



HAL
open science

The Laws and Regulation of AI and Autonomous Systems

Anthony Wong

► **To cite this version:**

Anthony Wong. The Laws and Regulation of AI and Autonomous Systems. Leon Strous; Roger Johnson; David Alan Grier; Doron Swade. Unimagined Futures – ICT Opportunities and Challenges :, AICT-555, Springer International Publishing, pp.38-54, 2020, IFIP Advances in Information and Communication Technology, 978-3-030-64245-7. 10.1007/978-3-030-64246-4_4 . hal-03194304

HAL Id: hal-03194304

<https://inria.hal.science/hal-03194304>

Submitted on 9 Apr 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

The Laws and Regulation of AI and autonomous systems

Anthony Wong^{1,2,3}

¹Managing Director, AGW Lawyers & Consultants, Sydney, Australia

²Vice-President, IFIP

³Past President, Australian Computer Society (ACS)

anthonywong@agwconsult.com

Abstract. Our regulatory systems have attempted to keep abreast of new technologies by recalibrating and adapting our regulatory frameworks to provide for new opportunities and risks, to confer rights and duties, safety and liability frameworks, and to ensure legal certainty for businesses. These adaptations have been reactive and sometimes piecemeal, often with artificial delineation on rights and responsibilities and with unintended flow-on consequences. Previously, technologies have been deployed more like tools, but as autonomy and self-learning capabilities increase, robots and intelligent AI systems will feel less and less like machines and tools. There is now a significant difference, because machine learning AI systems have the ability to learn, adapt their performances and ‘make decisions’ from data and ‘life experiences’. This chapter provides brief insights on some of the topical developments in our regulatory systems and the current debates on some of the risks and challenges from the use and actions of AI, autonomous and intelligent systems. [1]

Keywords: AI, Robots, Automation, Regulation, Law, Job Transition, Employment, Data Ownership, Data Portability, Access, Control, Intellectual Property, Legal Personhood, Liability, Transparency, Explainability, Data Protection, Privacy.

1 Introduction

The base tenets of our regulatory systems were created long before the advances and confluence of new technologies including AI (artificial intelligence), IoT (Internet of Things), blockchain, cloud and others. With the rise of these new technologies we have taken many initiatives to address their consequences by recalibrating and adapting our regulatory frameworks to provide for new opportunities and risks, to confer rights and duties, safety and liability frameworks, and ensure legal certainty for business.

Sector-specific regulation has also been adopted and adapted to address market failures and risks in critical and regulated domains. These changes have often been reactive and piecemeal, with artificial delineation of rights and responsibilities. There have been many unintended consequences. More recently we have begun to learn from past mis-

haps, and these regulatory adaptations are now more likely to be drafted in technologically neutral way avoiding strict technical definition, especially when the field is still evolving rapidly.

AI and algorithmic decision-making will over time bring significant benefits to many areas of human endeavour. The proliferation of AI systems imbued with increasingly complex mathematical and data modelling, and machine learning algorithms, are being integrated in virtually every sector of the economy and society, to support and in many cases undertake more autonomous decisions and actions.

How much autonomy should AI and robots have to make decisions on our behalf and about us in our life, work and play? How do we ensure they can be trusted, and that they are transparent, reliable, accountable and well designed?

Previously, technologies have often been deployed more like tools, as a pen or paintbrush, but as autonomy and self-learning capabilities increase, robots and intelligent AI systems feel less and less like machines or tools. AI will equip robots and systems with the ability to learn using machine-learning algorithms. They will have the ability to interact and work alongside us or to augment our work. They will increasingly be able to take over functions and roles and, perhaps more significantly, the ability to make decisions.

When I reviewed AI ethical frameworks in 2019, there were more than 70 in existence. The number continues to grow. In 2019, jurisdictions including Australia [2] and the EU [3] published their frameworks, adding to the lists of contributors including the OECD Principles on Artificial Intelligence [4], the World Economic Forum AI Governance: A Holistic Approach To Implement Ethics Into AI [5] and the Singapore Model AI Governance Framework [6]. The debates have matured significantly since then, beyond ethical principles to more detailed guidelines on how such principles can be operationalised in the design and implementation to minimise risks and negative outcomes. But the challenge has always been putting principles into practice.

Emerging technologies are rapidly transforming the regulatory landscape. They are providing timely opportunities for fresh approaches in the redesign of our regulatory systems to keep pace with technological changes, now and into the future. AI is currently advancing more rapidly than the process of regulatory recalibration. Unlike the past, there is now a significant difference—we must now take into consideration, machine learning AI systems that have the ability to learn, adapt their performances and ‘make decisions’ from data and ‘life experiences’.

This chapter provides brief insights on some of the topical developments in our regulatory systems and the current debates to address some of the challenges and risks from the use and actions of AI, autonomous and intelligent systems. [1]

2 Automation, Jobs and Employment law implications

Over the past few years we have been inundated with predictions that robots and automation will devastate the workplace, replacing many job functions within the next 10 to 15 years. We have already seen huge shifts in manufacturing, mining, agriculture, administration and logistics, where a wide range of manual and repetitive tasks have

been automated. More recently, cognitive tasks and data analyses are increasingly being performed by AI and machines.

Historically, new technologies have always affected the structure of the labour market, leading to a significant impact on employment, especially lower skilled and manual jobs. But now the pace and spread of autonomous and intelligent technologies are outperforming humans in many tasks and radically challenging the base tenets of our labour markets and laws. These developments have raised many questions.

Where are the policies, strategies and regulatory frameworks to transition workers in the jobs that will be the most transformed, or those that will disappear altogether due to automation, robotics and AI?

Our current labour and employment laws, such as sick leave, hours of work, tax, minimum wage and overtime pay requirements, were not designed for robots. What is the legal relationship of robots to human employees in the workplace? In relation to workplace safety— what liabilities should apply if a robot harms a human co-worker? Would the ‘employer’ of the robot be vicariously liable? What is the performance management and control plan for work previously undertaken by human employees working under a collective bargaining agreement, now performed or co-performed with AI or robots? How would data protection and privacy regulations apply to personal information collected and consumed by robots? Who would be responsible for cyber security and the criminal use of robots or AI?

Are there statutory protection and job security for humans displaced by automation and robots? Should we tax robot owners to pay for training for workers who are displaced by automation, or should there be a universal minimum basic income for people displaced? Should we have social plans, such as exist in Germany and France, if restructuring through automation disadvantages employees?

There are many divergent views on all these questions. All are being hotly debated. Governments, policy makers, institutions and employers all have important roles to play in the development of digital skills, in the monitoring of long-term job trends, and in the creation of policies to assist workers and organisations adapt to an automated future. If these issues are not addressed early and proactively, they may worsen the digital divide and increase inequalities between countries and people.

ICT professionals are also being impacted as smart algorithms and other autonomous technologies supplement software programming, data analysis and technical support roles. With AI and machine learning developing at an exponential rate, what does the future look like?

2.1 Case study - line between human and robo advisers in financial services

FinTech (financial technology) start-ups are emerging to challenge the roles of banks and traditional financial institutions. FinTechs are rapidly transforming and disrupting the marketplace by providing ‘robo-advice’ using highly sophisticated algorithms operating on mobile and web-based environments. The technology is called robotic process automation (RPA) and is becoming widespread in business, and particularly in financial institutions. Robo-advice or automated advice is the provision of automated

financial product advice using algorithms and technology and without the direct involvement of a human adviser [7].

Robo-advice and AI capabilities have the potential to increase competition and lower prices for consumers in the financial advice and financial services industries by radically reshaping the customer experience. They are designed, modelled and programmed by human actors. Often they operate behind the scenes 24/7 assisting the people who interact with consumers. There are considerable tasks and risks involved in writing algorithms to accurately portray the full offerings and complexity of financial products.

In 2017 Australia, after a number of scandals, introduced professional standards legislation for human financial advisers [8]. These regulations set higher competence and ethical standards, including requirements for relevant first or higher degrees, continuing professional development requirements and compliance with a code of ethics. The initiatives were introduced into a profession already under pressure from the robo environment.

Because robo-advice is designed, modelled and programmed by human actors, should these requirements also apply to robo-advice? Should regulators also hold ICT developers and providers of robots and autonomous systems to the same standards demanded from human financial advisers? What should be the background, skills and competencies of these designers and ICT developers?

Depending on the size and governance framework of an organisation, various players and actors could be involved in a collaborative venture in the development, deployment and lifecycle of AI systems. These might include the developer, the product manager, senior management, the service provider, the distributor and the person who uses the AI or autonomous system. Their domain expertise could be in computer science, or mathematics or statistics, or they might be an interdisciplinary group composed of financial advisers, economists, social scientists or lawyers.

In 2016 the Australian regulator laid down sectoral guidelines [9] for monitoring and testing algorithms deployed in robo-advice. The regulatory guidance requires businesses offering robo-advice to have people within the business who understand the “rationale, risk and rules” used by the algorithms and have the skills to review the resulting robo-advice. What should be the competencies and skills of the humans undertaking the role?

The EU General Data Protection Regulation (GDPR) [10] went further, by placing an explicit onus on the algorithmic provider to provide “meaningful information about the logic involved” [11]. In addition, GDPR provides an individual with explicit rights including the rights to obtain human intervention, to express their point of view and to contest the decision made solely by automated systems [12] that has legal or similarly significant impact. GDPR applies only when AI uses personal data within the scope of the legislation.

Revealing the logic behind an algorithm may potentially risk and disclose commercially sensitive information and trade secrets used by the AI model and on how the system works.

The deployment of robo-advice raises many new, interesting and challenging questions for regulators accustomed only to assessing and regulating human players and actors.

3 Do robots and AI dream of owning Intellectual Property?

AI and machine-learning systems have already developed to the point where they can write music, generate automated reports, create art or even display human traits such as curiosity and conduct experiments to self-learn and develop [13]. Humans excel in creativity, imagination, problem solving, collaboration, management, and leadership which, at least for now, are very far off for AI and automation.

Will AI eventually outpace human capability and creativity? This may happen, but there is no consensus on when. Whatever the case, we are seeing more examples of original works created not by humans, but by autonomous AI. Businesses are increasingly investing in new AI and robotics technologies, and in research and innovation to enhance competitiveness.

AI has introduced extra dimensions to the complexity of intellectual property (IP). Investors should tread with caution while questions remain about the ownership of works generated or supplemented by AI. Who owns intangible outputs which could be perceived as IP when they are generated by a robot or AI? Who owns the IP—the manufacturer, the developer or the programmer? Could ownership fall to the user who provided the data for the robot to create the output? Or alternatively, could the robot own its creations?

But what happens when inventions, source code, objects or other assets are created autonomously and are directed by non-human entities, as will increasingly be the case in the future? The distinction between human-generated works and AI-generated works is emerging to be a controversial topic.

Our current regulatory framework generally assumes that IP is created by natural persons. The UK [14], European [15] and US [16] patent offices, recently rejected patent applications in which an AI machine ‘DABUS’ was designated as the inventor.

Commentators have long distinguished between computer-assisted [17] and computer-generated works. In many countries, including Australia, the former category has created few copyright problems, but computer-generated works with little or no human involvement pose a challenge to copyright’s subsistence. Any works created by autonomous AI and robots will suffer serious hurdles in securing copyright protection. They might not have sufficient human authorial contribution for copyright to subsist. Given that technological research and progress are often driven by the promise of financial rewards, this uncertainty around IP ownership could be a disincentive for commercial entities to invest in AI development.

Some jurisdictions have implemented specific provisions to protect literary, dramatic, musical or artistic work which is computer-generated. [18] Section 178 of the UK Copyright Designs and Patents Act 1988 defines “computer-generated work” to mean work “generated by computer in circumstances such that there is no human author of the work”. The author is the person who undertook the arrangements necessary for the creation of the work [19].

The WIPO’s Second Session of the Conversation of Intellectual Property and Artificial Intelligence have disclosed the significance of the debate and that the “attribution of copyright to AI-generated works will go to the heart of the social purpose for which the copyright system exists” [20].

4 Data Fuels AI But Who Owns Data?

Data is at the centre of the operation of many AI machine learning models. Industrial and public data, as well as personal data, are important sources of input for the training and evaluation of AI machine learning models.

The deployment of advanced intelligent algorithmic software, in conjunction with the rapid declining cost of digital storage, is fuelling the assembly and combination of vast datasets (known as ‘Big Data’) for automated data processing and interrogation. These algorithmic programs are more cost effective and efficient than human readers and are being progressively deployed across all domains of our society. Their aims are to unlock and discover new forms of value, to connect previously unseen linkages, and provide insights to stimulate growth and innovation in the digital economy [21].

Economies have formed around data, irrespective of whether an adequate regulatory framework has been built around it. In their relentless technological development, the AI and Big Data phenomena have overtaken the slow march of our law and have embraced and encapsulated some of the facets of our concepts of property without giving due regard and serious thought to the implications of treating data as property. In an attempt to create order from a runaway phenomenon, should there be underlying policy reasons to accord some form of property rights in the context of Big Data, and if not, some ‘bundles of rights’? [22]

Property rights evolve and change to address the practical needs of a given epoch in our society. Those needs change as our values and norms evolve. There is abundant literature on the different senses in which the term ‘property’ has been used to encapsulate the move from the traditional notions of property, such as land and chattels, to the notion of property in intangibles, such as artistic works. We are embarking on yet another significant leap, this time regarding property or ‘property-like’ considerations in data.

It is difficult to define property with any precision as the “notions of property inevitably change to reflect their context” [23]. Property law deals with rights and if recognised under established heads of law are claims ‘good against the world’, often described as ‘rights to exclude others’ [24].

Protecting value and proprietary rights in data involves a balancing act between many vested interests, including the interests of the purported owner, the interests of the custodian, the interests of competing third parties, and the interests of the public to access and use data. The debate on data ownership rights, and the layered complexities and issues pertaining to the granting of property rights in data, has intensified as the use and control of data assets become more and more critical to our economy and our ability to innovate. This requires a balancing of the commercial, private and public interests in data, as well as data protection and privacy concerns.

Existing laws in relation to copyright, patent, confidential information and trade secret, and trademark all relate to and protect rights involving information.

As observed by Nimmer, “copyright law has become a primary source of property rights in information in the 1990s” [25]. But existing copyright law is an inadequate framework for the consideration of property rights in data, because it provides owners with only a limited property right in the expression of the information [26]. Copyright

law does not concern itself with the control or flow of ideas, facts or data per se. The data components contained in the copyrighted work may not be protected, no matter how valuable. Ideas and facts are generally regarded to be in the public domain [27].

The right to control use of information may also arise under patent or other laws. Patent protects the use of ideas or information contained in the patent, by restricting the practice of the invention for a period of time.

In Australia and elsewhere, the question of whether information can be properly characterised as property in the context of confidential information has been subjected to much academic and judicial commentary over the last half century [28]. But if the owner of the confidential information places it in the public domain and accessible for Big Data mining and analysis, the inherent 'secrecy' may be lost. In Australia, as in the United Kingdom, there is authority which supports the proposition that information is not property [29].

AI, Big Data and our society's dependence on the digital economy have emerged comparatively rapidly. This has heightened the debate on our ability and freedom to use and extract value from data without fear of prosecution as we try to gain insights into new discoveries, innovation and growth. Granting separate property rights to discrete collections of data (datasets) would create a substantial barrier to the evolution of Big Data and our ability to mine valuable information from these datasets.

In the world of Big Data these datasets can be created, collected and obtained (sometimes even verified) automatically, or as a by-product of another business function. Some will require the investment of time, capital and labour, while others may only require computer processing time. It will depend upon the types and forms of datasets, how they are derived, and the purpose they serve.

The different types and forms of Big Data will continue to challenge our thinking and concepts around the question of data ownership. They will also continue to create uncertainty about the boundaries of control and data ownership.

Rights in data come in many forms and from a variety of sources. For the most part, traditional intellectual property law has proven to be inadequate in providing protection. [30] These traditional intellectual property regimes do not provide adequate cover for data and information-based products. Indeed, these laws exclude most Big Data datasets (in whole or in part) from protection.

With the pervasive use of technology today, a rapidly growing percentage of our information is created automatically from the use of IoT devices, mobile and GPS devices, smart meters, systems collecting transactional data, and many other sources. Most of these sources generate factual information, so it is unlikely that they would be protected under our traditional intellectual property laws. Should rights be left to the realms of contract, confidential information, trade secrets, unfair competition laws and other mechanisms? Or should government provide the custodianship to enhance researchers' access to Big Data?

In 2006, the European Union adopted the Database Directive [31] in recognition of the fact that copyright is inadequate to protect the investment made by database owners. The database directive provides for two levels of protection:

- a) a sui generis database protection where a substantial investment has been undertaken (financial, technical or human) in “obtaining, verifying, or presenting the contents of the database” [32].
- b) in addition to that provided by copyright law, where by reason of the selection or arrangement of their contents constitute the author's own intellectual creation [33].

Article 1 of the directive defines a database as a “collection of independent works, data or other materials arranged in a systematic or methodical way and individually accessible by electronic or other means”.

In the USA, the tort of misappropriation allows owners some control over the use that can be made of their databases.

4.1 Is it about Data Portability, Access and Control?

In the era of AI, machine learning models, data portability and the right to control access to data are also relevant. The right to control another's access to information can involve several distinct bodies of law, including contract law, the law of confidential information and trade secrets, computer and cyber crime law, communications law, and various laws relating to privacy.

Recently we have seen examples of government intervention using the regulatory framework to regulate interest in data in the digital environment, without the requirement to establish ownership in the data held or restricted by an access control system associated with a function of the computer.

The Australian Consumer Data Right (CDR) regulations [34] give individuals and businesses greater control over their data, including the ability to access particular data in a usable form and to direct a business to securely transfer that data to a trusted third party. The consumer right will roll out across sectors of the economy, commencing in the banking sector from July 2020 followed by the energy and telecommunications sectors. The data regulatory framework also imposes significant additional privacy and data sharing obligations and penalties for breach.

In the EU, the Free Flow of Non-Personal Data Regulation [35] and the General Data Protection Regulation [36] allow users of data processing services to use the data gathered in different EU markets to improve their productivity and competitiveness. Both EU Regulations refer to data portability and aim to make it easier to port data from one IT environment to another one, to enable switching of service providers and to foster competition.

5 Legal personhoods for AI

Historically, our regulatory systems have granted rights and legal personhood to slaves, women, children, corporations and more recently to landscape and nature. Two of India's rivers, the Ganga and the Yamuna, have been granted legal status. In New Zealand

legislation was enacted to grant legal personhoods to the Whanganui river, Mount Taranaki and the Te Urewera protected area. Previously, corporations were the only non-human entities recognised by the law as legal persons.

“To be a legal person is to be the subject of rights and duties” [37]. Granting legal personality [38] to AI and robots will entail complex legal considerations and is not a simple case of equating them to corporations.

Who foots the bill when a robot or an intelligent AI system makes a mistake, causes an accident or damage, or becomes corrupted? The manufacturer, the developer, the person controlling it, or the robot itself? Or is it a matter of allocating and apportioning risk and liability?

As autonomic and self-learning capabilities increase, robots and intelligent AI systems will feel less and less like machines and tools. Self-learning capabilities for AI have added complexity to the equation. Will granting ‘electronic rights’ to robots assist with some of these questions? Will human actors use robots to shield themselves from liability or shift any potential liabilities from the developers to the robots? Or will the spectrum, allocation and apportionment of responsibility keep step with the evolution of self-learning robots and intelligent AI systems? Regulators around the world are wrestling with these questions.

The EU is leading the way on these issues. In 2017 the European Parliament, in an unprecedented show of support, adopted a resolution on Civil Law Rules on Robotics [39] by 396 votes to 123. One of its key recommendations was to call on the European Commission to explore, analyse and consider “a specific legal status for robots ... so that at least the most sophisticated autonomous robots could be established as having the status of electronic persons responsible for making good any damage they may cause, and possibly applying electronic personality to cases where robots make autonomous decisions” [40].

The EU resolution generated considerable debate and controversy, because it calls for sophisticated autonomous robots to be given specific legal status as electronic persons. The arguments from both sides are complex and require fundamental shifts in legal theory and reasoning.

In an open letter, experts in robotics and artificial intelligence have cautioned the European Commission that plans to grant robots legal status are inappropriate and “non-pragmatic” [41].

The European Group on Ethics in Science and New Technologies, in its Statement on Artificial Intelligence, Robotics and Autonomous Systems, advocated that the concept of legal personhood is the ability and willingness to take and attribute moral responsibility. “Moral responsibility is here construed in the broad sense in which it may refer to several aspects of human agency, e.g. causality, accountability (obligation to provide an account), liability (obligation to compensate damages), reactive attitudes such as praise and blame (appropriateness of a range of moral emotions), and duties associated with social roles. Moral responsibility, in whatever sense, cannot be allocated or shifted to ‘autonomous’ technology” [42].

In 2020, the EU Commission presented its “White Paper on Artificial Intelligence—A European approach to excellence and trust for regulation of artificial intelligence (AI)” [43] and a number of other documents including a “Report on the safety and

liability implications of Artificial Intelligence, the Internet of Things and robotics” [44] for comments. The White Paper is non-committal on the question of endowing robots with specific legal status as electronic persons. It proposes a risk-based approach to create an ‘ecosystem of trust’ as one of the key elements of a future regulatory framework for AI in Europe, so that the regulatory burden is not excessively prescriptive or disproportionate.

I concur with the conclusions reached by Bryson et al [45] that the case for electronic personhood is weak and the negatives outweigh the benefits—at least for the foreseeable future.

As evidenced by the historical debates on the status of slaves, women, corporations and, more recently landscape and nature, the question of granting legal personality to autonomous robots will not be resolved any time soon. There is no simple answer to the question of legal personhood, and one size will not fit all.

Should legal personhood for robots or autonomous systems eventuate in the future, any right invoked on behalf of robots, or obligation enforced against them, will require new approaches and significant recalibration of our regulatory systems. Legal personhood could potentially allow autonomous robots to own their creations, as well as being open to liability for problems or negative outcomes associated with their actions.

6 Responsibility and Liability for damages caused by AI

How should regulators manage the complexity and challenges arising from the design, development and deployment of robots and autonomous systems? What legal and social responsibilities should we give to algorithms shielded behind statistically data-derived ‘impartiality’? Who is liable when robots and AI get it wrong?

There is much debate as to who amongst the various players and actors across the design, development and deployment lifecycle of AI and autonomous systems should be responsible and liable to account for any damages that might be caused. Would autonomy and self-learning capabilities alter the chain of responsibility of the producer or developer as the “AI-driven or otherwise automated machine which, after consideration of certain data, has taken an autonomous decision and caused harm to a human’s life, health or property” [46]?

Or has “inserting a layer of inscrutable, unintuitive, and statistically-derived code in between a human decisionmaker and the consequences of that decision, AI disrupts our typical understanding of responsibility for choices gone wrong”? [47] Or should the producer or programmer foresee the potential loss or damage even when it may be difficult to anticipate—particularly in unusual circumstances, the actions of an autonomous system? These questions will become more critical as more and more autonomous decisions are made by AI systems.

One of the more advanced regulatory developments in AI is in the trialling of autonomous vehicles [48] and in the regulatory frameworks for drones [49].

The rapid adoption of AI and autonomous systems into more diverse areas of our lives—from business, education, healthcare and communication through to infrastructure, logistics, defence, entertainment and agriculture—means that any laws involving liability will need to consider a broad range of contexts and possibilities.

We are moving rapidly towards a world where autonomous and intelligent AI systems are connected and integrated in complex IoT environments in the mesh and “the plurality of actors involved can make it difficult to assess where a potential damage originates and which person is liable for it. Due to the complexity of these technologies, it can be very difficult for victims to identify the liable person and prove all necessary conditions for a successful claim, as required under national law” [50]. The burden of proof in a tort fault-based liability system in some countries could significantly increase the costs of litigation.

We will need to establish specific protections for potential victims of AI-related incidents to give consumers confidence that they will have legal recourse if something goes wrong.

One of the proposals being debated is for the creation of a mandatory insurance scheme to ensure that victims of incidents involving robots and intelligent AI systems have access to adequate compensation. This might be similar to the mandatory comprehensive insurance that owners need to purchase before being able to register a motor vehicle [51].

Another approach is for the creation of strict liability rules to compensate victims for potential harm caused by AI and autonomous systems along the lines of current product liability laws in the EU and Australia. Strict liability rules would ensure that the victim is compensated regardless of fault. But who amongst the various players and actors should be strictly liable?

Whether the existing mixture of fault-based and strict liability regimes are appropriate is also subject to much debate.

Introducing a robust regulatory framework with relevant input from industry, policymakers and government would create greater incentive for AI developers and manufacturers to reduce their exposure by building in additional safeguards to minimise the potential risks to humanity.

7 Transparency and Explainability of AI

Algorithms are increasingly being used to analyse information and define or predict outcomes with the aid of AI. These AI systems may be embedded in devices and systems and deployed across many industries and increasingly in critical domains, often without the knowledge and consent of the user. Should humans be informed that they are interacting with AI, on the purposes of the AI, and on the data used for the training and evaluation?

To ensure that AI based systems perform as intended, the quality, accuracy and relevance of data are essential. Any data bias, error or statistical distortion will be learned and amplified. In situations involving machine learning—where algorithms and deci-

sion rules are trained using data to recognize patterns and to learn to make future decisions based on these observations, regulators and consumers may not easily discern the properties of these algorithms. These algorithms are able to train systems to perform certain tasks at levels that may exceed human ability and raise many challenging questions including calls for greater algorithmic transparency to minimise the risk of bias, discrimination, unfairness, error and to protect consumer interests.

Over the last few years legislators have started to respond to the challenge. In the EU, Article 22 of the General Data Protection Regulation (GDPR) [52] gives individuals the right not to be subject to a decision based solely on automated decision-making (no human involvement in the decision process), except in certain situations including explicit consent and necessity for the performance of or entering into a contract. The GDPR applies only to automated decision-making involving personal data.

In the public sector, AI systems are increasingly being adopted by governments to improve and reform public service processes. In many situations, stakeholders and users of AI will expect reasons to be given for transparency and accountability of government decisions which are important elements for the proper functioning of public administration. It is currently unclear how our regulatory frameworks would adjust to providing a meaningful review by our courts of decisions undertaken by autonomous AI systems, or in what circumstances a sub-delegation by a nominated decision-maker to an autonomous AI systems would be lawful. We may need to develop new principles and standards and “to identify directions for thinking about how administrative law should respond ... that makes sense from both a legal and a technical point of view. [53].

As machine learning evolves, AI models [54] often become even more complex, to the point where it may be difficult to articulate and understand their inner workings—even to people who created them. This raises many questions: what types of explanation are suitable and useful to the audience? [55] How and why does the model perform the way it does? How comprehensive does the explanation need to be—is an understanding on how the algorithmic decision was reached required, or should the explanation be adapted in a manner which is useful to a non-technical audience?

In the EU, the GDPR explicitly provides a data subject with the following rights:

- a) rights to be provided and to access information about the automated decision-making; [56]
- b) rights to obtain human intervention and to contest the decision made solely by automated decision-making algorithm; [57] and
- c) places explicit onus on the algorithmic provider to provide “meaningful information about the logic involved” in algorithmic decision, the “significance” and the “envisaged consequences” of the algorithmic processing [58].

But how would these rights operate and be enforced in practice? With recent and more complex non-linear black-box AI models, it can be difficult to provide meaningful explanations, largely due to the statistical and probabilistic character of machine learning and the current limitations of some AI models—raising concerns including accountability, explainability, interpretability, transparency, and human control.

What expertise and competencies would be required from a data subject to take advantage of the rights or for the algorithmic provider to provide the above rights?

“In addition, access to the algorithm and the data could be impossible without the cooperation of the potentially liable party. In practice, victims may thus not be able to make a liability claim. In addition, it would be unclear, how to demonstrate the fault of an AI acting autonomously, or what would be considered the fault of a person relying on the use of AI” [59].

This opacity will also make it difficult to verify whether decisions made with the involvement of AI are fair and unbiased, whether there are possible breaches of laws, and whether they will hamper the effective access to the traditional evidence necessary to establish a successful liability action and to claim compensation.

Should organisations consider and ensure that specific types of explanation be provided for their proposed AI system to meet the requisite needs of the audience before starting the design process? Should the design and development methodologies adopted have the flexibility to embrace new tools and explanation frameworks, ensuring ongoing improvements in transparency and explainability in parallel with advancement in the state of the art of the technology throughout the lifecycle of the AI system?

While rapid development methodologies may have been adopted by the IT Industry, embedding transparency and explainability into AI system design requires more extensive planning and oversight, and requiring input and knowledge from a wider mix of multi-disciplinary skills and expertise.

New tools and better explanation frameworks need to be developed to instill the desired human values and to reconcile the current tensions and trade-off between accuracy, cost and explainability of AI models. Developing such tools and frameworks is far from trivial, warranting further research and funding.

8 Summary and Looking Beyond

This chapter raises some of the major topical regulatory issues and debates relating to job transition and employment law; data ownership, portability, access and control; legal status of AI and personhood; intellectual property ownership by AI; AI liability; transparency and meaningful AI explanation; and aspects of data protection and privacy.

In the wake of the 2020 “black lives matter” protests, a number of technology companies have announced limitations on plans to sell facial recognition technology. There have also been renewed calls for a moratorium on certain uses of facial recognition technology that has legal or significant effects on individuals until appropriate legal framework has been established [60].

The need to address AI and autonomous system challenges has increased in urgency as the adverse potential impact could be significant in specific critical domains. If not appropriately addressed, human trust will suffer, impacting on adoption and oversight and in some cases posing significant risks to humanity and societal values.

From this brief exploration, it is clear that the values and issues outlined in the chapter will benefit from much broader debate, research and consultation. There are no definitive answers to some of the questions raised—as for many, it is a matter of perspective. I trust that this chapter will embark you on your own journey as to what our future

regulatory systems should encapsulate. Different AI applications create and pose different benefits, risks and issues. The solutions that might be adopted in the days ahead, will potentially challenge our traditional beliefs and systems for years to come. We are facing a major paradigm shift which will require significant rethink of some of our long-established legal principles, as we must now take into consideration, machine learning AI systems that have the ability to learn, adapt and ‘make decisions’ from data and ‘life experiences’.

ICT professionals understand better than most in relation to the trends and trajectories of technologies and their potential impact on the economic, safety and social constructs of the workplace and society. Is it incumbent on ICT professionals and professional societies to raise these issues and ensure they are widely debated, so that appropriate and intelligent decisions can be made for the changes, risks and challenges ahead? ICT professionals are well placed to address some of the risks and challenges during the design and lifecycle of AI-enabled systems. It would be beneficial to society for ICT professionals to assist government, legislators, regulators and policy formulators with their unique understanding of the strengths and limitations of the technology and its effects.

Historically, our regulatory adaptations have been conservative and patchworked in their ability to keep pace with technological changes. Perhaps the drastic disruptions that COVID-19 has caused in our work, life and play beyond the normal will provide sufficient impetus and tenacity to consider and re-think on how our laws and regulatory systems should recalibrate with AI and autonomous systems, now and into the future.

Acknowledgment

I would like to acknowledge and express my appreciation to Graeme Philipson for his editorial assistance. He is an ICT editor, writer and publisher, and author of ‘The Vision Splendid: The History of Australian Computing’. www.philipson.info

References

1. This chapter is for general reference purposes only. It does not constitute legal or professional advice. It is general comment only. Before making any decision or taking any action you should consult your legal or professional advisers to ascertain how the regulatory system applies to your particular circumstances in your jurisdiction
2. Australian AI Ethics Framework (2019). <https://www.industry.gov.au/data-and-publications/building-australias-artificial-intelligence-capability/ai-ethics-framework>, last accessed 2020/6/6
3. European Commission: Ethics guidelines for trustworthy AI (2019). <https://ec.europa.eu/digital-single-market/en/news/ethics-guidelines-trustworthy-ai>, last accessed 2020/6/6
4. OECD, OECD Principles on Artificial Intelligence (22 May 2019), <https://www.oecd.org/going-digital/ai/principles/>, last accessed 2020/6/20
5. World Economic Forum: AI Governance: A Holistic Approach to Implement Ethics into AI, <https://www.weforum.org/whitepapers/ai-governance-a-holistic-approach-to-implement-ethics-into-ai>, last accessed 2020/6/20

6. Singapore Model AI Governance Framework, <https://www.pdpc.gov.sg/-/media/files/pdpc/pdf-files/resource-for-organisation/ai/sgmodelaigovframework2.pdf>, last accessed 2020/6/20
7. Definition from the Australian Securities & Investments Commission: Regulatory Guide 255 - Providing digital financial product advice to retail client, <https://asic.gov.au/regulatory-resources/find-a-document/regulatory-guides/rg-255-providing-digital-financial-product-advice-to-retail-clients/>, last accessed 2020/6/6
8. Corporations Amendment (Professional Standards of Financial Advisers) Act 2017
9. Australian Securities & Investments Commission: Regulatory Guide 255 - Providing digital financial product advice to retail client, <https://asic.gov.au/regulatory-resources/find-a-document/regulatory-guides/rg-255-providing-digital-financial-product-advice-to-retail-clients/>, last accessed 2020/6/6
10. Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation), 2016 O.J. (L 119/1) [GDPR].
11. *Ibid* art. 15(1)(h)
12. *Ibid* art. 22(3)
13. This section is based on the article, Wong, Anthony,: Do robots and artificial intelligence think about copyright?. *The Australian*, September 5, 2017
14. UK Intellectual Property Office, refer patent decision BL O/741/19 of December 2019, https://www.ipo.gov.uk/p-challenge-decision-results/p-challenge-decision-results-bl?BL_Number=O/741/19, last accessed 2020/07/10
15. European Patent Office, refer decision of January 2020 <https://www.epo.org/news-issues/news/2020/20200128.html>, last accessed 2020/07/10
16. US Patent and Trademark Office, refer to decision of April 2020 on Application No. 16/524,350 https://www.uspto.gov/sites/default/files/documents/16524350_22apr2020.pdf, last accessed 2020/07/10
17. Here the computer is used as a tool equivalent of the painter's brush or the writer's pen by the author in the creation of the work
18. Similar provisions have been replicated in New Zealand, Ireland, India, Hong Kong and South Africa
19. Copyright Designs and Patents Act 2014 (UK) s 9(3)
20. World Intellectual Property Organisation (WIPO), Conversation of Intellectual Property and Artificial Intelligence, Revised Issues paper on Intellectual Property and Artificial Intelligence, May 2020, paragraph 23, https://www.wipo.int/meetings/en/doc_details.jsp?doc_id=499504, last accessed 2020/07/20
21. In recognition of the importance of the 'Digital Economy', the US President Obama requested a study to examine how the US can benefit from the data economy in January 2014. The report *Big Data: Seizing Opportunities, Preserving Values* concluded that data can be a driver for economic growth and innovation ('Big Data: Seizing Opportunities, Preserving Values')
22. For an overview on data ownership, refer to Wong, Anthony,: *Big Data Fuels Digital Disruption and Innovation, But Who Owns Data?* In: Chaikin, David., Coshott, Derwent. (eds.) *Digital Disruption Impact of Business Models, Regulation & Financial Crime* ch 2, Australian Scholarly Publishing, Australia (2017)
23. Beverley-Smith, Huw,: *The Commercial Appropriation of Personality*, p 296. Cambridge University Press (2002)
24. Merges, Robert P.: *Justifying Intellectual Property*, p 100. Harvard University Press (2011)

25. Nimmer, Raymond T.: Information Law, [2:8]. Thomson Reuters (May 2014)
26. The nature of the copyright in a literary, dramatic or musical work is defined in copyright legislation in the respective jurisdictions and in Australia, under the Copyright Act 1968 (Cth) s 31
27. Samuelson, Pamela.: Is Information Property?. In: Communications of the ACM (1991) 34(3), p 16
28. For an introduction to the protection of information using the law of confidential information, see Lahore, LexisNexis: Patents, Trade Marks & Related Rights (at 25 April 2016) [30,000]
29. See, eg, *Federal Commissioner of Taxation v United Aircraft Corp* (1943) 68 CLR 525 at 534; *Moorgate Tobacco Co Ltd v Philip Morris Ltd (No 2)* (1984) 156 CLR 414 at 438; *Breen v Williams* (1996) 186 CLR 71 at 81, 90, 111, 125; and *Australian Broadcasting Corporation v Lenah Game Meats Pty Ltd* (2001) 208 CLR 199 at 271
30. See Osenga, Kristen.: Information May Want to Be Free, But Information Products Do Not: Protecting and Facilitating Transactions in Information Products. *Cardozo Law Review* (2009) 30(5) 2099. p 2101
31. Directive 96/9/EC of the European Parliament and of the Council of 11 March 1996 on the legal protection of databases, OJ L 077, 27/03/1996
32. *Ibid* art 7
33. *Ibid* art 3
34. Treasury Laws Amendment (Consumer Data Right) Act 2019, <https://www.accc.gov.au/focus-areas/consumer-data-right-cdr-0>, last accessed 2020/6/2
35. Regulation (EU) 2018/1807 of the European Parliament and of the Council of 14 November 2018 on a framework for the free flow of non-personal data in the European Union, OJ L 303, 28.11.2018
36. Article 20 of the Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation)
37. Smith, B.: Legal personality. *Yale Law J* 37(3), 283-299 (1928), p 283
38. For a discussion on the concept and expression “legal personality” refer to Bryson, J. J., Diamantis, M. E., Grant, T.D.: Of, for, and by the people: the legal lacuna of synthetic persons. *Artificial Intelligence and Law*. 25(3) (2017), p. 277
39. European Parliament: European Parliament resolution of 16 February 2017 with recommendations to the Commission on Civil Law Rules on Robotics (2015/2103(INL)), <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52017IP0051>, last accessed 2020/6/9
40. *Ibid* paragraph 59(f)
41. Refer <http://www.robotics-openletter.eu/>, last accessed 2020/6/9
42. European Group on Ethics in Science and New Technologies: Statement on Artificial Intelligence, Robotics and ‘Autonomous’ Systems, p 10. European Commission, Brussels (2018), http://ec.europa.eu/research/ege/pdf/ege_ai_statement_2018.pdf, last accessed 2020/6/9
43. European Commission: White Paper on Artificial Intelligence - A European approach to excellence and trust, COM(2020) 65 (Feb. 19, 2020), https://ec.europa.eu/info/sites/info/files/commissionwhite-paper-artificial-intelligence-feb2020_en.pdf, last accessed 2020/6/9

44. European Commission: Report on the safety and liability implications of Artificial Intelligence, the Internet of Things and robotics, COM (2020) 64 (Feb. 19, 2020), https://ec.europa.eu/info/files/commission-report-safety-and-liability-implications-ai-internet-things-and-robotics_en, last accessed 2020/6/9
45. Bryson, J. J., Diamantis, M. E., Grant, T.D.: Of, for, and by the people: the legal lacuna of synthetic persons. *Artificial Intelligence and Law*. 25(3) (2017), pp. 273-291
46. The World Economic Forum; White Paper on AI Governance A Holistic Approach to Implement Ethics into AI, p. 6. Geneva, Switzerland (2019), <https://www.weforum.org/white-papers/ai-governance-a-holistic-approach-to-implement-ethics-into-ai>, last accessed 2020/6/9
47. Selbst, Andrew D.: Negligence and AI's Human Users. In: *Public Law & Legal Theory Research Paper No. 20-01*, p 1. UCLA School of Law (2018)
48. For a brief rundown of the regulatory frameworks and developments in selected countries refer to the Australian National Transport Commission 2020, Review of 'Guidelines for trials of automated vehicles in Australia': Discussion paper, NTC, Melbourne, pp. 16-18, <https://www.ntc.gov.au/sites/default/files/assets/files/NTC%20Discussion%20Paper%20-%20Review%20of%20guidelines%20for%20trials%20of%20automated%20vehicles%20in%20Australia.pdf>, last accessed 2020/6/6. For examples of Australian legislation refer to: Motor Vehicles (Trials of Automotive Technologies) Amendment Act 2016 (SA), Transport Legislation Amendment (Automated Vehicle Trials and Innovation) Act 2017 (NSW), Road Safety Amendment (Automated Vehicles) Act 2018 (Vic)
49. For the new European Union drone rules refer to: <https://www.easa.europa.eu/domains/civil-drones-rpas/drones-regulatory-framework-background>. For the Australia drone rules refer to: <https://www.casa.gov.au/knowyourdrone/drone-rules> and the Civil Aviation Safety Amendment (Remotely Piloted Aircraft and Model Aircraft—Registration and Accreditation) Regulations 2019
50. European Commission: Report on the safety and liability implications of Artificial Intelligence, the Internet of Things and robotics, COM (2020) 64 (Feb. 19, 2020), p 14, https://ec.europa.eu/info/files/commission-report-safety-and-liability-implications-ai-internet-thingsand-robotics_en, last accessed 2020/6/
51. Australian National Transport Commission 2020, Review of 'Guidelines for trials of automated vehicles in Australia': Discussion paper, NTC, Melbourne, pp. 26-27, <https://www.ntc.gov.au/sites/default/files/assets/files/NTC%20Discussion%20Paper%20-%20Review%20of%20guidelines%20for%20trials%20of%20automated%20vehicles%20in%20Australia.pdf>, last accessed 2020/6/6
52. General Data Protection Regulation (GDPR) art.22; Recital 71; see also Article 29 Data Protection Working Party, 2018a, Guidelines on Automated individual decision-making and Profiling for the purposes of Regulation 2016/679, 17/EN WP251rev.01, p.19. http://ec.europa.eu/newsroom/article29/item-detail.cfm?item_id=612053, last accessed 2020/6/4
53. Cobbe, Jennifer.: *Administrative Law and the Machines of Government: Judicial Review of Automated Public-Sector Decision-Making*. *Legal Studies*, p 3 (2019)
54. For the interpretability characteristics of various AI models, refer to ICO and Alan Turing Institute: *Guidance on explaining decisions made with AI* (2020), annexe 2, <https://ico.org.uk/media/for-organisations/guide-to-data-protection/key-data-protection-themes/explaining-decisions-made-with-artificial-intelligence-1-0.pdf>, last accessed 2020/6/6
55. For the types of explanation that an organisation may provide, refer to ICO and Alan Turing Institute: *Guidance on explaining decisions made with AI* (2020), p. 20, <https://ico.org.uk/media/for-organisations/guide-to-data-protection/key-data-protection->

themes/explaining-decisions-made-with-artificial-intelligence-1-0.pdf, last accessed 2020/6/6

56. General Data Protection Regulation (GDPR) art. 15
57. General Data Protection Regulation (GDPR) art. 22
58. General Data Protection Regulation (GDPR) arts. 13-14
59. European Commission: Report on the safety and liability implications of Artificial Intelligence, the Internet of Things and robotics, COM (2020) 64 (Feb. 19, 2020), p 15, https://ec.europa.eu/info/files/commission-report-safety-and-liability-implications-ai-internet-thingsand-robotics_en, last accessed 2020/6/9
60. Australian Human Rights Commission: Discussion Paper on Human Rights and Technology (2019), p 104, <https://humanrights.gov.au/our-work/rights-and-freedoms/publications/human-rights-and-technology-discussion-paper-2019>, last accessed 2020/6/20; For a US perspective, refer to Flicker, Kirsten: The Prison of Convenience, The Need for National Regulation of Biometric Technology in Sports Venues In: 30 Fordham Intell. Prop. Media & Ent.L.J. 985 (2020), p 1015, <https://ir.lawnet.fordham.edu/iplj/vol30/iss3/7/>, last accessed 2020/6/20