SYMPOSIUM: HOW WILL ARTIFICIAL INTELLIGENCE AFFECT INTERNATIONAL LAW?

TAMING THE DIGITAL LEVIATHAN: AUTOMATED DECISION-MAKING AND INTERNATIONAL HUMAN RIGHTS

Malcolm Langford[*]

Enthusiasm abounds about the potential of artificial intelligence to automate public decision-making. The rise of machine learning and computational text analysis together with the proliferation of digital platforms has raised the prospect of "robo-judging" and "robo-administrators." From a human rights perspective, the reaction has been mixed, and on balance negative. Optimists herald the possibilities of democratizing legal services and making decision-making more predictable and efficient.[1] Critics warn, however, of the specter of new forms of social control, arbitrariness, and inequality.[2] This essay examines the concerns over the turn to automation from the perspective of two international human rights: the rights to social security[3] and a fair trial. It argues that while the critiques deserve a full hearing, they should be evidence-based, informed by an understanding of "technological systems," and cognizant of the trade-offs between human and machine failure.

The Long Road to Automation

The dream of automating judicial and administrative processes is not new. It dates to at least the first wave of law and artificial intelligence in the 1970s. Drawing on the similarity between the deductive logic in law and computer programming, scholars and others in the "expert design" movement developed a range of prototypes and rudimentary applications.[4] For example, Sergot "automated" the British Citizenship Act by guiding users through an ordered set of questions to the correct legal result.[5,6] And already in 1972 Norway was using "fully automated legal decision-making" to calculate benefits under housing laws.

However, progress was dampened in the 1990s by the onset of the so-called winter in artificial intelligence and law. The complexity and bespoke nature of law challenged the programming paradigm while application was

[1] Richard Susskind, Tomorrow's Lawyers: An Introduction to Your Future (2d ed. 2017).

[2] *See, e.g.,* Frank Pasquale, *A Rule of Persons, Not Machines: The Limits of Legal Automation*, 87 Geo. Wash. L. Rev. (2019).

[3] For example, Article 9 of the International Covenant on Economic, Social and Cultural Rights recognizes the "right of everyone to social security." The United States has signed but not ratified the Covenant. The Universal Declaration of Human Rights also includes the right to social security.

[4] Kevin Ashley, Artificial Intelligence and Legal Analytics: New Tools for Law Practice in the Digital Age (2017).

[5] M.J. Sergot et al., *The British Nationality Act As a Logic Problem*, 29 Communications of the ACM 370-86 (1986).

[6] Dag W. Schartum, *Law and Algorithms in the Public Domain*, Nord. J. Appl. Ethics 15, 19 (2016).

141

hampered by an absence of digital platforms and financial investment. However, these constraints have loosened. In the private sector today, a US$20 billion legal technology market is fueling a range of software applications, automating to varying degrees many aspects of lawyering.[7] Advances in machine learning permit, for example, automated legal research in some fields, the drafting of text for new contracts, and the identification of documents for discovery requests.

In the public sector, automation is equally central in legal technology discourse. Government departments, international organizations, and judicial bodies are increasingly moving from mere digitization to experiments with automation.[8] This is accompanied by a growing research literature that pilots data-driven techniques to predict judicial and administrative decision-making, potentially paving the way for more ambitious future applications of artificial intelligence.[9]

Yet the enthusiasm is not shared by all. It is important to ask: what are the current and future implications for human rights, and should we be worried? Consider two examples.

The Digital Welfare State and the Right to Social Security

In late 2019, Philip Alston, the UN Special Rapporteur for Extreme Poverty and Human Rights, announced that the world was "stumbling zombie-like into a digital welfare dystopia."[10] Pushing back against the "cheerleaders" of digitalization and promises of improved access and transparency, Alston reported that, in partnership with the private sector, governments are digitalizing the welfare state to "automate, predict, identify, surveil."[11] He called for a "sober reflection on the downsides" of the transformation of social protection and assistance.[12]

The report outlined a range of concerns with the rise of automated eligibility assessments, calculation of benefits, fraud detection, and risk scoring. First, Alston pointed to the lack of accuracy. He catalogued numerous scandals, from 1,132 eligibility errors affecting US$101 million worth of payments in Ontario to the automatic issuing of half a million flawed debt notices to social security beneficiaries in Australia to the tune of US$0.85 billion.[13] Second, Alston highlighted that technologies overlook structural disadvantages based on inequality, poverty, and racism.[14] An individual's rights may be determined on the basis of predictions derived from the behavior of a general population group, which can be exacerbated by secret algorithmic processing, risk-scoring and need categorization.

Finally, he warned of ideological appropriation. The digital welfare state, unwittingly or not, provided a useful "neutral" cover for long-standing neoliberal policies that challenged the right to social security, whether by reducing welfare budgets, narrowing the beneficiary pool, or enhancing sanctions.[15] He described this digitalization as

---

[7] Bob Ambrogi, *The 20 Most Important Legal Technology Developments of 2018*, LawSites (Dec. 26, 2018); Jonathan Marciano, LegalTech Hits $1 Billion Investment [in 2018] As Lawyers Belatedly Embrace Automation, Becoming Human (Dec. 3, 2018).

[8] Ashley Deeks, *High-Tech International Law*, 88 Geo Wash. L. Rev. (forthcoming 2020).

[9] Wolfgang Alschner et al., *The Data-Driven Future of International Economic Law*, 20 J. Int'l Econ. L. 217 (2017); Jens Frankenreiter & Michael Livermore, *Computational Methods in Legal Analysis*, Ann. Rev. L. & Soc. Sci. (forthcoming 2020).

[10] UN Press Release, World Stumbling Zombie-Like into a Digital Welfare Dystopia, Warns UN Human Rights Expert (Oct. 17, 2019).

[11] Report of the Special Rapporteur on Extreme Poverty and Human Rights, UN Doc. A/74/48037, para. 8 (Oct. 11, 2019) [hereinafter Alston].

[12] *Id.* at paras. 1, 78.

[13] *See* Terry Carney, *Robo-Debt Class Action Could Deliver Justice for Tens of Thousands of Australians Instead of Mere Hundreds*, The Conversation (Sept. 17, 2019).

[14] Alston, *supra* note 11, para. 28.

[15] UN Comm. on Econ., Soc. & Cultural Rights, General Comment No. 19 on the Right to Social Security, UN Doc. E/C.12/GC/19, para. 11 (Feb. 4, 2008) [hereinafter General Comment No. 19].

reversing the "traditional notion that the State should be accountable to the individual"[16] because digitalization makes the individual transparent to the state.[17]

Indeed, this latter point is an important recognition that technology possesses constitutive power. In 1978, Carolyn Miller argued that technology, not unlike law, creates its "own forms of consciousness," making us view it as "truer, or more transparent, or more objective than others."[18] Technology begins as an instrumental means and becomes an inevitable end,[19] reshaping how we see the world. It privileges "linear, incremental, causal forms of thought" in understanding social phenomena and legitimates "efficiency" narratives.[20]

Alston's critique is comprehensive but not new. While the early adoption of algorithmic governance in policing and security has garnered the most attention,[21] its arrival in the welfare state has not gone unnoticed.[22] Harlow and Rawlings worry that "the good governance triad of transparency, accountability and participation may be restricted, even reversed," especially through the loss of reason-giving and discretion;[23] Larkin argues that the absence of digital literacy can hinder access to social services;[24] Burton demonstrates that face-to-face and telephonic services may be more appropriate for serious and urgent cases;[25] and Tomlinson indicates how the digitalization of appeals may be transforming administrative process into formal adjudication.[26]

The Digital Rule of Law and Right to Fair Trial

Scholars have raised similar concerns about automation's effect on civil rights, such as the right to fair trial. Digitalization and automation are reshaping legal proceedings. A growing number of countries have digitized aspects of formal dispute resolution, with an increasing use of video, online portals, and e-documentary systems—a process only likely to be accelerated by COVID-19-related restrictions that have illustrated the contingency of physical proceedings. In the private sector, online dispute resolution (ODR) platforms have grown and increasingly attracted public interest. A public-private partnership in the Netherlands provided ODR for divorce and housing cases until 2017[27] and many predict that digital resolution of disputes will become common.[28]

---

[16] Alston, *supra* note 11, at para. 5.

[17] *Id.* at para. 3.

[18] Carolyn Miller, *Technology As a Form of Consciousness: A Study of Contemporary Ethos*, 29(4) COMM. STUDS. 228, 236 (1978).

[19] *Id.* at 230.

[20] *Id.* at 231-32.

[21] Rosamunde Van Brakel & Paul De Hert, *Policing, Surveillance and Law in a Pre-Crime Society: Understanding the Consequences of Technology Based Strategies*, 20(3) J. POLICE STUDS. 163 (2011).

[22] *See, e.g.*, Joanna Redden, *Democratic Governance in an Age of Datafication: Lessons from Mapping Government Discourses and Practices*, 2 BIG DATA & SOC'Y 1 (2018); Caroline Sheppard and John Raine, *Parking Adjudications: The Impact of New Technology* in Administrative Justice in the 21st Century (Michael Harris and Martin Partington eds, 1999).

[23] Carol Harlow & Richard Rawlings, *Proceduralism and Automation: Challenges to the Values of Administrative Law*, *in* THE FOUNDATIONS AND FUTURE OF PUBLIC LAW (IN HONOUR OF PAUL CRAIG) (Elizabeth Fisher et al. eds., 2019).

[24] Paul Larkin, *Universal Credit, "Positive Citizenship", and the Working Poor: Squaring the Eternal Circle?*, 81 MOD. L. REV. 114 (2018).

[25] Marie Burton, *Justice on the Line? A Comparison of Telephone and Face-to-Face Advice in Social Welfare Legal Aid*, 40 J. SOC. WELFARE & FAMILY L. 195 (2018).

[26] Joe Tomlinson, *The Policy and Politics of Building Tribunals for a Digital Age*, UK CONST. L. BLOG (July 21, 2017).

[27] Karalina Mania, *Online Dispute Resolution: The Future of Justice*, 1(1) INT'L COMP. JURISPRUDENCE 76-86 (2015); Marit Barendrect, *Rechtwijzer: Why Online Supported Dispute Resolution Is Hard to Implement*, HIIL (June 21, 2017).

[28] Ernest Ryder, *The Modernisation of Access to Justice in Times of Austerity* 24 (5th Annual Ryder Lecture, University of Bolton, 2016).

Further, many see prospects for automated judging. Using machine learning methods on past jurisprudence, researchers have been able to predict with increasing confidence outcomes in judgments.[29] In the United States, many courts use COMPAS, an apparent machine learning software, to predict recidivism when imposing criminal sentences.[30] In New Zealand, a computer-based prediction model helps handle claims and profile claimants under the country's accident compensation scheme.[31] Others eye the potential for digital-friendly legislation and institutional reform that would permit greater automated decision-making through "expert design." These court-centric developments are likely to be complemented by attempts by parties in litigation to gain an advantage through use of data-driven legal research and prediction.[32]

Rights-based concerns about automated judging are growing, and new civil society organizations such as Algorithm Justice League are on the rise. The critiques fall into four main categories and mirror many of the critiques of the digital welfare estate. First, there is the potential for arbitrariness and discrimination. For example, while the literature is divided, there is some evidence that the COMPAS algorithm discriminates against African-American defendants by using structural background data.[33] Second, there are concerns about legal accuracy. Many doubt that either expert design or machine learning can master the bespoke and complex nature of legal decision-making, and worry about the rush to simplify law to reduce this obstacle.[34] Third, there is a lack of transparency over algorithm-based methods. Litigants may be deprived of reasons in automated decision-making and the algorithms in some software may be inaccessible due to intellectual property restrictions. Fourth, there may be an increase in the justice divide inter partes. If some litigants are better able to game or predict automated decision-making, they may obtain an unfair advantage in legal systems already plagued by strong disparities among parties.

Grounded Technological Critique

In the move to automated legal decision-making, the critical reflex of human rights (and often of doctrinal) scholars is strong. There is a legitimate concern that, in a Gramscian manner, new digital hegemonies create new subalterns and threaten old ones.[35] However, in the spirit of critical empiricism, it is worth reflecting on *how* to critique the march of new digital technologies.[36]

The first question concerns how we *frame* technology in general and automation in particular. Many critics adopt an "artifact"-centric approach, which is focused on digital techniques and methods. Yet, in science and technological studies, technology is commonly understood as the complex assemblages of components and know-how

---

[29] Daniel Katz et al., *A General Approach for Predicting the Behavior of the Supreme Court of the United States*, 12(4) PLoS ONE (2017); Masha Medvedeva et al., *Judicial Decisions of the European Court of Human Rights: Looking into the Crystal Ball*, PROC. OF THE CONF. ON EMPIRICAL LEG. STUDS. IN EUROPE (2018).

[30] Ben Green, *"Fair" Risk Assessments: A Precarious Approach for Criminal Justice Reform* (5th Workshop on Fairness, Accountability, and Transparency in Machine Learning, 2018).

[31] Jin Yoshikawa, *Sharing the Costs of Artificial Intelligence: Universal No-Fault Social Insurance for Personal Injuries*, 21 VAND. J. ENT. & TECH. L. 1155 (2018-2019).

[32] Deeks, *supra* note 8, at pt. III.

[33] Green, *supra* note 30.

[34] Tania Sourdin, *Judge v Robot? Artificial Intelligence and Judicial Decision-Making*, 41(4) UNSW L.J. 1114 (2018).

[35] Esteve Morera, *Gramsci's Critical Modernity*, 12 RETHINKING MARXISM: A JOURNAL OF ECON., CULTURE & SOC'Y 16 (2000).

[36] Steve Redhead, *Toward a Theory of Critical Modernity: The Post-Architecture of Claude Parent and Paul Virilio*, 14 TOPIA: CANADIAN J. CULTURAL STUDS. 37 (2005).

that make up "technological systems."[37] Thus, frames such as the "digital" welfare state or "robo"-judging obscure as much as they enlighten.

Modern states have been long based on "systems" and various forms of "automation." Indeed, a distinctive aspect of the General Comment No. 19 on the Right to Social Security under the International Covenant on Economic, Social and Cultural Rights is that the first element of the right is access to a "system" to provide coverage for predetermined risks.[38] Likewise, the right to a fair trial in a court process is based on a complex combination of actors, processes, and rules. Yet, at the same time, these bureaucratic and judicial systems have long been tools for control, exclusion, informal punishment, and surveillance. As Duncan Kennedy has observed, it is this dark side of the welfare state that helped spur the turn to rights across the political spectrum from the 1970s.[39]

If we retain an artifact-centric conception of technology, we risk reifying and romanticizing the imaginary of a "human/e state." Public administration should be understood as a form of technology, a complex and hierarchical amalgam of rules, algorithms, institutions, and spaces—that can both liberate and repress. Humans in their physical, affective, and cognitive states are just one element of this system. It is thus essential that the emerging critiques of digital welfare states not succumb to an atemporal reflex, but rather be viewed in the longue durée. In this respect, Fleur Johns's approach, which highlights the fusion of old and new technologies—the "list-as-algorithm"—is helpful in identifying the transformation rather than arrival of new forms of technological power.[40] The same can be said for Alston's linking of neoliberal forms of governance and new control-oriented technologies. The key question for advocates concerned with international human rights law is thus to ask how digital technologies strengthen or relieve the long-standing abusive aspects of the governmental state.

The second question concerns *evidence*. What do we know of the ills of the digital and automated state? One perennial challenge in critical theory and human rights fact-finding is the preference for the anecdotal and qualitative. While various studies do support several of Alston's conclusions, we should guard against cherry-picking. For example, the eligibility errors in the Ontario software were clear in 2015 but it is difficult to find evidence of the same problem in later Auditor-General reports. Did the Ontario government fix the automation errors after an experimental phase?

This potential slippage goes to the heart of the debate on the uptake of automation technologies. By what metrics should we evaluate its upsides and downsides? We have long known that there is a "black box" in human decision-making: administrative and judicial cognition is inflected and shaped by implicit bias, racial animus, arbitrariness and custom, laziness, and error. This is sometimes lost in discussions of the dangers of the computational "black box," with its structural bias that is compounded by the determinism and atheorism of data-driven approaches and the bluntness and inflexibility of "expert design" programming. Discussions of automation and digitalization should be guided by a logic of minimizing danger, regardless of whether its origin is machine or human.

The third question is how to effectively *regulate* the dark sides of automated decision-making. Alston sets out a classic human rights approach, a Pareto-optimization logic in which no individual is made worse off through efficiency improvements. In the case of welfare, he argues that states must ensure that there is a legal basis for digital welfare reforms; promote digital literacy and non-digital access;[41] maintain eligibility fairness and human dignity in

---

[37] JAMES MCCLELLAN & JAMES DORN, SCIENCE AND TECHNOLOGY IN WORLD HISTORY (2015).

[38] General Comment No. 19, *supra* note 15, at para. 11.

[39] Duncan Kennedy, *Three Globalizations of Law and Legal Thought*, *in* THE NEW LAW AND ECONOMIC DEVELOPMENT 19, 61 (David Trubek & Alvaro Santos eds., 2006).

[40] Fleur Johns, *Global Governance Through the Pairing of List and Algorithm*, 34 ENVT. & PLANNING D: SOCIETY AND SPACE 126 (2017).

[41] Alston, *supra* note 11, at 44.

procedures; protect civil rights through privacy constraints and limits on use of data to harass and surveil; democratize policy-making on digitalization; and hold both public and private actors accountable. Similar demands are made elsewhere on the emerging automation of judging, although with a greater emphasis on accuracy.

Is this enough to address the dark sides of digitalization? In my view, holding the digital Leviathan to account will also require new digital tools. Indeed, the robo-debt scandal in Australia traversed by Alston was tackled by advocacy groups through a website for automating complaints, which also served as a platform for digital mobilization. The legal tech movement is slowly developing public interest technologies,[42] and applications like the new JustBot application help individuals in Europe apply more easily to the European Court of Human Rights and potentially avoid customary summary rejection.

In sum, the human rights community must not only ready itself to challenge digital developments, but also must develop digital weapons that match new forms of bureaucratic and judicial power. Yet despite the promise of a legal assistance revolution through technology, legal technology is rarely directed towards public interest, rights-enhancing projects.[43] Alston is right when he observes that the automation agenda is mostly one of costs savings and efficiency.[44] Public and private investment in digital accountability will be crucial therefore in ensuring that automation advances rather than retards international human rights.

---

[42] Shannon Farley, *Legal Tech is Opening the System to Those Who Need Legal Representation Most*, TECHCRUNCH (Mar. 13, 2018).

[43] Brian Sheppard, *Incomplete Innovation and the Premature Disruption of Legal Services*, 2015 MICH. ST. L. REV. 1797 (2015).

[44] *See also* JANNICK SCHOU & MORTEN HJELHOLT, DIGITALIZATION AND PUBLIC SECTOR TRANSFORMATIONS (2018).