Horizon Scanning Series

The Effective and Ethical Development of Artificial Intelligence: An Opportunity to Improve Our Wellbeing

AI Liability and Challenging Algorithmic Decisions

This input paper was prepared by Gary Lea

Suggested Citation


The views and opinions expressed in this report are those of the author and do not necessarily reflect the opinions of ACOLA.
AI LIABILITY AND CHALLENGING ALGORITHMIC DECISIONS:

GARY LEA, SCHOOL OF REGULATION AND GLOBAL GOVERNANCE (REGNET), AUSTRALIAN NATIONAL UNIVERSITY

Introduction

As the adoption of Artificial Intelligence (AI) expands and its sophistication grows, so have questions relating to the regulatory effects of the change. For this present report, three questions have been posed:

1. What is the civil liability regime in Australia for property damage or personal injury applicable to an AI system?
2. What is the criminal liability regime in Australia for an AI system in charge of a machine [i.e. a smart robotic system]?
3. How can algorithmic decisions be challenged?

In relation to the first two questions, the more AI systems make decisions for or about us and, in some cases, become embodied actors in our physical environment (e.g. autonomous vehicles), the greater the possibilities for harm to humans. This expansion creates corresponding greater potential for breaches over an expanding range of regulation but also greater uncertainty as to practicable applicability of principles and rules developed before the advent of the technology in question. The issue of modification and renewal of the relevant regulatory framework – including but not limited to civil and criminal liability – then arises, together with issues of the appropriate principles and methodologies to achieve such change (Petit 2017).

In the longer term, questions arises as to when, why and to what extent AI and smart robotic systems might be recognised as persons under the law, including assuming civil and criminal liability either with others or even alone; at present, under Australian law, individual humans are natural persons but some other entities are legal persons, either generally (e.g. a company registered under the Corporations Act 2001 (Cth.) is a separate legal entity from its shareholders that has the legal capacity and powers of an individual per s124 of the Act) or for more limited purposes (e.g. a partnership is deemed to be a person for the purposes of Part XIC of the Competition and Consumer Act 2010 (Cth.) (CCA) on telecommunications access arrangements. This issue is also considered further below.

The Tangled Web – Liability Cannot be Viewed in Isolation

Taking enhanced vehicle automation as a reference point, it is clear that liability is a small and dependent part of a much bigger regulatory picture: leaving aside liability (criminal or civil), even current vehicle regulation in Australian is a complex, decentred amalgam of rules, standards and norms, including but not limited to road rules, driver licensing, vehicle type approval and insurance (Dent 2018). For convenience, vehicle automation levels can be classified in accordance with the Society of Automotive Engineers (SAE) table as follows with Level 0 representing most current vehicles:
Major car manufacturers project that their Level 4 automated vehicles will ready for mass production between 2019 and 2021 with Level 5 sometime between 2021 and 2028 (Walker 2018). However, at present, the combined effect of State/Territory road rules requiring a human in control, vehicle type approval and insurance rules is that, in the absence of special provisions for trials or other more permanent changes, only vehicles at Levels 0-2 might lawfully be used on Australian roads (National Transport Commission 2016, pp32-33). Thus, Tesla vehicles in Australia currently operate at Level 2 while specific statutory provisions for public road trial schemes covering Level 3+ vehicles had been introduced in SA, Vic and NSW by February 2018 (Maschmedt and Searle 2018) with alternative testing arrangements in place elsewhere. The changes necessary beyond the trial phase are discussed further below.

The Current Civil Liability Regime and Its Gaps

At present, the heads of civil liability potentially most relevant to personal injury, death or damage to property arising in connection with AI systems and smart robotic systems are:

(a) common law negligence as modified by the various State and Territory civil liability statutes (e.g. the Civil Liability Act 2002 (NSW)) (H1);
(b) statutory liability under other legislation operating in place of, alongside or beyond common law negligence (e.g. liability to relatives of a deceased person in respect of death under the Compensation to Relatives Act 1897 (NSW)) or the special liability regime applicable to motor vehicle injuries under the Motor Accident Injuries Act 2017 (NSW)) (H2);
(c) defective goods liability under Schedule 2, CCA (otherwise known as the Australian Consumer Law (ACL)), in particular ss138-141 (relating to loss or damage suffered by an injured individual, a person other than the injured individual, suffered by another person if other goods are destroyed or damaged, suffered by another person if land, buildings or fixtures are damaged respectively) a (H3);
(d) consumer guarantees under the ACL, in particular sections 54 relating to supply of goods and 60 relating to supply of services (H4); and
(e) implied terms under State/Territory statute (e.g. merchantable quality for goods sold by description or fitness for purpose disclosed by person acting in reliance per s19, Sale of Goods Act 1923 (NSW)) or common law (e.g. application of due care and skill in respect of services) (H5).

In relation to H1, establishing liability in negligence requires careful and, in the case of AI and smart robotic systems, potentially burdensome identification (Gerstner 1993) of a person or
persons owing a duty of care, the nature and scope of the duty/duties, breach of duty/duties, damage and causation. Focussing on motor vehicles, the practical difficulties of such identification even in simple factual scenarios with limited underlying technology issues has already driven changes under H1 and H2. Similarly, claims by injured persons in respect of loss or damage arising from vehicle defects increasingly more likely to be dealt with under H3 where applicable.

With the projected road legalisation of Level 3+ vehicles from 2020 onwards (National Transport Commission 2018, p68), recourse to actions in negligence may become rarer as other options are (or become) available under H2 or H3; in addition to issues of foreseeability (Lea 2015), identification of persons owing a duty of care will become significantly harder at each successive level of vehicle automation as technical errors leading to a collision could arise from any one or more of:

‘ … errors in the [system’s] function that could have been detected by the developer; an incorrect or inadequate knowledge base [or training set]; incorrect or inadequate documentation or warnings; not keeping the knowledge base [or training] up to date; the user supplying faulty input; the user relying unduly on the output; or using the [system] for an incorrect purpose.’ (Kingston 2018).

Turing to H3, under the existing defective goods provisions, both a vehicle and computer software are goods as defined (s2, ACL). Noting that firmware is treated as software in the defective goods context (*Ipstar Australia Pty Ltd v APS Satellite Pty Ltd* [2018] NSWCA 15), AI systems supplied as vehicle firmware are likely (but by no means certain) to be treated as goods. In relation to defective goods, the defendant is the actual or deemed manufacturer under s7, ACL (e.g. person holding out as manufacturer, importer where manufacturer not established in Australia, etc.) and possible actions regarding manufacturer's liability in respect of safety defects (defined at s9, ACL) set out at ss138, 139 and 141 as outlined above. However, note that H3 may not apply in the event of an accident involving a trial vehicle: defective goods have to have been supplied 'in trade or commerce' for the relevant sections to engage. Furthermore, without change, the applicability of H3 going forward may also be limited by s142; in particular, under the 'state of the art' defence, it is a defence that the defect could not have been discovered at the time the manufacturer supplied the goods because there was insufficient scientific or technical knowledge at that time, noting that knowledge is more than mere suspicion (*Merck Sharp & Dohme (Australia) Pty Ltd v Peterson* [2011] FCAFC 128).

Turning to H4, where a person supplies goods or services to a consumer as defined (s3, ACL) in the course of trade or commerce (again noting the issue of trial vehicles) then, among other requirements, the goods must be acceptable quality as defined (ss54-(2)-(3), ACL) including consideration of freedom from defects and safety, and the services must be rendered with due care and skill (s60, ACL). However, it must be emphasised that acceptable quality is not an absolute but what a consumer might reasonably expect; again, this will have impacts in relation to advanced driving systems. Note in this context the recent review of some of the consumer guarantees (Commonwealth Treasury 2018) and suggestions of stricter rules specifically covering defective vehicles (Corones 2018).

Finally turning to H5, although H4 has largely displaced the operation of these in respect of consumer transactions, they still apply in other areas (e.g. B2B transactions). One issue that arises here is the more technical and restrictive definition of goods still used in many State/Territory sale of goods legislation such that software is not treated as goods unless deemed merged with the goods (*Gammasonics Institute for Medical Research Pty Ltd v Conrad Medical Systems Pty Ltd* [2010] NSWSC 267), leaving AI systems in a grey area unless changes are made. It is also unclear the extent to which provision of cloud or other remote AI systems could or might be treated as services under this heading at present.
The Current Criminal Liability Regime and Its Gaps

At present, criminal law potentially applicable to AI and smart robotic systems is even more piecemeal and potentially subject-matter specific than civil law: by way of respective brief illustrations:

(a) although some States/Territories have statutory regimes against computer hacking and misuse modelled on the Commonwealth’s Cybercrime Act 2001 (Cth.), WA, Queensland and Tasmania retain separate regimes based on older approaches; and
(b) the civilian use of drones is subject to its own set of air rules under the Civil Aviation Safety Regulations 1998 (Cth.) which at present typically prohibit fully autonomous drone operation because of the drafting separation of ‘person’ from ‘unmanned aircraft’ and the requirement that the aircraft be operated within the person’s visual line of sight (r101.073).

For convenience and consistency, then, criminal liability in relation to a smart robotic system will be considered in relation to motor vehicles.

Under the Geneva Convention on Road Traffic 1949 to which Australia is party, motor vehicles must have a driver and drivers must be able to control their vehicles at all times. Although the Convention references ‘person’ (Article 4), the drafting of rules and reference to ‘himself’ and ‘he’ at Article 7 indicate a human individual. Under the resulting Australian model road rules and State/Territory implementations, the drafting separation of ‘person’ from ‘vehicle’ and the drafting assumption that such a person is human (e.g. left and right side of the vehicle defined from the driver’s perspective) currently rules out an AI or advanced robotic system as the driver (National Transport Commission, pp64-66). Up to Level 2, this is not problematic but, in effect, at each level from Level 3 inclusive upwards, the user/occupant of a vehicle becomes more like a passenger in a Level 0-2 vehicle than a driver.

Given the above, under the current road rules, excepting special statutory provision for vehicle trials, engagement of Level 3+ automation would be a criminal offence in so far as the (human) driver must have proper control of the vehicle whilst driving (e.g. r297, Road Rules 2014 (NSW)). Conversely, without road rule changes, since it is the (human) driver that is regulated, other road rules could hypothetically cease to operate if higher levels of automation were to be engaged: for example, a Level 4 or 5 vehicle executing a stop because of a machine-unidentifiable hazard on an intersection which a human could identify, with the user/occupant not being a driver and the vehicle not being a driver, there could potentially be no criminal liability for obstruction of that intersection (Tranter 2016).

Necessary Changes Moving Forward

In fact, the changes hinted at above are already starting to happen: the Vienna Convention on Road Traffic 1968 (the sister convention to the Geneva Convention noted above) has already been amended at Article 8 to allow for systems ‘influencing’ driving and further changes to the Convention framework are being discussed. At the domestic level, the National Transport Commission and Transport and Infrastructure Council are seeking national harmonised approach to regulate advanced driver systems (ADSs) by specifying responsible ‘advanced driver system entities’ (ADSEs) at Level 3+: this forms part of a projected ‘gapless’ system of driving laws, safety assurance, insurance and a permitted access to data regime for Cooperative-Intelligent Transport System (C-ITS) plus vehicle onboard data. For example, as part of gathering the safety benefits of Level 5 automation, it would not be a criminal offence to use a Level 5 vehicle whilst under the influence of alcohol or drugs but (a) interfering with the safe operation of such a vehicle and (b) improper use of emergency functions (e.g.
emergency stop, emergency beacons) would be (National Transport Commission 2018, pp58-60)

In relation to insurance, present State/Territory schemes for Compulsory Third Party insurance (CTP) could not work for Level 3+ vehicles going forward given the current drafting presumptions about the ‘driver’ and ‘vehicle’ in relation to vehicle registration; this is why for current Level 3+ trials in Australia, car manufacturers such as Volvo have been obliged to meet claims for death, personal injury and property damage via their own public liability insurance, product liability insurance or self-insurance (Wright 2018). For Level 3+ vehicles to be allowed on Australian roads beyond trials, there would have to be national harmonisation on such insurance issues as fault/no fault claiming, vehicle classification, registration conditions, extended cover issues (e.g. cyberattacks on vehicles) and so forth.

Turning next to changes under H3, H4 and H5, the following appear to be the minimum necessary even now in order to accommodate AI and smart robotic systems: clarifying application of the categories of goods and services (possibly including introduction of the separate third category of downloaded digital content as used under the Consumer Rights Act 2015 (UK) or similar) for all three, redefining acceptable quality under H4, restricting the scope of the ‘state of the art’ defence under H4, modernising implied terms as to goods under H5 to mirror H4 and, as part of clarifying application and modernisation, putting supply of services under State/Territory statute law under H5. In the longer term, as they are introduced, safety ratings and other classifications of AI and smart robotic systems are likely to at least indirectly influence the operation of these provisions e.g. what constitutes acceptable quality under H4.

Finally, returning to personhood, there is no inherent technical barrier to rendering an AI or smart robotic system a person under Australian law today: a statutory deeming provision could do it, even if simply to capture a system as a regulated entity for the purposes of that statute. A more elaborate and complete way, detailed by Solum (1992), would be to incorporate an AI or smart robotic system as a corporation sole and vest it with legal capacity and powers. However, at this stage, even full personhood proponents concede there may be little practical point (Dvorsky 2018) while, if done in the longer term, there are two redlines:

(a) AI and smart robotic personhood and accompanying rights must not be drafted or implemented in such a way as to derogate from human rights and human dignity; and
(b) unless and until AI and smart robotic systems (i) can appreciate and uphold civic rights and responsibilities and (ii) can be appropriately deterred, punished or rehabilitated for criminal law purposes, the individuals and existing legal entities that design, build, distribute and use them must be held completely responsible for them by analogy to rules on children or potentially dangerous animals (cf. Hallevy 2013).

In short, designers, manufacturers, distributor and users should never be allowed to evade liability by simply saying ‘the robot did it’.

**Challenging Algorithmic Decisions**

In relation to challenging algorithmic decisions, there are five basic issues that need to be addressed:

(a) communicating whether or not algorithmic decision-making (ADM) is being used (H6);
(b) identification of the person using ADM (H7);
(c) some basic level of explanation of the relevant algorithm(s) and their use (H8);
(d) having a clear and simple process for challenging decisions made (H9); and
(e) ensuring that materials are made available to support a challenge (whether by way of legal evidence or otherwise) (H10).
At the present time in Australia, in addition to issues of fairness under H8, H9 and H10 (see further below), there is frequently a significant lack of transparency under all headings. As elsewhere, this is partly attributable to the piecemeal way in which ADM has been and is being adopted by multiple actors across multiple sub-sectors. However, another factor is the extent to which Commonwealth, State and Territory legislatures have previously focussed narrowly on privacy; thus, while the Privacy Act 1988 (Cth.) and State/Territory counterparts legislation can contingently affect how ADM is conducted, there is not (yet) a direct equivalent to Article 22, General Data Protection Regulation 2016/679 (GDPR). Article 22 provides that, subject only to tightly-controlled exceptions, ADM, including profiling, is prohibited if ‘based solely on automated processes’ and producing ‘legal effects concerning [the relevant individual] or similarly significant affects [that individual]’. The exceptions relate to entry into or performance of contracts, EU or Member State laws with specified safeguards (including as a minimum the right to require human intervention, to comment and to contest a decision) and the explicit consent of the relevant individual.

In so far as H6 and H7 are concerned, even the public sector in Australia does not automatically act transparently e.g. leaving aside sensitive areas such as defence and national security, there no list of Commonwealth departments or agencies using ADM (Elvery 2017). Turning to H8, while security, privacy or commercial-in-confidence considerations are often cited as a reason not to explain the workings and use of the relevant algorithm(s), some basic level of explanation is necessary for a challenge to be seriously contemplated. Although even the GDPR does not, as is often asserted, contain an express ‘right to explanation’ (Wachter et al. 2017), to the extent that ADM involves processing of personal information, that processing must be carried out transparently (Article 5(1)(a), GDPR). A related issue here is algorithmic fairness, including but not limited to bias or error in inputs, design of processing and selection of outputs: reflecting their designers and builders, ADM systems can – albeit unintentionally – discriminate such that developing and implementing standards of fairness need to be prioritised (Zou and Schiebinger 2018). Thus, while the Commonwealth Better Practice Guide 2007 on ADM expressly deals with accuracy, accountability and transparency, fairness is currently simply assumed.

Turning to H9, at present, the processes for challenging vary significantly depending on the sector and regulatory framework the particular ADM user operates in. Thus, although there are potential difficulties in relation to delegation and decision maker identification (Hogan-Doran 2017, p7), a decision made by a computer program under section 495A, Migration Act 1958 (Cth.) is deemed to be the Minister’s and, accordingly, is legally and administratively challengeable in the same way that the Minister’s decision would be. In contrast, unless provided for under a particular agreement, policy or Code of Conduct, or contingently via a contractual or specific statutory requirement, there may not necessarily be a process to challenge an ADM decision of a private sector entity; even if so, it may be very limited. Looking to the future, consolidating challenge pathways, creating standardised challenge processes and providing clear, simple information about those processes should be a priority.

Turning finally to H10, the making available of material to challenge is potentially difficult in that some kinds of AI that might be used in an ADM render both explanations and provision of material inherently problematic e.g. although algorithmic at base, deep learning artificial neural networks (ANNs) are currently a ‘black box’ in terms of the structure of the function those networks carry out. However, technical solutions to these problems are being progressively identified (Castelvecchi 2016) and so no relief from obligations to explain or provide materials should be given – simply do not use for ADM until the solutions are available. Other objections to making material available based on time, cost and even in some cases sensitivity are far less compelling again: as a matter of fairness, an individual subject to ADM cannot be reasonably expected to carry additional burdens merely to benefit the ADM user.
REFERENCES

Castelvecchi D ‘Can We Open the Black Box of AI?’ (Nature, October 5, 2016) accessed 21 July 2018 at <https://www.nature.com/news/can-we-open-the-black-box-of-ai-1.20731>


Tranter K (2016) ‘The Challenges of Autonomous Motor Vehicles to Queensland Road and Criminal Laws’ 16(2) Queensland University of Technology Law Review 59


