The robot judge will see you now

Could artificial intelligence regulate itself?

Simon Chesterman

For The Straits Times

As artificial intelligence (AI) transforms modern life, regulators around the world are grappling with how to reap its benefits while minimising its risks.

Earlier this year, the European Union proposed a draft regulation on certain aspects of AI. Last month, China adopted its new state-centric Data Security Law, even as the Group of Seven pushed its alternative vision of an open and "human centric" approach to AI through the Global Partnership for AI (GPAI).

These government initiatives complement the hundreds of guides, frameworks and principles put forth by think-tanks and industry associations – even the Pope endorsed a set of principles last year in the Rome Call For AI Ethics.

As the difficulty of actually applying such rules to fast, opaque and autonomous systems becomes clearer, however, a radical suggestion is beginning to gain credence. Could the machines regulate themselves? And might they one day regulate us?

1. All rise

The idea that legal practice – seen as the logical application of rules to established facts – could be automated has been around for decades. In the 1980s, researchers developed prototype systems of rules in machine-readable form. The enthusiasm was characteristic of the time, preceding as it did one of the "AI winters" that has periodically seen inflated expectations crash against reality.

Subsequent decades did see transformations in legal research and document management. These increased lawyers' access to information and their efficiency in using and sharing it, but did not fundamentally alter their role.

Even those encouraging the adoption of technology believed that the inability of AI to emulate human qualities limited its scope for taking on the higher functions of lawyers – the role of judges in particular.

As we have seen in other areas, however, emulating human methods may not be the right or the best approach for reaping the benefits of AI. Autonomous vehicles, to pick an obvious case, are not driven by humanoid robots controlling speed and direction with mechanical hands and feet in substitution of their absent "drivers".

The DoNotPay chatbot, launched in 2015, offered an indication of what might be possible. Written by a 17-year-old Stanford student, it followed a series of rules to appeal against parking fines. Similar technology now facilitates other simple tasks – from the making of wills to reporting suspected discrimination – yielding efficiencies as well as offering greater access to basic legal services for the wider public.

Many lawyers long assumed that litigation would be the last part of legal practice to be automated. Yet, online dispute resolution (ODR) has been around since the 1990s and, for smaller claims in particular, has been embraced not only by online traders like eBay and PayPal, but also in the legal systems of Canada and Britain.
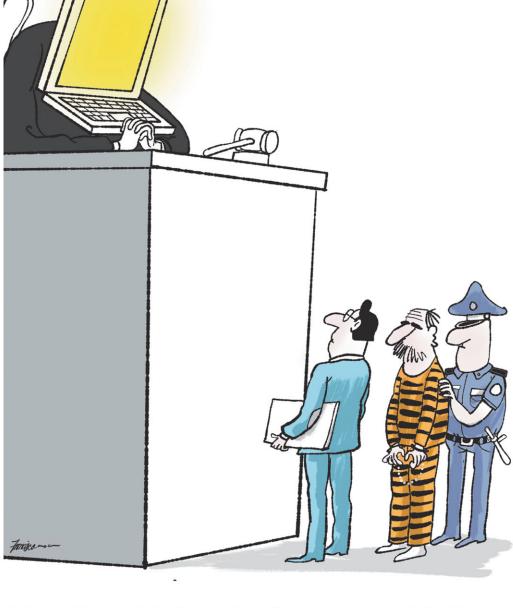
At its simplest, ODR merely helps parties settle their own disputes more efficiently. An algorithm might, for example, serve as the go-between to parties as they negotiate a settlement, nudging them until they reach a mutually acceptable result. A more elaborate system might analyse relevant data and propose its own settlement. In either case, the outcome is binding because the parties themselves agree to it.

Deeper inroads have been made in China.

In late 2019, Hangzhou's Internet Court began allowing parties to appear virtually before an avatar judge. The avatar can handle online trade disputes, copyright cases, and e-commerce product liability claims.

Hangzhou was chosen because it is the home of Alibaba, enabling integration with trading platforms like Taobao for the purpose of evidence gathering as well as "technical support".

Meanwhile, the Wujiang District of Suzhou has trialled a "one-click" summary judgment process for less complex criminal cases, automatically generating proposed grounds of decision complete with sentence.

Singapore's Chief Justice Sundaresh Menon has said that such developments in China are making "machine-assisted court adjudication a reality". At the same time, he noted, the use of AI within the justice system gives rise to a "unique set of ethical concerns, including those relating to credibility, transparency and accountability".

To this one might add considerations of equity, since the drive towards greater automation is being dominated by deep-pocketed clients and ever-closer ties to technology companies, with uncertain consequences for the future administration of justice.

2. Objection!

Inherent in many of the debates over how far this should go are fundamental differences in the understanding not of AI, but of law.

If law is understood in a narrowly formalistic way – the blind application of rules to uncontested facts – then processing it through algorithms makes sense, in the same way that it would be inefficient to have regulators or judges doing long division by hand instead of using a calculator.

But, to state the obvious, law is not long division.

The simplest of cases aside, the regulation of behaviour and the resolution of disputes involve values and meaning that are necessarily contested. As American jurist Oliver Wendell Holmes famously said, "The life of the law has not been logic: it has been experience."

Ah yes, the computer scientist might respond. But experience is precisely what machine learning can replicate now.

Indeed, more recent innovations reflect a shift in the approach to the law analogous to the move in AI research towards machine learning.

Rather than trying to code legal rules in fixed systems that can then be applied to sanitised facts – top down, as it were – key achievements have been made in analysing large amounts of data from the bottom up.

This approach does not seek to answer an individual case, but to offer an estimate based on past experience. By feeding in relevant case law and then comparing it to the facts at hand, the computer system looks for similarities and projected outcomes.

The turn to AI in this context has proven useful in identifying relevance for the purposes of legal research, contract review, and discovery.

Yet if extended to regulation and adjudication, it would fundamentally change the task from making a decision to predicting it.

Rather than being part of an ongoing social process in the development of the law, it would be more like forecasting the weather – or the manner in which Netflix suggests you might like a movie based on what you have enjoyed in the past.

3. Sentence first, verdict afterwards

There may, however, be a case for AI playing a larger role in regulating AI itself.

The speed, opacity and autonomy of AI systems do occasionally give rise to practical and conceptual difficulties for human regulators.

In some cases, the response has been to slow them down, as in the case of high-frequency trading.

In others, it has been to ensure the possibility of accountability through requiring that actions be attributable to traditional legal persons – typically the owner, operator, or manufacturer. In still others, it has been to call for prohibiting certain activities entirely – such as the use of lethal force.

The unique features of AI do suggest two avenues for a form of self-regulation.

First, regulatory objectives can be built into the software itself.

Like Isaac Asimov's famous "laws" these would not really be rules saying what a machine should do, but design features limiting what it could.

Second, AI systems allow for interrogation of mistakes and adverse outcomes in a manner not possible with traditional legal actors.

If you ask a machine whether it is biased, for example, it will generally tell you the truth.

Provided the instructions were clear, a system could report on its compliance with rules and policies, among other things examining its conduct with a degree of candour not possible with humans.

Problems disclosed in this way would also point to a need to rethink the remedies available – not as sins to be punished, but errors to be debugged.

4. Court adjourned

It remains to be seen whether China represents the future of regulation by AI or its limit case.

Even if AI systems are more efficient and consistent than human regulators and judges, that would not justify the handover of their powers more generally.

For the authority of law depends not only on its processes in a formal sense but in a substantive sense also. Regulation, legal decisions are not mere Turing Tests in which we speculate whether the public can guess if the regulator or judge is a person or a robot.

Legitimacy lies in the process itself, the ability to tie the exercise of discretion to a being capable of weighing uncertain values and standing behind that exercise of discretion.

Accepting otherwise would be to accept that legal reasoning is not a mix of doctrinal, normative and interdisciplinary scholarship. Rather, it would come to be seen as a kind of history – the emphasis on appropriate categorisation of past practice rather than participation in a forward-looking social project.

As Robert H. Jackson, another US Supreme Court judge, once observed: "We are not final because we are infallible, but we are infallible only because we are final."

Many decisions might therefore properly be handed over to the machines. But the final exercise of discretion, public control over the legal processes that regulate our interactions with the world around us, should be transferred only when we are also prepared to transfer political control also.

stopinion@sph.com.sg