The algorithms of the law must keep pace with new and emerging technologies". The pace of technological progress is accelerated by war; and many have used the language of war to describe the current pandemic, with talk of frontlines, field hospitals and the battle against an invisible enemy. It is certainly as deadly. But as the economic and social costs of the coronavirus lockdown are increasingly seen as too high a price to pay for the reduction in direct mortality rates, to aid with exiting the lockdown, 'new and emerging technologies' are being developed and adopted hastily – even in societies which would previously have balked at the prospect of, for example, biometric surveillance.

At this time of global emergency, as the historian Yuval Noah Hariri recently put it, "Immature and even dangerous technologies are pressed into service, because the risks of doing nothing are bigger." So as UK plc fumbles its way towards 'a new normal', this article aims to explore some of the surveillance technologies likely to be contemplated by organisations on that journey, and the privacy and data protection risks associated with their use.

An artificially intelligent exit

Before the World Health Organisation (WHO) issued its official warning on the outbreak of a novel coronavirus in early January, it seems that algorithms on the other side of the world had already detected it.

BlueDot and Metabiota, start-ups specialising in infectious disease surveillance, leverage artificial intelligence (AI) and machine learning techniques such as natural language processing (NLP) to sift through various datasets such as flight itineraries, news outlets and official healthcare reports in different languages around the world. BlueDot is reported to have notified its clients of the outbreak days before the WHO's warning; Metabiota seemingly went on to predict the countries most likely to report new cases.

Big data and AI have been causing a paradigm shift in most sectors. Advances in technology have resulted in a dramatic increase in the quantity of data created through the digitisation of pretty much everything. Advanced algorithms and improvements in computing power have improved our ability to collect, store and analyse those data to produce actionable insights. Telehealth, medical imaging and drug discovery are just a few areas where AI is being applied in the fight against coronavirus. But it is AI-powered surveillance, to screen and monitor populations with a view to controlling the outbreak, which has become a hot topic.

Body temperature detection systems, for example, use thermal cameras whose algorithms can detect elevated facial temperatures – potentially interesting since a fever is one of the key symptoms of coronavirus. These systems can continuously scan high volumes of people in real-time (avoiding delays through manual testing) and can do so remotely (avoiding contact). They are commonly combined with FRT to identify individuals for building access control.

Other AI-powered tools can detect whether someone is wearing a mask – something Uber is reportedly rolling out to enforce mandatory mask wearing by its drivers. And, in case you thought that FRT (as opposed to detection) does not work on partially covered faces, think again: AI has been developed which is reportedly able to operate effectively on partially covered faces with a high degree of accuracy. It is also worth noting that, when it comes to using computer vision uniquely to identify people, some solutions do not even use facial features – instead, their neural networks use other visible parameters such as clothing, body shape, hairstyle and items in their possession.

Location and contact tracing apps are another type of tool being considered as potentially useful in the context of the health crisis. Some such apps are intended for use by the general public, while others are intended only for closed user groups, such as within the workplace.

The apps tend to fall into three groups:

(i) to inform users (e.g. advising through self-diagnosis questionnaires);

(ii) to inform on users (e.g. monitoring and enforcement of quarantines); or

(iii) to warn and trace users who have been in proximity to an infected person.

But even those apps work in different ways. For example, different types of location technology might be used: some apps use GPS to track users' locations; others use short-range technologies such as Bluetooth or WiFi to monitor and collect data on other enabled devices nearby. Some apps are designed to rely on a centralised server to store and process data; others take a decentralised approach where the processing takes place locally on users' smartphones. Some solutions, avoid smartphones altogether. Instead, individuals are issued with badges, key-rings or wristbands embedded with Bluetooth beacons which record when they come into close proximity with another.

> ***"irrespective of the solution proposed, or who is proposing it, trust is key. Without trust, the relationship between the surveillant and the surveilled risks being undermined – whether in the private sector or in the context of our 'social contract' with the state"***

Who am (A)I?

Since most solutions currently in contemplation will be powered by AI, it is worth unpicking that and other related terms before looking at some of the risks. AI is defined by the ICO in its 'explAIn' guidance (co-badged with The Alan Turing Institute) as "an umbrella term for a range of algorithm-based technologies that solve complex tasks by carrying out functions that previously required human thinking." Examples of such tasks include visual perception, speech recognition, decision-making, and translation between languages. AI systems are built by creating an algorithm that uses data to model some aspect of the world, and then applying that model to new data in order to make predictions about it.

Machine learning is a subset and most widely used form of AI which trains a machine how to learn by looking for patterns in data and then applying them. Machine learning can be supervised, where an algorithm is trained using labelled datasets which map input variables onto desired outputs. Through those examples, the algorithm is able to identify patterns that link inputs to outputs and then, once trained, reproduce those patterns to transform new inputs it receives into predictions. Karen Hao of MIT Technology Review says: "Think of it as something like a sniffer dog that will hunt down targets once it knows the scent it's after."

In more sophisticated techniques, such as unsupervised and reinforcement machine learning, data are not labelled. In unsupervised learning, the machine is left to look for whatever patterns it can find in the input data, without any instructions about what to look for. Hao likens unsupervised machine learning to "letting a dog smell tons of different objects and sorting them into groups with similar smells." Reinforcement learning involves an objective being set, and the algorithm learning through trial and error by trying out lots of different things and being rewarded or not depending on whether what it tries out helps or hinders achieving that objective. Continuing with Hao's example, it is like "giving and withholding treats when teaching a dog a new trick."

Deep learning is a subset of machine learning – described by Hao as "machine learning on steroids" – and is behind many scientific breakthroughs. These include computer vision where the AI can 'see' and identify what it is seeing; and NLP where it can 'read' text and interpret what it is reading. Most deep learning models are based on artificial neural networks, inspired by the inner workings of the human brain – they are 'deep' because of the multiple layers of computational nodes in the network. This gives machines an enhanced ability to find and amplify even the smallest patterns. The trouble is, however, that the inner workings and rationale of these so-called 'black box' algorithms can be opaque and inaccessible to human understanding.

'Face' the music

Concerns about surveillance technologies interfering with privacy and data protection rights have, in recent years, been before regulators and the courts. Such was the case, for example, in relation to the British government's collection and use of bulk personal datasets and bulk communications data. FRT has, more recently, been the lightning rod. South Wales Police's use of live FRT at various large public events to locate persons of interest was considered in the case of *R (Bridges) v South Wales Police [2019] EWHC 2341 (Admin)* – the first time any court in the world had considered that technology.

It involves capturing digital images of faces of members of the public from live CCTV feeds. Those images of faces are then processed in real time to extract unique facial features and to create unique biometric templates which are then compared with the unique biometric templates of persons on an existing watch list and, if there is no match, immediately deleted. The Court found that there had been an interference with the claimant's right to privacy given "the intrinsically private" nature of biometric data and the intrusive nature of the processing which "goes much further than the simple taking of a photograph". The interference was, however, justified by the public interest in harnessing new

(Continued on page 12)

(Continued from page 11)

technologies to aid the detection and prevention of crime, and had been deployed in a lawful and proportionate way.

The case came at a time when FRT regularly made the headlines, with an investigation launched by the ICO into its use at a development near King's Cross, and news reports about its use by the Home Office for passport photo checking apparently in the knowledge that it was unable to recognise the faces of some ethnic minorities. Racial bias is a well-documented issue and the Information Commissioner has observed that "facial recognition systems are yet to fully resolve their potential for inherent technological bias; a bias which can see more false positive matches from certain ethnic groups." Onfido, an identity verification company, is using its participation in the ICO's Sandbox as an opportunity to address this issue.

The *Bridges* case also highlights differing world attitudes to FRT. San Francisco banned use of the technology by police and other city agencies in May 2019, with other US cities swiftly following suit.

On the other hand, the Datainspektionen, Sweden's Data Protection Authority (DPA), has approved its use to fight crime. Earlier that year, the Datainspektionen also issued its first GDPR fine to a municipality which ran a short pilot where FRT was used in a school to keep track of student attendance. Although consent had been sought, it was not a valid legal basis given the imbalance between the data subject and the controller (an issue which is also likely to be relevant in the workplace). Further, the data protection impact assessment ('DPIA') was inadequate.

CNIL, the French DPA, had to grapple with a similar issue following the submission of a DPIA by a regional authority where FRT was proposed to control access to two high schools. Despite students' consent, the project was deemed unlawful. CNIL focused on the lack of proportionality given that the purpose could, in its view, be achieved in a less privacy-intrusive manner, such as by checking badges. In reaching its decision, CNIL considered the particular sensitivity of biometric data, the intrusiveness of FRT, and the importance of special protection afforded to children. The Administrative Court of Marseille subsequently heard a case against the regional authority. It found that the students' consent was not freely given (echoing the Datainspektionen) and that FRT was a disproportionate way to control school access.

What are the privacy and data protection risks?

Privacy and data protection risks will ultimately depend on the type of technology under consideration, as well as its purpose and the context of its use. But irrespective of the solution proposed, or who is proposing it, trust is key. Without trust, the relationship between the surveillant and the surveilled risks being undermined — whether in the private sector or in the context of our 'social contract' with the state. When it comes to surveillance, in particular, a good level of trust is crucial: a 2013 study cited by the ICO found there already to be a "widespread wariness" about being spied on by government, corporations and criminals.

> ***"the starting point, when considering tools which are likely to be privacy intrusive, is a risk assessment and, where required, a DPIA. These help analyse, identify and minimise a project's data protection risks. DPIAs are an essential part of an organisation's accountability obligations under the GDPR, and are mandatory where processing is 'likely to result in a high risk'"***

Trust needs transparency, to ensure that individuals understand the 'whys' and 'hows' around the collection and use of their personal data. It needs fairness, to ensure that those data are used in line with individuals' reasonable expectations and in considering the effect on them. Discrimination is a real risk, not just through algorithmic bias, but also in the way individuals are treated if, in the context of the current health crisis, they are linked with the disease, be it through stigmatisation or limiting access to spaces and opportunities.

Control granted to the surveilled should not be illusory either, which is why consent will often be inappropriate where there is no real choice in the matter. 'Garbage in, garbage out' is a known Achilles heel of AI, and inaccurate predictions, recommendations or classifications from it risk creating a false sense of security. As we have seen with serological testing: a bad test is worse than none.

Trust also means being able fully to justify the collection and use of personal data in the first place. Especially when those data are inherently sensitive, as is the case with biometric, health and location data. Recent CNIL guidance reminds us that verifying temperatures through the use of a contactless manual infrared thermometer at a site entrance, without making any record, would not be subject to data protection law. In fact, on-site temperature testing might be avoided altogether by asking individuals to self-monitor and stay home in the event of a raised temperature or feeling unwell. The effectiveness of such strategy will depend on organisational culture, as well as factors such as presenteeism resulting from heightened job insecurity.

Not only should organisations be able to make the case for using surveillance but — as seen above in relation to FRT — it needs to be necessary and proportionate to the end being pursued. If there is a less intrusive way to achieve a particular purpose, then proportionality is likely to be an issue. Necessity includes carefully assessing whether technology can in fact help meet a stated purpose, otherwise it risks being adopted solely for optics, without any real utility. Technology's limitations should be understood, as it is rarely a silver bullet. For example, 'Google Flu Trends', big data's former poster child which sought to predict flu outbreaks in real time by analysing search engine queries, did not work as hoped. And whilst body temperature detection systems might be useful to screen in certain environments, they are not appropriate for diagnosis given that imaging only detects skin (not internal) temperature, the coronavirus has a long incubation period with asymptomatic carriers, and not all sufferers will experience a fever.

The length of time data are kept – mindful, in particular, of the period of infection and the eventual passing of the threat – is also relevant to trust. Temporary adoption often becomes the new normal, and institutionalisation means that technologies remain in place even after the problems for which they were developed have gone away. Without explicit 'sunset' provisions, systems are less likely to be decommissioned once the coronavirus threat has passed.

Minimising the amount of data collected, and limiting the purpose for that collection, helps to allay concerns about 'scope creep' where data collected for one purpose are used for another – a particular problem for big data analytics. Concerns have been raised, for example, about the NHSX contact tracing app being used to build a 'social graph' of physical contacts which might later be used – either by the state, private sector or hackers – to spy on citizens. Take from it what you will that GCHQ has been advising on the app's architecture. Following the Cambridge Analytica scandal and a volley of 'mega' data breaches which have made headlines in recent years, individuals are also now, more than ever, concerned to ensure that their personal data are protected; i.e. not shared without first putting safeguards in place, and kept secure.

(A)I need a DPIA

All of this means that the starting point, when considering tools which are likely to be privacy intrusive, is a risk assessment and, where required, a DPIA. These help analyse, identify and minimise a project's data protection risks. DPIAs are an essential part of an organisation's accountability obligations under the GDPR, and are mandatory where processing is 'likely to result in a high risk'. This term is undefined but the ICO's list of illustrative triggers (which need to be combined with one of the other nine in European Guidelines) include:

- use of 'innovative technologies' (AI, machine learning and deep learning are specifically called out);
- processing biometric data (FRT is specifically mentioned); and
- tracking individuals' geolocation or behaviour (processing employees' location data is cited). However, in some cases an intended use of innovative technology requires a DPIA even without other triggers.

DPIAs also support compliance with another GDPR obligation: data protection by design and default. This requires organisations to bake data protection into their processing and business activities. The obligation extends to selecting and using applications, services and products which involve processing personal data. So when looking to purchase a new system, the ICO is clear that you should look to choose one where the designers and developers have taken data protection into account.

Users and vendors of AI systems will also need to consider the risk of bias, and may find the ICO's AI Auditing Framework helpful. In some circumstances, an Equality Impact Assessment may be required.

Where processing involves special category personal data, such as biometrics and information about health, an 'appropriate policy document' will also likely be needed. Once implemented, systems should be regularly tested to ensure compliance. And remember that DPIAs are living documents which should be subject to continual review.

This pandemic is clearly an unprecedented challenge to our healthcare system, way of life, economic stability and values. But with machine-assisted ingenuity and innovation, coupled with human resolve and our ability to adapt, the challenge will surely be met. Regulators are often keen to emphasise that privacy is not a 'zero-sum game', the proposed trade-off usually being between privacy and security, or privacy and innovation.

Today, the choice we are seemingly presented with is between privacy and health; and for many, the answer is simple: health trumps. But whether it is through the "algorithms of the law" or accountability tools such as DPIAs, we have the means to try to find a healthy balance. After all, as Harari warns: "we should ask ourselves not only how to overcome the immediate threat, but also what kind of world we will inhabit once the storm passes."

Ali Vaziri
Lewis Silkin LLP
ali.vaziri@lewissilkin.com