BOOK REVIEW

Cambridge Handbook of Artificial Intelligence. Global Perspectives on Law and Ethics
Larry A. DiMatteo, Cristina Poncibò and Michel Cannarsa (eds) (2022), 400pp., £145 hardback, Cambridge University Press, Cambridge, ISBN 9781009072168

It is a gargantuan task to compile a handbook on artificial intelligence (AI) that focuses on global perspectives of law and ethics. This is not only because new developments in AI are constantly reshaping understandings about what might be possible, but also because the approach to law and AI differs so significantly across jurisdictions. These jurisdictional differences have resulted in a wealth of ever-expanding literature and regulatory material that requires thoughtful analysis. Analysing this material, contemplating the 'known unknowns' and the 'unknown unknowns', becomes even more difficult when academics have differing approaches to AI evaluation. On the one hand, many, often with a European frame of reference, are very focused (if not obsessed) with regulatory measures, while others, often with a common law background, consider AI from the perspective of longstanding legal principles.

In addition, attempting to group thematically and link a range of diverse contributions is a Sisyphean task, particularly as concepts explored by the various contributors to the *Cambridge Handbook* not only overlap, but at times may be in opposition to one another. Despite these substantive issues, the *Cambridge Handbook* provides much food for thought, in part because of its focus on broad principles, presented symbiotically with specific examples. The editors are to be congratulated on this excellent work. This equilibrium of comprehensive and focused analysis will enable lawyers (and others) to craft thoughtful approaches to AI developments that are already present, and those which may emerge in the future.

In the various contributions to this extraordinary book, it is possible to discern the difficulty that many of the authors face when attempting to draw upon conventional legal theory and processes to accommodate the new AI world. In most chapters, there are admissions that, despite the emergence of a range of regulatory instruments and the proposed Artificial Intelligence Act (AI Act) in the EU, boundaries will need to be redrawn. Some issues will require the creation of new, AI-related norms that may be difficult to develop, manage and maintain. Overall, though, most authors consider that the approaches can be grafted onto existing approaches. Without considering whether new theoretical constructs will be required, this approach will be problematic as new technologies continue to resist codification and regulation.

Definitions of AI, autonomy and agency

There are several omissions in the book. This is partly because within the authors' focus on AI, there is a lack of clarity relating to the scope of the handbook and various technological developments are not explored. This lack of clarity is perhaps a result of the current uncertainty surrounding the definition of AI, the subject of discussion in the second chapter of the book. While the authors refer to the Turing test to determine what AI might include – that is, does the machine mimic human intelligence? – flaws in this approach arise in the context of narrow AI, and what is referred to as a 'broader AI field' definition:

> denoting the set of digital artifacts (hardware and software, possibly combined) that contains at least one learning or learned component, that is a component that can change its behaviour based on presented data and the patterns induced from that data. (p.25)

DOI:10.13169/prometheus.39.1.0060

By adopting a definition of AI that is linked to the notion of 'agents that are situated in an environment and interact with that environment while showing a certain degree of autonomy' (p.25), some chapter contributors appear to assume that the AI agent will be separate from a human. So, developments in neuro technology (as well as many other new technology areas) that raise a myriad of ethical and regulatory issues are not explored. Although complementary medical robotics are touched on (mostly from a liability perspective), there is very little discussion of human enhancements (external or internal) that may impact on the way we think and behave, and which may support a blending of intelligence.

In addition, new technology developments may be partly (but not always) driven by AI (as narrowly defined) and may further impact on decision making. As AI advancements are the major focus of the handbook, it is the autonomy issue with which a number of the contributors grapple. For example, the contract-law-focused authors spend much time examining the capacity of theoretical approaches to 'accommodate' the deployment of AI. However, they fail to consider that, because of advances in AI, many concepts in contract law may be completely overtaken by a complex blend of newer technologies.

At a basic level, if for example my virtual assistant acts independently based on my past preferences with your virtual assistant, can it really be suggested that the forms of AI are 'only' executing prior human instruction? In addition, as some authors concentrate on supportive rather than replacement AI (where some or many human tasks are completely undertaken by forms of AI), there is little focus in parts of the handbook on the legal capacity of semi or autonomous forms of AI that may undertake a range of tasks with little human input.

Although understandable, it is unfortunate that authors often start their chapters with a definition of AI, only to have them differ from each other. This is perhaps to be expected given the current definition inconsistency within the field. Even the EU struggles to define AI, with a range of definitions appearing in various regulatory instruments, exemplified by the continuously changing definition in the proposed AI Act. Nevertheless, in putting together a handbook of this nature, it might have been useful for the various chapter authors to agree upon a definition (perhaps a simple as well as another complex one) in addition to offering some thoughts about an excluder definition. This issue becomes particularly problematic when some contributors, such as Eric Tjong Tjin Tai, limit AI to a 'coded neural network' while others appear to include a range of expert and other algorithmic systems, with one contributor remarking that 'AI seems to be anything and everything'.

Thankfully, Pinar Caglayan Aksoy delves into this issue by considering AI as an agent. Aksoy concludes that 'foreseeable acts of an AI agent will be attributable to the operator: unforeseeable ones will not'. Aksoy also grapples with AI as a separate legal personality (having 'partial legal capacity') and notes that, given the current state of AI, this issue does not yet need to be determined. This is concluded in combination with commentary that appears throughout the book suggesting that, if an AI agent exceeds its authority, it may be necessary to consider revised compulsory insurance arrangements. Joshua Davis, in a chapter exploring ethics, attempts to distinguish between conscious and unconscious AI. In some ways, this approach assists. The dichotomy produces some philosophical conundrums with some very insightful commentary on free will. Once again, however, AI is described as 'other' and conversations about transhumanism might result in additional discussion and an extension of this work.

Florian Moslein, while steering clear of futuristic uncertainties, also attempts to tackle the issues that emerge in corporate law in the context of delegation. Moslein notes 'it is difficult to draw the line between merely advisory use of the technology, which is supposed to be always permissible and legally unproblematic, and improper delegation to AI with binding effect'. In a later chapter, there is also a discussion of AI legal personhood from an ethical perspective. Despite this personhood approach not garnering much support, there is a lot to recommend it as a conceptualization of AI.

In discussing tort law reform (and concluding that 'the core constructs of old language have remained malleable to changes brought by "new" technologies'), Robert Heverly essentially argues that existing tort theories are 'ready for AI'. In suggesting this, he concludes that 'given the general flexibility inherent in the tort system, and the existing methods of making incremental change,

allowing time for the tort system to respond to AI-related harms is the most prudent course'. One might ask whether we have time to enable this gradual catch-up and whether this approach can apply to all jurisdictions, particularly as different tort law cultures exist with a reduced focus on philosophical underpinnings in continental Europe. Jonas Knetsch ably argues that it is undoubtedly this factor that has led to a flurry of tort regime regulatory directives in Europe, while skilfully indicating the issues that emerge with various forms of AI that might resist clear rule-making in the tort area. In conclusion, 'the real issue at stake is that the rules governing the liability of manufacturers of AI products and algorithms, especially based on the EU product liability regime, are not fit for a widespread use of those technologies'.

It is not surprising that a discussion about autonomous vehicles and liability arises in some contributions to the handbook. Monot-Foultetier notes that this potentially disruptive technology requires a revision in insurance arrangements and the introduction of a compensation fund to meet the needs of any accident victims so there is no need 'to specifically determine which one is responsible'. However, Howells and Twigg-Flesner, when considering the internet of things (IoT) and arguing for the creation of a strict liability approach via a 'network liability' response, consider that there is still scope for a legal system response. It is conceded that this will require the redefinition of terms such as 'producer' and 'organiser' and a response to complex liability issues so that consumers (although not commercial entities) do not face an insurmountable evidentiary burden. I must say I found this proposal to be somewhat confusing as it reflects the existing European approach where commercial entities bear far greater evidentiary burdens (in this case, to deal with very complex liability issues) while consumers will bear no such burden. Is it realistic to suggest that all commercial entities should bear this burden? An unwanted side effect is no doubt the creation of ever-larger commercial entities as only very large entities will be able to manage the ever-increasing regulatory burdens and potential litigation costs that AI-related disputes will spawn.

Data protection and automated decision making

No handbook in this area could neglect data protection approaches and there is a particular focus on the EU General Data Protection Regulation (GDPR) in this one. As Spiker and Dohmann point out, the GDPR is not focused on AI, but on data. However, AI approaches are captured by the GDPR because the processing of data must be linked to a clear purpose and it is not legal to process data outside the purposes for which it was initially processed. This significant limitation in the EU setting is, of course, not replicated around the world. In my travels since the lockdowns lifted, I am always surprised by the EU's approach. It differs so significantly from approaches adopted in other jurisdictions. In contrast, a more holistic and strict regime operates within the EU and handbook contributors warn that AI developers and AI users should not neglect the bindings of the GDPR.

The other highlight of the GDPR is the prohibition against solely automated decision making unless it 'produces only positive consequences'. The interpretation of the GDPR provisions is the subject of regular EU court attention with the development of much case law in this area. This development was sadly not the subject of much commentary in this handbook. No doubt contributors have been hampered by the constant emergence of new interpretive material reflected in (sometimes inconsistent) case law since contributions were invited. This may point to the need for a second edition of this Handbook in the near future.

Europe vs the rest of the world

The intellectual property rights section of the work highlights some of the jurisdictional divide evident throughout much of the handbook. This divide involves the Europeans taking a regulatory approach (explored through regulation and some case analysis) and the Americans adopting much more of a wait and see attitude. Contributors from each jurisdiction concede that there are problems with both approaches and significant gaps, particularly when AI is an inventor or creator. In an

excellent summary of the arguments for and against AI inventor recognition, Christian Mammon takes the reader back to science fiction works before considering whether inventiveness and creativity are a core part of how we define ourselves as human. Mammon concludes that 'The prognostications reflect our collective ambivalence about an AI-fuelled future'. Without doubt, inventing the machine that invents things does raise the potential for harm. Mammon speculates:

> Maybe our future will be filled with benevolent, ethical cyborgs who are all too appreciative of the inclusivity of our legal system and will devote their resources to human improving innovation. Or maybe it will be much darker, populated by AI that has no interest in ethics, fairness, or humans' quality of life – as Stephen Hawking famously observed in 2014, 'Success in creating AI would be the biggest event in human history. Unfortunately, it might also be the last, unless we learn how to avoid the risks'.

The discussion about AI and copyright law is, of course, much more hopeful as it is primarily concerned with art, music and other creative endeavours. Gerald Spindler returns to a human delegation and agency view before concluding that, although close monitoring is required, it is not necessary to extend or modify copyright protection. This conclusion is reached after a discussion of whether AI is truly intelligent and whether AI in its present forms can 'determine the preferences or goals to achieve'. This is interesting as many would argue that AI can do this already and there is unsupervised AI now built that is having this impact.

The different perspectives on AI, with the Europeans spending much time detailing what is now an extensive regulatory regime, while others note that US laws are still sparse and that existing laws will apply, clearly highlight the issues confronting legal minds. Can AI be regulated effectively? What might this look like? And – perhaps the question that no one likes to ask – how much would it cost? Migle Lautkyte explores some of these issues in an engaging chapter called 'Keeping AI legal'. He considers how AI clashes with existing legal frameworks, with particular reference to intellectual property law and consumer protection law. His approach is basically that to 'keep AI legal means to keep it within the current legal framework and, if that is not feasible (as seems likely), then the ethical framework should be a temporal guide, but not a permanent alternative'. One suggestion he proffers – one that I agree with and have also written about – is to adopt a human-centred approach so that human wellbeing is prioritized whenever possible, and to think of AI and the law in the context of agreed ethical principles.

This approach is consistent with some other influential commentary that highlights the need for AI to be sustainable and for human wellbeing to be at the forefront of AI development. In addition, the following five principles that were settled on by the OECD to guide the use of AI require further examination and extension: inclusive growth, sustainable development and wellbeing; human-centred values and fairness; transparency and explainability; robustness, security and safety; and accountability. It might also be useful to include proportionality and complexity; that is, the need to consider how AI can be supported without layers of regulatory complexity that may be costly, unwieldy and inconsistent. This does not mean that regulatory approaches are not appropriate, but asks whether the many regulatory approaches to AI are beneficial. There is some interesting discussion in the handbook on whether technological reform (for example, through the use of emerging anti-bias tools) might sometimes address emergent AI problems.

In one chapter, John McGinnis attempts to reframe a cost-benefit analysis and speaks out for 'friendly AI'. In doing so, he notes that the adoption of a precautionary approach with AI regulation could stifle helpful innovation and 'create harm itself'. Discussing the benefits of AI during the pandemic and the everyday benefits of AI, McGinnis brings the reader back to a focus on benefit rather than risk (which is, of course, where most legal scholars stray). The thoughtful analysis by two of the book editors in the final chapter raises issues that emerge in the context of open discourse about regulatory arrangements. They conclude that 'ethics becomes more important than formal law in the future world of advanced AI'.

Ethics and the future

The ethical framework section of the handbook grapples with the proposed AI Act before exploring emotional AI, including the 'legal status of emotions, surveillance and interrogation, discrimination and decisional interference'. In considering right to psychological privacy and the gold standard of the GDPR, there is a fascinating, although very brief, discussion of the potential for AI to improve human wellbeing. The chapter by Mateja Durovic and Jonathon Watson suggests that, consistent with most discussions about AI, there is a potential dark side:

> The more concerning aspect is the extent to which the technology exploits or seeks to alter or take advantage of an individual's state of mind – emotions are after all 'potent, pervasive, predictable, sometimes harmful and sometimes beneficial drivers of decision making.' Emotional AI could therefore be applied to exploit or provoke particular emotions and 'nudge' the individual towards or subject them to detrimental decisions, for example by encouraging 'retail therapy' where inferences point towards sadness.

In noting that the proposed AI Act could include emotion recognition systems as high risk, there is certainly scope for more protection for consumers. But what might this capture? Arguably, every website and many forms of AI would be included. In attempting to limit the scope of work required, there is some suggestion that where power and information asymmetry is 'of such proportions that its use in some sectors ought to be prohibited'. In addition, there is reference to the proposed AI Act and 'material distortion of behaviour'. Again, this is problematic, in part because it requires consideration of vulnerable individuals and groups and exploration of how people may have been influenced (when the AI influence may be subtle or completely undetectable).

Again, much of this discussion highlights the need to consider these issues from an ethical perspective. A heavy-handed regulatory response is unlikely to attend to the many issues that will arise with new AI developments which might stifle 'friendly AI' in the justice sector. For example, research and support for the vulnerable in our society can be enabled through the development of sophisticated AI that can tailor communication approaches and even support positive behavioural change.

Impacts on lawyers

Much of the handbook explores legal and regulatory approaches to AI, although the foreword, written by Roger Brownswood, plots two disruptive phases that have ensured that lawyers pay attention to technological change. The first is linked to developments in biotechnologies and the second to developments in AI and machine learning. To these disruptive phases, I would add a third which is less about the content of the law and more about how legal practice takes place. This phase is partly linked to technological change. but is also linked to an expanded appreciation of the role and function of lawyers in modern society. Brownswood notes that:

> When lawyers had no interest in technology, thinking like a lawyer was inward-looking, doctrinal and guided by general principles. By the time lawyers saw technology as a challenge, thinking like a lawyer needed to become more 'regulatory', more policy-oriented, and more outward-looking (with a view to learning from economics, sociology and philosophy). (p.xxiii)

The way lawyers think about the world is necessarily changing because of the impact of newer technologies. There is a need to think more broadly. I further suggest that the way lawyers operate is changing and that lawyer ethics, professionalism and business models, in combination with justice processes, are all being fundamentally altered by newer technologies. If I have one criticism of this excellent book, it is that there is little focus on how AI and digitized systems will impact on lawyering, the development of dispute resolution mechanisms and the types of disputes that might

require legal attention. While there is a significant focus on content, this omission is significant as these are the human processes (including lawyer processes!) that will be impacted by AI.

While there are challenges in how we 'continue to write the rules to preserve social life and democratic institutions', one must conclude in reviewing this book that the AI revolution will create a significant amount of work for lawyers, regulators and others. The challenge is correctly summarized by McGinnis: we need to ensure that our response is not too precautionary, but is enabling these technologies, with such enormous potential, to benefit society. We must create common definitions that are comprehensive enough to incorporate developments, consider ethical and value-based approaches, and be cautious of a 'regulate everything' approach. An overly regulated approach will soon be out of date, disproportionate and ultimately not supportive of a future world so different from that of the past.

In summary, this is an outstanding book that provides much food for thought. The various contributions highlight the chasm that exists amongst legal scholars when considering AI. Some adopt a dystopian view with a focus on risk and regulation, while others are much more optimistic and encourage return to an ethical framework to underpin AI responses. It is difficult to see where the future of AI legal regulation lies. The proposed AI Act in Europe may be extended beyond the EU as a result of conventions between countries so it seems probable that much of the thoughtful regulatory material will be considered beyond the EU, perhaps supporting the low impact/high impact evaluation approach that is being fostered to reduce regulatory load. Some academics have recently suggested that some organizations will comply with EU regulations simply because it is easier to adopt a global approach (what is referred to as the 'Brussels effect'). It seems likely however that the significant jurisdictional and philosophical issues that have been highlighted above will mean that lawyers will continue to grapple with AI, and will continue to adopt a reactive rather than a proactive response. Some organizations will not be prepared to comply with EU approaches, particularly as many have already fostered separate systems with geo-blocking capacity.

Overall, the contributions in the handbook are certainly not a reflection of 'AI euphoria'. While some of the authors accept that AI will produce efficiencies and benefits, in most of the contributions AI is definitely not perceived as a panacea for the many problems facing our societies. As a consequence, AI continues to be perceived via a legal framing approach as a problem in need of hard governance and regulation. It is not seen as an opportunity to reframe how we interact with one another in consideration of innovative governance options in a world that is much more connected than ever before and where human-like 'intelligence' is not restricted to humans.

Tania Sourdin
Newcastle Law and Justice School, University of Newcastle, NSW, Australia
Tania.Sourdin@newcastle.edu.au