BOOK REVIEW

Advanced Introduction to Law and Artificial Intelligence by Woodrow Barfield and Ugo Pagallo (2020) Edward Elgar, Cheltenham, 208pp., hardback £85, ISBN 978 178990 5144.

This is a condensation of the same authors' *Research Handbook on the Law of Artificial Intelligence* (2018), which presents legal issues arising from the advent of artificial intelligence (AI) to a broad audience. Barfield and Pagallo wrote this abridged version to offer an easily accessible introduction to law and AI for law students, legal practitioners and non-legal experts.

AI is pervading business as well as our social and private lives – up to the point where we interact with some form of AI daily. These systems are becoming increasingly smart, and we must deal with the consequences of their incremental autonomy. Our legal systems are not yet equipped to handle these novel challenges and are in need of new and innovative regulatory strategies.

The authors start with a brief account of the fundamental concepts of AI to introduce lawyers to the subject matter. Subsequently, they explain relevant legal notions and legal situations in a comprehensible manner for both law students and non-lawyers. They present a concise overview of human rights frameworks and constitutional law, although their emphasis is on business-related law (data protection, intellectual property, antitrust and business law). In doing so, they analyse EU and US law comparatively, which provides valuable insights for law students into other legal systems as well as a felicitous approach to discussion of different regulatory strategies.

Key concepts of artificial intelligence

The authors' introduction to AI revolves around the core idea: the mimicking of human behaviour, and its implementation by using either rule-based (expert systems) or probabilistic (machine learning) approaches. This introduction will suffice for readers from technical disciplines and may provide a general idea to lawyers. However, more examples and illustrations of the general functioning as well as more cases might have helped readers without much knowledge of computer science. Additionally, a more detailed understanding of the process of machine learning and data science might have explained the features, idiosyncrasies, and risks associated with the use of the machine learning methods referenced in the subsequent chapters (see, for instance, the introductions to data science and machine learning in Surden, 2014, pp.89 ff. and Lehr and Ohm, 2017). The probabilistic nature of machine learning could have been contrasted more lucidly with rule-based expert systems to clarify that machine learning may expose statistical correlations without discerning causality (for a more in-depth discussion of correlation versus causality and counterfactual explanations, see Pearl 2019, especially pp.286 ff.). Given that the target audience consists partly of law students, a more thorough introduction would have set the stage for ensuing legal discussion. For example, legal redress against algorithmic decisions hinges on their transparency and explainability (see Doshi-Velez *et al.*, 2017, pp.7 ff.; Wachter *et al.,* 2018, pp.861 ff.). In contrast, the authors portray expert systems comprehensively in chapter 1, including issues of liability, professional misconduct and consumer protection. These concerns may also arise within the legal technology domain.

A vivid explanation of several AI systems and algorithmic decision-making systems follows with regard to discrimination (chapter 2) and patent law (chapter 9). This compensates for the initial short explanation by confining the discussion of specific technical properties to the relevant legal context. The variety of case law enriches the analysis of AI's impact on different legal fields. In chapter 3, with regard to free speech rights, the lack of examples obscures the nature of the AI

entities as speakers. Specific applications would have elucidated the abstract elaboration on freedom of speech. The quoted source on the matter refers mainly to such personal digital assistants as Siri (see Massaro and Norton, 2016). Another potential example in this context could be bots on social media (see Lamo and Calo, 2018). Furthermore, though the curating and pre-selection of information in the feeds of Facebook, Twitter, Google, etc. might not qualify as AI speech, it still impacts the right to freedom of expression (Sander, 2020). This deficiency of examples is only sporadic; for example, the next section (on searching and seizing evidence) mentions smart home gadgets, and the many court decisions in subsequent chapters offer instructive insights into practically relevant cases.

Legal status of AI

A central and reoccurring theme is the current and future status of AI within our legal system. Manifestly, the autonomy of AI impacts the distribution of roles in legal interactions: might AI entities be granted human or constitutional rights, be awarded tortious or criminal responsibility, or earn the role of inventor or creator?

In an abstract, conceptional discussion on legal personality (chapter 4), the authors introduce three main points of view: the believers are in favour of granting legal personhood to AI, either to avoid a scenario analogous to slavery or because of an (assumed) quality of the AI entity, e.g., awareness or the capability to suffer. The open-minded do not rule out a future legal personhood for AI, but based on rational choice and empirical evidence. Much like historic iconoclasts destroyed the icons and images of their gods, the AI iconoclasts reject the idea of awarding legal personhood to AI and thereby elevating it.

The authors themselves argue for a rigorous distinction of agency and personhood: while the notion of AI agency and accountability might be helpful in business contexts, full legal personhood would require that AI hold rights and duties of its own. As of now, the authors conclude, AI does not exhibit the level of autonomy, consciousness and intentionality needed to meet the requirements for legal personhood. In search of a benchmark for awarding legal personhood, the authors point to hearings on a (natural) person's legal capacity today. In the future, judges could determine the capacity of AI entities in a similar way, perhaps by means of the Turing test. Furthermore, mental states (e.g., intent) could be ascribed to an AI entity by taking the intentional stance (Sartor 2009, pp.260 ff.; Chopra and White, 2011, pp.12 ff.). In this context, an interdisciplinary reflection on agency and communication with AI systems would have complemented the discussion (e.g., Esposito, 2017; Gunkel, 2018). However, in the light of the short format of the book, the omission is understandable. In addition, AI can be a source of responsibility for other agents in the system. The current legal landscape might not be adequate to cope with the human–AI interaction and distributed responsibilities. Therefore, the authors call for regulation to address the impending loopholes and inconsistencies.

The legal status of AI is the book's most dominant red threat: chapter 6 transforms the theme of AI liability through AI personhood into a general discussion of allocating liability. The discussion on torts nearly seamlessly transitions into the question of whether AI could be criminally liable and is capable of being guilty (chapter 7). In chapter 5, the authors ponder on whether AI systems might qualify as data controllers under general data protection regulation (GDPR) since they collect and process data (however, the authors critically opine that the functional concept of a controller is intended to allocate responsibility to the person exercising the factual influence on the collection and processing of data). Furthermore, businesses conduct legal transactions with the use of automated contracting based on predictive analytics and pattern recognition, and AI systems are able to negotiate contracts or monitor compliance by analysing contractual relationships and identifying anomalies (chapter 10).

Encroaching on a domain generally thought to be dominated by fundamentally human qualities, AI may become a creator under copyright law (chapter 8) or an inventor under patent law

(chapter 9). The field of copyright law will certainly evolve within the next decades as natural language processing and generation of text will advance beyond the scope of OpenAI's GPT-3 (Romero, 2021). Additionally, the works-made-for-hire doctrine, paralleling the human–AI relationship with employers and employees, may provide an adequate legal framework for AI-generated creations. Besides that, an inventor AI system might be more skilled in a specific subject domain than most (or any) humans. Nevertheless, the idea of AI inventors is questionable because of the brute force, trial-and-error approach of many machine learning models.

Regulation of AI

Another recurring theme is the regulation of AI technology. This topic is usually associated with tort law and liability for AI systems (chapter 6). However, the authors draw on a wide array of regulatory domains and frameworks. Throughout the book, the authors present many examplary strategies: e.g., the sectorial introduction of data protection law (such as the health insurance portability and accountability act in the US) versus a broad regulation (such as the EU's GDPR). AI technology might be regulated by industry-specific law, e.g., autonomous driving or drones. Moreover, high-risk AI systems may be subjected to a pre-market review, much like medical devices and pharmaceuticals are approved by the US food and drug administration (FDA, briefly discussed in chapter 1). Pre-market review is a potential approach that has been repeatedly proposed with regard to regulating AI (see Tutt, 2017). Intriguing in this regard (from my point of view) is the review of changes and updates of software in medical devices by the FDA (chapter 1). Dynamic systems capable of learning and adapting need to be continuously reviewed to ensure their proper functioning and conformity with legal requirements. Often, their performance is continually improving and they keep learning. Besides that, AI systems are susceptible to cyberthreats, much like any other IT system. In the future, new regulation or adjusted existing frameworks, e.g., product liability (see chapter 6), must address both the intangibility of AI software and the need for continuous updating. Within the EU, two new directives will introduce a new framework for digital content and a duty to supply such updates; for examples, the content and digital services directive (EU 2019/770) and the consumer sales and guarantees directive (EU 2019/771).

Regulating AI is complicated by the infamous black box nature of AI, which the authors address throughout the book (and most prominently in chapters 3 and 6). The lack of transparency and explainability challenges the legal system in several ways. Potential harm and damage can hardly be foreseen, and thereby unavoidable. Moreover, the persons affected may encounter difficulties when pursuing legal redress. Apart from evidentiary problems due to proprietary software, it might be nearly impossible to disprove the inferences drawn by the AI system based on (accurate) data. Nevertheless, due process, a fair trial and legal redress against decisions must be guaranteed.

Another hot topic is data protection. With the emergence of ubiquitous computing and data-hungry machine learning models, people's privacy must be safeguarded effectively (chapter 5). Besides presenting the legal situation in the EU and the US, the authors analyse common problems and approaches, and detect a global legislative convergence in this domain. With new regulation in the US (the California consumer privacy act, with more states to follow) and the striking down of the EU-US privacy shield by the court of justice of the European Union, privacy regulation will remain a dynamic topic. Moreover, the book addresses some of the repercussions induced by the massive data collection and processing that has become endemic to our data-driven economy. The collective aspect of privacy is particularly interesting with regard to data mining, aggregation of data and ultimately profiling.

Legal protection of AI

Another main theme, resonating through several chapters (especially chapter 8–10), is the protection of AI systems by intellectual property law or as trade secrets. Barfield and Pagallo set out the purpose of intellectual property and its conflict of objectives in chapter 9. Intellectual property law

strives to optimize social welfare by guarding against under- and over-protection of information. It seeks to achieve a balance between the public and the rightsholders. After all, if there is no protection of innovative technology, companies and other actors may lack the incentive to innovate. Therefore, the protection of AI systems under intellectual property law should be evaluated based on these considerations.

Chapters 8 and 9 demonstrate the obstacles to protecting AI systems under copyright and patent law. In a logical flow, chapter 10 offers a remedy for the gap in protection caused by the specific standards in copyright and patent law. AI systems may be protected as trade secrets, so that the software's owner is entitled to refuse its disclosure. However, such software may be used to render decisions that affect people and their rights. Several of the cases discussed by Barfield and Pagallo pertain (at least partly) to the difficulties of successfully challenging algorithmic decisions rendered by proprietary software. A review of such software is either denied or carried out *in camera* to respect private companies' interest in non-disclosure of business secrets – even in criminal proceedings (chapter 7). A separate issue altogether is what kind of explanation humans need to understand algorithmic decisions and how best to elucidate the logic of an automated decision-making system (see Edwards and Veale, 2017, pp.64 ff.).

Future challenges

No book on AI would be complete without a little bit of futuristic science fiction and a salute to Nick Bostrom's *magnum opus Superintelligence* (2014). In their last chapter, the authors look ahead to the future when AI might surpass humans in (general) intelligence. In such a scenario, the most pressing legal issues may pertain to a super-intelligent entity's legal standing and the rights of humans (unless human laws are no longer effective). A list of previously unmentioned topics complements the preceding chapters.

In the end, AI triggers a plethora of legal problems that need to be addressed by legal systems. Convergence of regulatory approaches, international cooperation and collaboration may help to develop new strategies and solve many of the problems. Potential laws may be empirically tested in regulatory sandboxes. Nevertheless, many newly arising problems with AI have known loopholes or vulnerabilities within law and society – the perpetuation of social inequality, discrimination based on gender, sexual orientation, etc., government surveillance and consumer manipulation are a few of the most prominent. However, as the authors demonstrate, AI is capable of magnifying and intensifying these pre-existing problems to a new dimension.

Conclusions

Barfield and Pagallo's book offers a great overview on the most discussed and practically relevant legal discussions about AI. The abridged format has prevented the authors expanding on every issue of the current debate. However, they chose those topics which will be the most relevant for their readers: the fundamental question of legal status as well as constitutional and human rights, and the impact of AI on business. The authors portray the currently applicable laws and the relevant decisions comprehensibly for law students and non-lawyers. The references throughout the book as well as a list of additional topics will assist readers who would like to expand their knowledge. In many segments, further examples of AI would have helped comprehension of the real-world issues arising with AI. However, the many legal decisions do provide real-world examples and enrich abstract explanations. The book is too short and concise to accommodate much interdisciplinary scholarship. Instead, the authors limit their references to the most seminal works on AI and law. They present an overview and offer law students several carefully chosen gateways through which readers may explore the vast and steadily growing literature in the field. If you are looking for a concise book on the manifold issues of artificial intelligence and law, Barfield and Pagallo's *Law and Artificial Intelligence* is a great starting point.

References

Barfield, W. and Pagallo, U. (2018) *Research Handbook on the Law of Artificial Intelligence*, Edward Elgar, Cheltenham.

Bostrom, N. (2014) *Superintelligence: Paths, Dangers, Strategies*, Oxford University Press, Oxford.

Chopra, S. and White, L. (2011) *A Legal Theory for Autonomous Artificial Agents*, University of Michigan Press, Ann Arbor.

Doshi-Velez, F., Kortz, M., Budish, R., Bavitz, C., Gershman, S., O'Brien, D., Scott, K., Schieber, S., Waldo, J., Weinberger, D., Weller, A. and Wood, A. (2017) 'Accountability of AI under the law: the role of explanation', arXiv preprint arXiv:1711.01134.

Edwards, L. and Veale, M. (2017) 'Slave to the algorithm: why a right to an explanation is probably not the remedy you are looking for', *Duke Law and Technology Review,* 16, pp.18–84.

Esposito, E. (2017) 'Artificial communication? The production of contingency by algorithms', *Zeitschrift für Soziologie,* 46, pp.249–65.

Gunkel, D. (2018) *Robot Rights*, MIT Press, Cambridge MA.

Lamo, M. and Calo, R. (2019) 'Regulating bot speech', *University of California Los Angeles Law Review,* 66, pp.988–1028.

Lehr, D. and Ohm, P. (2017) 'Playing with the data: what legal scholars should learn about machine learning', *University of California Davis Law Review,* 51, pp.653–717.

Massaro, T. and Norton, H. (2016) 'Siri-ously? Free speech rights and artificial intelligence', *Northwestern University Law Review*, 110, pp.1169–94.

Pearl, J. (2019) *The Book of Why: The New Science of Cause and Effect*, Penguin Books, Harmondsworth.

Romero, A. (2021) 'OpenAI opens GPT-3 for everyone', *Medium*, 20 November, available at https://towardsdatascience.com/openai-opens-gpt-3-for-everyone-fb7fed309f6 (accessed November 2021).

Sander, B. (2020) 'Freedom of expression in the age of online platforms: the promise and pitfalls of a human rights-based approach to content moderation, *Fordham International Law Journal,* 43, pp.939–1006.

Sartor, G. (2009) 'Cognitive automata and the law electronic contracting and the intentionality of software agents', *Artificial Intelligence and Law,* 17, pp.253–90.

Surden, H. (2014) 'Machine learning and law', *Washington Law Review,* 89, 1, 86–115.

Tutt, A. (2017) 'An FDA for algorithms', *Administrative Law Review,* 69, 1, pp.83–123.

Wachter, S., Mittelstadt, B. and Russell, C. (2018) 'Counterfactual explanations without opening the black box: automated decisions and the GDPR', *Harvard Journal of Law and Technology,* 31, 2, pp.841–87.

Carolin Kemper
German Research Institute for Public Administration
67346 Speyer, Germany
kemper@foev-speyer.de