

FEATURE ARTICLE

What's Inside the Black Box? AI Challenges for Lawyers and Researchers

Abstract: The Artificial intelligence revolution is happening and is going to drastically re-shape legal research in both the private sector and academia. AI research tools present several advantages over traditional research methods. They allow for the analysis and review of large datasets ('Big Data') and can identify patterns that are imperceptible to human researchers. However, the wonders of AI legal research are not without perils. Because of their complexity, AI systems can escape the control and understanding of their operators and programmers. Therefore, especially when run by researchers with insufficient IT background, computational AI research may skew analyses or result in flawed research. Premised thus, the main goals of this paper, written by Ronald Yu and Gabriele Spina Ali, are to analyse some of the factors that can jeopardize the reliability of AI-assisted legal research and to review some of the solutions to mitigate this situation.

Keywords: artificial intelligence; AI; legal research

INTRODUCTION: THE AI REVOLUTION HAS BEGUN

Artificial Intelligence (AI), i.e. the ability of computers to exhibit human-like cognitive abilities, is already changing the transportation, financial and medical businesses.¹ According to some forecasts, computers will replace humans in one-third of traditional jobs by 2025.² The legal business will be no exception in this revolution and the legal profession has already begun to adopt AI technology over the past several years.^{3,4} Factors driving the adoption of AI include the challenge lawyers face to act more strategically and make better use of technology, especially when confronted with a sluggish growth in the demand of legal services and steep declines in productivity. This scenario drives firms to adopt AI as part of a "more for less" cost-saving solution.⁵

A number of legal AI systems have been introduced recently. These can be loosely grouped into three macro areas:

- **Unstructured data analysis and due diligence:** Where AI tools are employed to uncover background information and have so far had the most impact on the field. This includes contract analysis, document review and electronic discovery. In the field of contract review, some of these systems claim time and cost reductions up to 90% when compared to manual review.^{6,7} JP Morgan's Contract Intelligence (or COIN) is even said to need only a few minutes to

perform tasks that would take thousands of hours of human work.⁸

- **Legal research & analytics:** For instance, CaseMine provides more accurate legal researches and even allows unravelling covert legal relationships by mapping linkages between different cases.⁹ Other applications analyse past case law, win/loss rates and a judge's history for trends and patterns. For example, there is ongoing work to apply AI to analyse the rulings of both judicial bodies (the U.S. Supreme Court)¹⁰ and quasi-judicial bodies (the U.S. Patent Trial and Appeal Board).¹¹ China's Supreme People's Court started using an AI-enabled tool to search for precedents and identify analogous decisions to guide judges.¹² Some companies already offer similar systems to private law firms,¹³ while other focused on prediction technology that tries to anticipate litigation outcomes and opposing arguments.¹⁴ In the field of intellectual property, some AI tools help lawyers navigating large IP portfolios, warning of existing legal conflicts with prior IP and identifying potential brand name threats.¹⁵
- **Practice management applications:** Including electronic billing¹⁶ and document automation, i.e., programs helping lawyers in the drafting and redaction of legal documents and briefings.¹⁷ Other applications are meant helping companies with regulatory compliance in specific jurisdictions, e.g. China.¹⁸

THE PRESENT IMPACT OF AI IN THE LEGAL PROFESSION AND ITS IMPLICATIONS

The legal profession has only begun realizing the benefits of AI.¹⁹ Legal teams are finally embracing AI and some firms have even started their own in-house big data analysis/AI teams,²⁰ or even introduced AI-based products.²¹ However, we are still far from a generalized diffusion of AI technology among firms and researchers. Evidence from the US shows that both private practice and in-house lawyers have so far been reluctant to accept AI and that there has even been a low rate of adoption of new technology in general.^{22,23}

As for the next future, over half of in-house counsels believe the impact of automation will be “significant” or “very significant”, while only 3% believe automation will have no impact at all.²⁴ U.S. consulting group McKinsey estimated that 22% of a lawyer’s job and 35% of a law clerk’s job can be automated.²⁵ Similarly, 49% of the 386 US firms participating in Altman Weil’s *2017 Law Firms in Transition* reported to have created special projects and experiments to test innovative ideas or methods, and that they were using technology to replace human resources with the aim of improving efficiencies.²⁶

The greater adoption and accessibility of legal AI could bring numerous benefits to many areas of law and to the society in general. For example, a study of hundreds of summary judgment briefs in employment discrimination found that the vast majority of plaintiffs’ briefs omit available case law rebutting key defence arguments. Many of these briefs fell far below basic professional standards with incoherent writing or no meaningful research.²⁷

There are many ways in which AI can remedy the problem of sub-standard lawyering. AI allows for the analysis and review of large datasets (commonly referred to as “Big Data”) and is able to identify patterns that would be inevitably overlooked by a human observer. For instance, legal analytics software applications are able to process millions of court documents and can offer lawyers insights on potential litigation strategies and even simulate how a specific judge may respond to a given motion.²⁸ AI can also improve access to legal information. It can enable lay users to pose sophisticated legal inquiries and get plain answers from cheap and easily accessible AI systems. In this way, AI legal applications overcome past problems of accessing and mastering costly online services that were both incomplete in coverage and cumbersome even for experts to use, resulting in better justice for more people.²⁹

AI-SPECIFIC PROBLEMS

However, as more lawyers, law students and legal researchers embrace AI, they need also be aware of the potential dangers of placing blind faith in the impartiality, reliability and infallibility of legal AI.³⁰ As already noted in

a 1970 Stanford Law Review paper: “Lawyers might rely too heavily on a restricted, and thus somewhat incompetent, system with a resulting decline in the quality of legal services”.³¹ Following are some of the features that undermine the reliability and accuracy of AI in the legal profession and academia.

The myth of impartiality

AI systems are programmed using a set of algorithms,³² and ‘learn’ by studying data to identify patterns.³³ They are thus subject to both biases inherent in the algorithms employed - as different sets of engineers bring very different biases and assumptions to the creation of algorithms - and the data sets used. Different legal AI systems operate with different algorithms and, in many cases, on different datasets. Thus, despite claims of comprehensive and all-encompassing coverage, it is not surprising that different legal AI systems can produce different results.

This observation was reinforced by researchers who compared the results of the same legal search entered into the same jurisdictional case databases of Casetext, Fastcase, Google Scholar, Lexis Advance, Ravel, and Westlaw. The databases returned widely divergent results: an average of 40% of the cases was unique to one database, and only about 7% of the (same) cases were returned by all six databases.³⁴

Even law-makers are starting to acknowledge the implications of AI biases. For instance, the 2016 EU General Data Protection Regulation (GDPR) is among the first laws to recognize the effects of algorithmic decision-making on the “fundamental rights and freedom of natural persons”³⁵ and to address the issue of potential AI abuses.³⁶ Recital 71 of the Regulation even speaks of the implementation of “technical and organizational measures” that “prevent, inter alia, discriminatory effects on natural persons on the basis of racial or ethnic origin, political opinion, religion or beliefs, trade union membership, genetic or health status or sexual orientation, or that result in measures having such an effect”.³⁷ Under this perspective, the EU Data Protection Regulation seems to prohibit processing data on the basis of membership to special categories. Accordingly, companies operating in the EU will have to utilise algorithms that do not take into account characteristics such as gender, race or religion.³⁸

Algorithmic bias

As alluded to earlier, algorithms codify human choices about how decisions should be made and thus are not immune from the human values of their creators.³⁹ They can reinforce human prejudices as they are written and maintained by people and because machine learning algorithms adjust what they do based on people’s behaviour.⁴⁰

Problems of algorithmic bias were highlighted in two famous cases concerning image recognition AIs. It was

revealed that Hewlett-Packard's implementation of a feature-based face localization algorithm did not detect Black people as having a face.⁴¹ The algorithm measured the difference in intensity of contrast between the eyes and the upper cheek and nose of a human face and because of the choice of these parameters it did not work properly on darker faces in certain light conditions.⁴² Similarly, Google Photo's image recognition algorithm started tagging black people as gorillas. In the short term, Google was unable to fix the algorithm and solved the problem by removing words relating to monkeys from Google Photo's search engine.⁴³

Data bias

In addition to algorithmic biases, poor or flawed datasets used by AI systems are also a cause of concern. AI based on neural networks identifies recurring patterns on existing datasets and makes future predictions based on these patterns. However, there is a strong risk that AI may reiterate and even amplify the biases and flaws in datasets, even when these are unknown to humans.⁴⁴ In this sense, AI has a self-reinforcing nature, due to the fact that the machine's outputs will be used as data for future algorithmic operations.⁴⁵

In a well-known experiment, researchers found that employers were 50% more prone to shortlist job applicants with white-sounding names rather than those with African sounding ones.⁴⁶ The fear that AI recruiters would start amplifying human hidden biases became reality as an AI used by Amazon to shortlist job applicants started discriminating women candidates.⁴⁷ Similarly, there is a widespread concern that AI-powered banking software might start applying higher interest rates on racial grounds.⁴⁸ Also, researchers at the University of Virginia, who tested two large collections of labelled photos used to "train" image-recognition software, not only discovered that the images displayed a predictable gender bias in their depiction of activities such as cooking and sports (e.g. images of shopping and washing were linked to women while coaching and shooting were tied to men) but that machine learning software trained on the datasets amplified those biases.⁴⁹ Considered the foregoing, it should not surprise that experts have been warning that outsourcing decisions to AI may reinforce human prejudices rather than lead to more impartial, fair or transparent decisions.⁵⁰

The EU GDPR tries to solve the issue of data bias through the means of "data sanitization", i.e. by preventing the inclusion of variables relating to categories such as race, gender or sexual preferences.⁵¹ The data sanitization requirement can be interpreted as relating to explicitly discriminatory variables (e.g., skin colour) or to variables that have an implicit or statistical discriminatory incidence (e.g., height as a proxy to distinguish between men and women).⁵²

Nevertheless, data sanitization is a difficult application to AI systems. Under the first reading of the GDPR,

sanitization implies cleansing an algorithm from explicit discriminatory instructions.⁵³ This is rarely the case for AI systems, which take decisions based on recurring patterns in large databases, rather than because of explicit variables embed in their initial algorithm. By contrast, cleansing the data from all variables having even a potential discriminatory impact is infeasible because it would mean depriving the AI of the necessary operating information and impair its ability to reach accurate or altogether meaningful conclusions.⁵⁴

There are also issues of statistical bias in the management of data,⁵⁵ translating into outdated data, selection bias,⁵⁶ sampling bias, or misleading statistics.⁵⁷ Finally, there are collection or modelling errors. For instance, the problem of collection errors was highlighted in the 2016 American presidential election, when pollsters incorrectly predicted the results just days before the election, due to, *inter alia*, improper sampling, limitations in data collection or an electorate too complex to poll with any accuracy.⁵⁸ Institutional bias can also become a problem. For example crime reports, which are assumed to be random and representative, show significant geographical biases, i.e. if the police concentrate their patrols in certain areas there will be more data generated in those neighbourhoods.⁵⁹

Bad data practice could be caused by honest statistical or computing error (which may range from spreadsheet formulae, overflow or format conversion errors), misunderstanding of data and its applicability to the task on hand, misapplication of methods or failures to normalize data.⁶⁰ Finally, there is also the issue of malicious manipulation or corruption of data caused by cybercriminal or malicious hacking activity. For example, some speculated that data tampering was behind the problems Microsoft Tay experienced as the Tay had been subjected to "a coordinated attack by a subset of people" who "repeatedly had the bot riff on racist terms, horrific catch-phrases, etc".⁶¹ Similarly, researchers at the University of Southern California found that nearly 20% of the conversations surrounding the 2016 US presidential election on Twitter may have actually been created by bots that they speculate had been created by parties seeking to manipulate election results.⁶²

Inference and prediction

Algorithms are very poor at distinguishing between causation and correlation, and thus there is always a risk of conclusions based on wrong inferences. For example, the purchase of a hang-gliding magazine could be correlated with a risky lifestyle when the purchaser's true motive is an interest in photography.⁶³ Analysis models might also be flawed by incorrect assumptions, proxies, or presumptions of causal relationships where none actually exist (this is the problem of p-hacking).⁶⁴ Users may also attribute greater predictive capability to AI systems than is justified. Illusions of predictability may be caused partly

by users' lack of understanding of the technology and the systems themselves.

AI-powered predictive analytics and language translation systems commonly use statistical methods for dealing with unknowns and data limitations,⁶⁵ but these methods are not always followed by investigative processes that resolve the related unknowns and validate the system assumptions.⁶⁶ Users need not only to understand both that probabilities presented are not outright forecasts and that past historic behaviour does not necessarily predict future outcomes but that the models used may have limitations. For instance, a model attempting to 'predict' the voting behaviour of American Supreme Court Justices by examining overall past results will not provide a complete or accurate picture if it does not also consider long-term trends, e.g. that Supreme Court Justices have, on the whole, become more 'liberal' over time.⁶⁷

Input and output limitations

Legal AI systems must also be able to correctly interpret users' inputs. Though AI systems have made significant progress in understanding human language,⁶⁸ there are still significant morphological and semantic challenges to be overcome – especially where non-English languages are involved.

With respect to morphological challenges – a system needs to correctly understand what the query is. Where a foreign language is involved, the system must not only provide correct translations of individual words but also distinguish the meaning of compound words. For example, a legal AI may need to know that the Chinese word for 'Canada' is a combination of the characters for "to add", "to hold" and "big" (加拿大) that, when read sequentially, phonetically approximates the English word Canada⁶⁹ and that the specific grouping of 加拿大 needs to be translated as 'Canada' rather than as individual component characters.

The system must also be able to correctly perceive the relationship between words, as this is key to understanding the entire meaning of a law or regulation. For instance, while the Chinese sentence 網下和網上投資者獲得配售後，應當按時足額繳付認購資金⁷⁰ can be translated as: "After offline and online investors receive the placement, they should pay the subscribed-to funds on time and in full", a semantic mistake made by a machine could confuse dependency of the clauses and translate this as requiring the funds to pay the investors once placement has been completed.⁷¹

Unpredictability

AI systems are capable of surprising behaviours, sometimes due to external inputs, sometimes because of their own internal structure. Complex AI neural networks consist of several layers of electronic synopsis, which process and convert a given input into an output.⁷² They

learn by themselves via a trial-and-error process, similarly to what happens in biological brains.⁷³ Nowadays, AI can teach themselves how to perform complex tasks that only a couple of years ago were thought to require the unique intelligence - or deceptive capabilities - of humans.⁷⁴

Part of the problem is that developers do not really know how the algorithms used by such systems operate. Deep learning machines can self-reprogram to the point that even their programmers are unable to understand the internal logic behind AI decisions. In this context, it is difficult to detect hidden biases and to ascertain whether they are caused by a fault in the computer algorithm or by flawed datasets.⁷⁵ For this reason, neural networks are commonly depicted as a black box: closed systems that receive an input, produce an output and offer no clue why.⁷⁶

To provide an example, a research group at Mount Sinai Hospital in New York applied deep learning to the hospital's database of patient records, creating a program named Deep Patient. When tested on new records, Deep Patient proved proficient at predicting diseases. Without any expert instruction, it discovered patterns hidden in the hospital data that seemed to indicate when people were on the way to a wide range of ailments, including liver cancer and even schizophrenia. Its developers, however, had no idea how Deep Patient learned to do this.⁷⁷

For all these reasons, when systems capable of self-learning are exposed to external inputs the results can be unpredictable and even whimsical. In October 2014, a bot tasked to buy random items from the web bought 10 pills of ecstasy from the dark web. Interestingly enough, the Swiss police arrested the robot while releasing the programmers from any wrongdoing.⁷⁸

Microsoft's TAY chatbot had to be shut down soon after its release because it began tweeting offensive comments.⁷⁹ The chatbot behaviour was later ascribed to the widespread diffusion of racist comments online. Some even suggested that TAY's experience validated the Internet adage that as an online conversation increases in length, the greater the likelihood a comparison to Hitler will materialize.⁸⁰ Others simply raised the possibility of hacking.⁸¹

Google DeepMind's AlphaGo achieved historic success when it became the first computer program to defeat a world champion at the ancient game of GO.⁸² AlphaGo playstyle had initially been described by some as 'creative and intriguing'⁸³ causing evident unsettlement to his human opponent.⁸⁴ Yet that version of AlphaGo was subsequently defeated by a more advanced version (AlphaGo Zero) that learned to play Go without the constraint of human knowledge.⁸⁵ As a result, Go players worldwide started to re-evaluate centuries of inherited knowledge.⁸⁶ Some perceived AlphaGo's success as the definitive evidence of the creative capabilities of AI, especially considering that its playstyle was not anticipated by its developers.⁸⁷

Finally, there are also spill-overs risks to consider. AI systems may be employed proactively: for example to review data as cases come up and send cases, laws and regulations directly to a lawyer interested in a particular area; or be integrated with other systems or databases for, e.g., policy decision making.⁸⁸ Thus there is a risk that the consequences resulting from improper information from a legal AI system will cascade to other areas.

POTENTIAL SOLUTIONS

Fears of AI systems as black boxes, as well as concerns regarding potential implementations of AI technology in automatized weaponry and cyber attacks have sparked calls to ensure greater reliability, transparency and ethical alignment of AI.⁸⁹ There are several potential solutions to the challenges posed by AI to legal researchers. These are not mutually exclusive. The first four solutions proposed below operate upstream, i.e., they require institutional intervention from ruling bodies, universities or corporations. Conversely, the last set of solutions tries to empower researchers and practitioners with an array of best practices to remedy to the unpredictability, randomness and unaccountability of AI-driven legal research.

Education

Law schools have recognized the trend towards the use of AI and have begun creating new programs to teach the next generation of lawyers how to use these platforms and speak intelligently to the people building them. For example, the Harvard and MIT law schools nowadays jointly offer a course entitled 'The ethics and governance of artificial intelligence'. The course covers, among the others, topics such as algorithmic bias, risk assessment and predictive policies.⁹⁰ In 2016, The Georgetown University Law Centre in Washington even pioneered the elective course "Programming for lawyers" where student teams, under the supervision of legal services organizations, are asked to build an application to facilitate access to justice.⁹¹ Several Australian universities have followed this lead.⁹² For instance, in 2017, The University of New South Wales launched "Designing technology solutions for access to justice", an elective meant to teach law students with no IT background how to design legal information systems.⁹³

Courses of this kind are still at a pioneering stage and are offered by only a handful of universities worldwide. They also rarely cover AI-powered legal technology, focusing either on the implications of AI technology in the legal systems (e.g. liability and IP) or on traditional programming. In the future, it will be important to educate law students about the functioning and limitations of legal AI systems; with IT/AI law courses covering topics such as a taxonomy of the different types of legal AI/Big data systems, AI inference and predictive errors and input and output limitations.

Audit/ rating services

Experts have proposed the idea of an audit rating service to validate and certify the quality and accuracy of AI systems.⁹⁴ Audits have already been successfully applied to fields such as automated online advertisement, interest rates and pricing.⁹⁵ As noted earlier, use of statistical and other methods for dealing with unknowns and data limitations commonly employed in AI systems should be followed by investigative processes that resolve the related unknowns and validate the system assumptions, and should perhaps extend to the data and the models used.⁹⁶ In the field of personal data protection, the EU GDPR takes a step towards the diffusion and implementation of third-parties audits. The regulation favours the establishment of mechanisms such as certifications, seals and trademarks as long as they are granted through a process that is both transparent and subject to periodic review.⁹⁷

Establishing a certification system will not be an easy task, due also to the potential resistance of traders. Commercial providers could be reluctant to share information on their models or have their systems openly compared to their competitors. For instance, relevant litigation in the United States has shown that website terms of service prohibit most of the activities needed to conduct strict auditing to unveil discriminations on the Internet.⁹⁸ Also, as it will be shown later, it remains an open question whether there should be a legal requirement in any jurisdiction to disclose AI proprietary information to facilitate third-party reviews.

There are also pragmatic problems related to funding, insofar as there could be limited demand for such an auditing service. Corporations might also be updating their models so frequently as to make operating such a rating service impractical or too expensive. There will also be difficulties related to setting appropriate review standards, finding qualified reviewers and operating the scheme. These could prove a significant challenge given that they would require expertise in both legal and technical fields. A last problematic factor would be keeping the review criteria secret, with a view of preventing vendors or other parties from manipulating the results.

Algorithmic transparency

There also have been calls for greater algorithmic transparency, i.e., to oblige companies to release some mandatory information on their AI algorithms, in order to detect potential bias.⁹⁹ In the US, these claims are usually confronted with the observation that algorithms have proprietary nature and are protected under trade secret law.¹⁰⁰ In the European Union, a first step towards transparency has been taken with the aforementioned General Data Protection Regulation. This instrument provides European citizens with a right of explanation, i.e., to be informed about the reasons behind any algorithmic decision affecting them.¹⁰¹ Again, this measure aims at

avoiding automated decisions based on discriminatory parameters, such as race, gender or religion.¹⁰²

Others emphasize that transparency alone is inadequate to solve the problem of AI discrimination. Indeed, both the complexity of neural networks and the size of the datasets on which they are trained make AI internal logic inaccessible to human scrutiny.¹⁰³ Moreover, greater transparency neither solves the problems of data bias nor of the quality of the overall results returned by a legal AI.

Self Explanatory AI

As explained earlier, because AI learns from the surrounding environment and past mistakes, even programmers struggle to understand intelligent machines' internal logic and decision-making. Against this background, calls for greater understanding of AI inner procedures are ubiquitous.

At the institutional level, the US Department of Defence is currently working on what is defined as 'Explainable Artificial Intelligence' (XAI), running a project that aims to create a suite of machine learning techniques that: a) Produce more explainable models, while maintaining a high level of learning performance (prediction accuracy); and b) Enable human users to understand, appropriately trust, and effectively manage artificial intelligence outputs.¹⁰⁴

XAI would help AI meeting the requirements set by the EU GDPR, which, as anticipated, entitles data subjects with a right to explanation, i.e., the right to obtain meaningful information about the logic involved in automated decision making.¹⁰⁵ XAI would be especially useful considering that the GDPR prescribes that the relevant information has to be provided in a concise, transparent, intelligible and accessible manner.^{106, 107}

Private firms and research institutes are also working on teaching AI systems how to make themselves more understandable. For instance, a team at Microsoft is trying to teach AI to show how it weighted every single variable in evaluating mortality risk factors.¹⁰⁸ Similarly, a team at Rutgers University is working on a deep neural network that provides users with examples that demonstrates why it took a specific algorithmic decision.¹⁰⁹ Another project at the University of Berkeley involves lashing two neural networks together, tasking one to describe the inner procedures running inside the other.¹¹⁰ Finally, an international team consisting, among the others, of researchers from Facebook, Berkeley and the University of Amsterdam has taught an image recognition software to show the evidence he relied upon to reach its decisions.¹¹¹

Scholars and experts have already emphasized how we cannot blindly outsource moral decisions to machines. Under this perspective, understanding AI internal logic is a first step towards ensuring full accountability for computational legal research and automated legal decisions.

Best practices for legal researchers

Finally, as for more traditional research methods, there are some best practices that legal researchers should observe to avoid skewed or flawed results. These may be of particular importance while AI systems get more transparent and easier to understand by legal researchers.

Utilize multiple AI systems

As mentioned earlier, different legal databases are likely to return different results for the same query. Randomness may easily escalate in the field of AI. Indeed, neural networks learn from past failures, are able to self-modify their internal algorithm and are potentially trained on different datasets, and therefore each AI shows unique traits. Given that, legal researchers might need to compare outputs from different programs to detect flaws in the AI utilized and increase research accuracy.

Trying different inputs

AI systems are able to analyse and formulate replies to more complex questions than traditional systems based on Boolean logic. For instance, instead of returning an array of documents containing the words searched for, legal AI systems may be able to come up with precise and definite answers to questions such as what are the statutory exceptions to copyright infringement in a given jurisdiction.¹¹² Applications of this kind are often structured as natural language processing software, i.e. systems that work by calculating the probability that words may be found close to one other based on statistical inferences.¹¹³

For instance, the word "parody" is to be found close to words such as "exception" or "limitation" across EU copyright statutes and case law.¹¹⁴ Therefore, modern AI research software may be able to provide a list of results when asked in which European countries parody is an exception to copyright infringement. However, natural processing systems may fail in jurisdictions where the statistical dependency between the words "exception" and "parody" is weaker. For instance, the Italian copyright statute does not include parody among the potential exceptions to copyright infringement. Nevertheless, courts have recognized that parodies have to be protected as fully independent creations on the basis of some of the guiding principles of the Italian Constitution, such as freedom of speech and of artistic expression.¹¹⁵

Once again, the circumstance that legal researchers are unaware of the internal logic of AI systems requires that they exercise extra-care before trusting AI outputs. Like traditional research methods, running multiple queries using different keywords, relaxing time constraints or rephrasing the question asked to the software can help to detect biased, inaccurate or flawed research results.¹¹⁶

Human monitoring

As advocated by some scholars, ex-post human monitoring might be a tool to detect and correct cases of

algorithmic discrimination.¹¹⁷ In the world of online video games, gamers have spontaneously established committees to adjudicate, correct and sanction game violations.¹¹⁸ Combined with greater algorithmic transparency, crowd-level human monitoring might constitute a potential remedy to tackle skewed results and provide feedback to software engineers.¹¹⁹

In fields such as legal analytics and prediction technology, human intervention is also essential to avoid outsourcing moral responsibilities to machines.¹²⁰ Similarly, legal researchers cannot over-trust computers. Even though human's ability for empirical analysis is greatly surpassed by machines' computational ability, researchers should crosscheck computer's outputs with results obtained through traditional research methods on smaller data samples. This might be a way to ensure that AI research aligns with human values and to try to spot data or algorithmic biases.

CONCLUSIONS: THE ROAD TO THE FUTURE

The present paper is meant as a warning against over-trusting AI outputs and has provided an overview of the problems affecting AI computational research and

reasoning in the legal field. As AI becomes progressively more accepted among legal practitioners and academics, the duty falls upon researchers to keep up with the diffusion of AI to ensure that algorithmic decisions align to human values and that AI-driven research conforms to appropriate qualitative and ethical standards.

As anticipated, this is a first step in the exploration of the challenges posed by legal AI. Future research should analyse more narrowly-scoped problems and look for more in-depth insights, for which the close collaboration of legal professionals and IT experts will be necessary. Only through their joint efforts, will it be possible to implement new techniques and strategies to remedy AI biases, as well as to explore new ways of introducing greater AI transparency without compromising the intellectual property of AI vendors.

In the era of crypto currencies, self-executing contracts and robots showing superhuman intelligence, there is a widespread fear that computers might one day replace humans in the legal industry and in the management of the legal system as a whole.¹²¹ It is too early to know whether this day will eventually come, but what is certain is that until then it is up to humans to fix both the legal system and artificial intelligence.

Footnotes

- ¹ See Horst Eidenmüller (2017) 'The Rise of Robots and the Law of Humans', Oxford Legal Studies Research, Paper No. 27/2017, p. 3 [online]. Available at <https://ssrn.com/abstract=2941001> or <http://dx.doi.org/10.2139/ssrn.2941001> [Accessed 22 December 2017].
- ² Christoffer Hernaes (2015) 'Artificial Intelligence, Legal Responsibility and Civil Rights' *Techcrunch.com* [online]. Available at <https://techcrunch.com/2015/08/22/artificial-intelligence-legal-responsibility-and-civil-rights/> [Accessed 31 December 2018].
- ³ Ed Silverstein (2015) 'Lawyers Are Turning to Big Data Analysis. Volume: Complexity of Data Collections Create Challenges for Companies Facing Litigation', *The National Law Journal*, July 20, 2015 [online]. Available at <https://www.law.com/nationallawjournal/almID/1202732493683/Lawyers-Are-Turning-to-Big-Data-Analysis/> [Accessed 22 May 2018]; Gabrielle Orum Hernández, 'Data (Gold) Mining: The Rise of the Law Firm Data Analytics Teams' *Legal Tech News*, March 7, 2018 [online]. Available at <https://www.law.com/legaltechnews/2018/03/02/data-gold-mining-the-rise-of-the-law-firm-data-analytics-teams/> [Accessed 22 May 2018].
- ⁴ For purposes of this paper, the term 'artificial intelligence' will include logical AI/inferencing, machine learning, natural language processing, robotics, speech, vision, and neural network technologies as well as expertise automation, image recognition and classification, question Answering, Robotics, Speech, text analytics, text generation and translation functions. See, for example: 'Demystifying Artificial Intelligence (AI): A legal professional's 7-step Guide through the Noise' [online]. Available at <https://legalsolutions.thomsonreuters.com/law-products/artificial-intelligence/demystifying-ai?cid=7011B000002Kbtn&chl=eb> [Accessed 10 May 2018].
- ⁵ Edgar Alan Rayo (2017) 'AI in Law and Legal Practice – A Comprehensive View of 35 Current Applications' *TechEmergence*, November 29, 2017 [online]. Available at <https://www.techemergence.com/ai-in-law-legal-practice-current-applications/> [Accessed 4 May 2018].
- ⁶ Lawgeex, [online]. Available at <https://www.lawgeex.com> [Accessed 4 May 2018]; Ebrevia, [online]. Available at <https://ebrevia.com/#overview> [Accessed 4 May 2018].
- ⁷ Other AI applications in this field are Kyrasystems, [online]. Available at <https://kirasystems.com/benefits/> [Accessed 4 May 2018]. Levertion, [online]. Available at <https://www.levertion.ai> [Accessed 4 May 2018]. Lelgalrobot, [online]. Available at <https://www.legalrobot.com> [Accessed 4 May 2018]. Exterro, [online]. Available at <https://www.exterro.com/about/news-events/exterro-revolutionizes-e-discovery-market-with-robotic-e-discovery-fusion-whatsun/> [Accessed 4 May 2018]. Brainspace, [online]. Available at <https://www.brainspace.com> [Accessed 4 May 2018]. Everlaw, [online]. Available at <https://www.everlaw.com> [Accessed 4 May 2018].

- ⁸ Hugh Son (2017) 'JP Morgan software does in seconds what took lawyers 360,000 hours', *Bloomberg.com*, 28 February 2017 [online]. Available at <https://www.bloomberg.com/news/articles/2017-02-28/jpmorgan-marshals-an-army-of-developers-to-automate-high-finance> [Accessed 4 May 2018].
- ⁹ Casemine, [online]. Available at www.casemine.com [Accessed 4 May 2018].
- ¹⁰ Hannah Fairfield & Adam Liptak (2014) 'A More Nuanced Breakdown Of The Supreme Court' *NY Times*, June 26, 2014 [online]. Available at <https://www.nytimes.com/2014/06/27/upshot/a-more-nuanced-breakdown-of-the-supreme-court.html> [Accessed 22 May 2018], also see: Daniel Martin Katz, Michael James Bommarito & Josh Blackman, (2017) 'A General Approach for Predicting the Behavior of the Supreme Court of the United States [online]. Available at SSRN: <https://ssrn.com/abstract=2463244> or <http://dx.doi.org/10.2139/ssrn.2463244> [Accessed 22 May 2018].
- ¹¹ Rajshekhar, Kripa, Zadrozny, Wlodek & Garapati, Sri Sneha (2017) 'Analytics of Patent Case Rulings: Empirical Evaluation of Models for Legal Relevance' (July 12, 2017). Proceedings of the 16th International Conference on Artificial Intelligence and Law (ICAIL 2017), London, UK, June 12-16, 2017. Available at SSRN: <https://ssrn.com/abstract=3002782> [Accessed 22 May 2018].
- ¹² Bill Novomisle (2018) 'Deploying AI in the legal department' *In House Counsel*, March 21, 2018, [online]. Available at <http://www.inhousecommunity.com/article/deploying-ai-legal-department/> [Accessed 22 May 2018].
- ¹³ Lexmachina, [online]. Available at <https://lexmachina.com> [Accessed 4 May 2018]; Ravel, [online]. Available at <http://ravellaw.com/products/> [Accessed 4 May 2018]; Judicata, [online]. Available at <https://www.judicata.com/about> [Accessed 4 May 2018]; Loomanalytics, [online]. Available at <https://www.loomanalytics.com> [Accessed 4 May 2018].
- ¹⁴ Intraspection [online]. Available at <https://intraspection.com> [Accessed 4 May 2018]; CARA [online]. Available at <https://case-text.com> [Accessed 4 May 2018].
- ¹⁵ Anaqua, [online]. Available at <https://www.anaqua.com/corporate/products/anaqua-studio> [Accessed 4 May 2018]. Trademarknow, [online]. Available at <https://www.trademarknow.com> [Accessed 4 May 2018].
- ¹⁶ Anaqua, [online]. Available at <https://www.anaqua.com/corporate/products/anaqua-studio> [Accessed 4 May 2018]; Brightflag, [online]. Available at <https://brightflag.com/product> [Accessed 4 May 2018].
- ¹⁷ Catalyst, [online]. Available at <https://catalystsecure.com> [Accessed 4 May 2018].
- ¹⁸ Artificial Lawyer (2018) 'Deloitte Legal develops Legal + Regulatory AI tool in China', Feb. 7, 2018, www.artificiallawyer.com/2018/02/07/deloitte-legal-develops-legal-regulatory-ai-tool-in-china [Accessed 4 May 2018].
- ¹⁹ Jeff Pfeifer (2017) 'How Analytics Is Shaping the Current and Future Practice of Law, The Nature of Legal Work Today and the Need to Consume Vast Amounts of Unstructured Text Make our Profession a Ripe Target for the Promise of Machine Learning and Artificial Intelligence', *Law Journal Newsletters*, July 2017, [online]. Available at <http://www.lawjournalnewsletters.com/sites/lawjournalnewsletters/2017/07/01/how-analytics-is-shaping-the-current-and-future-practice-of-law-4/> [Accessed 22 May 2018].
- ²⁰ Nick Hilborne (2017) 'Law Firm Launches Data Analytics Team to Help Lawyers Predict the Future', *Legalfutures*, June 7 2017, [online]. Available at <https://www.legalfutures.co.uk/latest-news/law-firm-launches-data-analytics-team-help-lawyers-predict-future> [Accessed 22 May 2018].
- ²¹ Riverlaw.com, [online]. Available at <http://www.riverviewlaw.com/virtual-assistants/> [Accessed 22 May 2018].
- ²² Edgar Alan Rayo (2017) 'AI in Law and Legal Practice – A Comprehensive View of 35 Current Applications', *TechEmergence*, November 29, 2017 [online]. Available at <https://www.techemergence.com/ai-in-law-legal-practice-current-applications/> [Accessed 22 May 2018].
- ²³ Altman Weil (2017) 'Law Firms in Transition Survey' [online]. Available at http://www.altmanweil.com/dir_docs/resource/90D6291D-AB28-4DFD-AC15-DBDEA6C31BE9_document.pdf [Accessed 22 May 2018].
- ²⁴ Erin Winick (2017) 'Intelligent Machines, Lawyer-Bots Are Shaking Up Jobs', *MIT Technology Review*, December 12, 2017, [online]. Available at <https://www.technologyreview.com/s/609556/lawyer-bots-are-shaking-up-jobs> [Accessed 22 May 2018].
- ²⁵ Winick (2017), *ibid*.
- ²⁶ Marlene Jia (2018) 'Now that Lawyers Have Lost to AI, What is the future of law?' *TopBots*, March 8, 2018 [online]. Available at <https://www.topbots.com/future-of-law-legal-ai-tech-lawgeex/> [Accessed 6 May 2017].
- ²⁷ Scott A. Moss (2013) 'Bad Briefs, Bad Law, Bad Markets: Documenting The Poor Quality Of Plaintiffs' Briefs, Its Impact On The Law, And The Market Failure It Reflects', *Emory Law Journal* Vol. 63, p. 59. Available at http://law.emory.edu/elj/_documents/volumes/63/1/articles/moss.pdf [Accessed 4 May 2018].
- ²⁸ See for instance, Lex Machina, [online]. Available at <https://lexmachina.com/what-we-do/how-it-works/> [Accessed 4 May 2018].
- ²⁹ See for instance, Eliot Wrenn (2017) 'Must See Legal Technology to Deliver Better Answers Faster' [online]. Available at <https://legalsolutions.thomsonreuters.com/law-products/westlaw-legal-research/insights/must-see-cutting-edge-legal-technology-to-deliver-better-answers-faster> [Accessed 22 May 2018].
- ³⁰ Stephen Mason (2017) 'Artificial Intelligence: Oh Really? And Why Judges and Lawyers are Central to the Way we Live Now — But they Don't Know it' [online]. Available at http://stephenmason.co.uk/wp-content/uploads/2017/12/Pages-from-2017_23_CTLR_issue_8_PrintNEWMASON.pdf [Accessed 4 May 2018].

- ³¹ Buchanan, B., & Headrick, T. (1970) 'Some Speculation About Artificial Intelligence and Legal Reasoning', *Stanford Law Review*, Volume 23, No. 1, November 1970.
- ³² For a definition of algorithm see Christian Sandvig, Kevin Hamilton, et al. (2016) 'Automation, Algorithms, and Politics. When the Algorithm Itself is a Racist: Diagnosing Ethical Harm in the Basic Components of Software', *International Journal of Communication*, Vol. 10 [online]. Available at <http://ijoc.org/index.php/ijoc/article/view/6182> [Accessed 22 May 2018].
- ³³ 'Artificial Intelligence and the Practice of Law', Oct, 27, 2017 [online]. Available at <http://200.hls.harvard.edu/events/hls-in-the-world/artificial-intelligence-practice-law/> [Accessed 22 May 2018].
- ³⁴ See Susan Nevelow (2016) 'The Algorithm as a Human Artefact: Implications for Legal {Re}Search', October 26, 2016 [online]. Available at <https://ssrn.com/abstract=2859720> or <http://dx.doi.org/10.2139/ssrn.2859720> [Accessed 22 May 2018].
- ³⁵ Article 1(2), 'Regulation 201/679 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data'.
- ³⁶ See Goodman (2016) 'A Step Towards Accountable Algorithms? Algorithmic Discrimination and the General Data Protection Regulation in the European Union', 29th Conference on Neural Information Processing Systems.
- ³⁷ Recital 71, 'Regulation 201/679 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data'.
- ³⁸ Megan Garcia (2016) 'Racist in the Machine: The Disturbing Implications of Algorithmic Bias' *World Policy Journal*, 33(4), p. 115.
- ³⁹ Ivana Bartoletti (2018) 'Women Must Act Now or Male-designed Robots Will Take Over our Lives', *The Guardian*, March 13, 2018, [online]. Available at <https://www.theguardian.com/commentisfree/2018/mar/13/women-robots-ai-male-artificial-intelligence-automation> [Accessed 22 May 2018].
- ⁴⁰ Claire Cain Miller (2015) 'When Algorithms Discriminate', *New York Times*, July 19, 2015 [online]. Available at <https://www.nytimes.com/2015/07/10/upshot/when-algorithms-discriminate.html> [Accessed 22 May 2018].
- ⁴¹ Christian Sandvig & al. (2016) *ibid*.
- ⁴² Albanesi Chloé (2009) 'HP Responds to Claim of Racist Webcams', *Pcmag.com* [online]. Available at <https://www.pcmag.com/article2/0,2817,2357429,00.asp> [Accessed on 12 October 2018].
- ⁴³ Alex Hern (2018) 'Google's Solution to Accidental Algorithmic Bias', *The Guardian*, Jan 12 2018 [online]. Available at <https://www.theguardian.com/technology/2018/jan/12/google-racism-ban-gorilla-black-people> [Accessed 22 May 2018].
- ⁴⁴ Aylin Caliskan, Joanna J. Bryson & Arvind Narayanan (2017) 'Semantics derived automatically from language corpora contain human-like biases', *Science*, 356, 183–6; Garcia, *ibid* (2016), p. 112.
- ⁴⁵ Garcia (2016) *ibid*, p. 113.
- ⁴⁶ Marianne Bertrand & Sendhil Mullainathan (2004) 'Are Emily and Greg More Employable than Lakisha and Jamal? A Field Experiment on Labor Market Discrimination', *Am. Econ. Rev.*, 94, pp. 997–9.
- ⁴⁷ Dustin, J. (2018) 'Amazon Scraps Secret AI Recruiting Tool that Showed Bias against Women', Reuters.com [online]. Available at <https://www.reuters.com/article/us-amazon-com-jobs-automation-insight/amazon-scraps-secret-ai-recruiting-tool-that-showed-bias-against-women-idUSKCN1MK08G> [Accessed 11 October 2018].
- ⁴⁸ Petrasic Kevin & Benjamin Saul (2017) 'Algorithms and Bias: What Lenders need to Know', *Whitecase.com*, January 20, 2017 [online]. Available at <https://www.whitecase.com/publications/insight/algorithms-and-bias-what-lenders-need-know> [Accessed 12 October 2018].
- ⁴⁹ Tom Simonite (2017) 'Machines Taught by Photos Learn a Sexist View of Women', *Wired*, Aug 21 2017 [online]. Available at <https://www.wired.com/story/machines-taught-by-photos-learn-a-sexist-view-of-women/> [Accessed 22 May 2018].
- ⁵⁰ See Anupam Chander (2016) 'The Racist Algorithm?', UC Davis Legal Studies Research Paper Series, Research Paper N. 498.
- ⁵¹ See Goodman, (2016) *ibid*, p. 2. See Article 9 and 22, Regulation 2016/679.
- ⁵² See Goodman, (2016) *ibid*, p. 2–3.
- ⁵³ See Goodman, (2016) *ibid*, p. 2–3.
- ⁵⁴ See Goodman, (2016) *ibid*, p. 2–3.
- ⁵⁵ See Tomi Mester (2018) 'Statistical Bias Types Explained' [online]. Available at <https://data36.com/statistical-bias-types-explained/> [Accessed 22 May 2018].
- ⁵⁶ According to Black's Law Dictionary, 'selection bias' is defined as statistical error that causes one sampling group to be selected more than other sampling groups. It will create a bias in an experiment. The Law Dictionary: Selection Bias [online]. Available at <https://thelawdictionary.org/selection-bias/> [Accessed 22 May 2018].
- ⁵⁷ Ward, M. (2017) 'How Fake Data Could Lead to Failed Crops and Other Woes' *BBC.com*, March 21, 2017 [online]. Available at www.bbc.com/news/business-38254362 [Accessed 22 May 2018].
- ⁵⁸ Adam Stone (2017) 'When Big Data Gets It Wrong' *Government Technology*, March 2017 [online]. Available at <http://www.govtech.com/data/When-Big-Data-Gets-It-Wrong.html> [Accessed 22 May 2018].
- ⁵⁹ Stone (2017) *ibid*.

- ⁶⁰ Kalev Leetaru (2018) 'How Bad Data Practice Is Leading to Bad Research', *Forbes*, February 19, 2018, [online]. Available at <https://www.forbes.com/sites/kalevleetaru/2018/02/19/how-bad-data-practice-is-leading-to-bad-research/> [Accessed 22 May 2018].
- ⁶¹ Devin Coldewey (2016) 'Microsoft Apologizes for Hijacked Chatbot Tay's Wildly Inappropriate Tweets' *TechCrunch*, March 26, 2016 [online]. Available at <https://techcrunch.com/2016/03/25/microsoft-apologizes-for-hijacked-chatbot-tays-wildly-inappropriate-tweets/> [Accessed 22 May 2018].
- ⁶² Alessandro Bessi & Emilio Ferrara (2016) 'Social Bots Distort the 2016 US Presidential Election Online Discussion', *Fast Monday*, 21(11) [online]. Available at <http://journals.uic.edu/ojs/index.php/fm/article/view/7090/5653> [Accessed 22 May 2018].
- ⁶³ Leslie Scism (2017) 'Life Insurers Draw on Data, not Blood', *Wall St. Journal*, Jan 12, 2017 [online]. Available at <https://www.wsj.com/articles/the-latest-gamble-in-life-insurance-sell-it-online-1484217026> [Accessed 4 May 2018].
- ⁶⁴ John Lucker, Susan K. Hogan & Trevor Bischoff (2017) 'Predictably Inaccurate: The Prevalence and Perils of Bad Big Data', *Deloitte Review*, 21, July 31, 2017, [online]. Available at <https://www2.deloitte.com/insights/us/en/deloitte-review/issue-21/analytics-bad-data-quality.html#endnote-24> [Accessed 22 May 2018]; Megan L. Head, Luke Holmann et al (2015) 'The Extent and Consequences of P-Hacking in Sciences', *Plos Biology*, (2015). Available at <http://journals.plos.org/plosbiology/article?id=10.1371/journal.pbio.1002106> [Accessed 22 May 2018].
- ⁶⁵ Adam Lopez (2008) 'Statistical Machine Translation', *ACM Computing Surveys*, 40, 1–49.
- ⁶⁶ W. Scott Palmer (2012) 'Predictive Analytics with Predictable Problems', *Injury Sciences LLC*, May 4, 2012 [online]. Available at http://www.cccis.com/wp-content/uploads/2013/11/Predictable_Analytics_with_Predictive_Problems_White_Paper.pdf?x34637 [Accessed 22 May 18].
- ⁶⁷ Oliver Roeder (2015) 'Supreme Court Justices Get More Liberal as They Get Older' *FiveThirtyEight*, October 5, 2015 [online]. Available at <https://fivethirtyeight.com/features/supreme-court-justices-get-more-liberal-as-they-get-older/> [Accessed 22 May 2018].
- ⁶⁸ Will Knight (2016) 'AI's Language Problem', *MIT Technology Review*, Aug 9, 2016 [online]. Available at <https://www.technologyreview.com/s/602094/ais-language-problem/> [Accessed 22 May 2018].
- ⁶⁹ The Chinese word for 'Canada is pronounced in Mandarin as "jianáda."
- ⁷⁰ 'China Securities Regulatory Commission Amendments to the Securities Issuance and Underwriting Decision Law (2014).
- ⁷¹ Bill Novomisle (2018) 'Deploying AI in the Legal Department' *In House Counsel*, March 21, 2018, [online]. Available at <http://www.inhousecommunity.com/article/deploying-ai-legal-department/> [Accessed 22 May 2018].
- ⁷² University of Toronto (2018) 'Artificial Neural Networks' [online]. Available at <http://www.psych.utoronto.ca/users/reingold/courses/ai/nn.html> [Accessed 4 May 2018].
- ⁷³ Carlos E. Perez (2017) 'Why We Should Be Deeply Suspicious of Backpropagation', *Medium.com* [online]. Available at <https://medium.com/intuitionmachine/the-deeply-suspicious-nature-of-backpropagation-9bed5e2b085e> [Accessed 22 May 2018].
- ⁷⁴ See Knight (2017) *ibid*; Lisa Calhoun, 'Artificial Intelligence Poker Champ Bluffs its Way to 17 Million', *Inc*, Feb. 6 2017, [online]. Available at <https://www.inc.com/lisa-calhoun/artificial-intelligence-poker-champ-bluffs-its-way-to-17-million.html> [Accessed 22 May 2018].
- ⁷⁵ Garcia (2016) *ibid*, p. 116.
- ⁷⁶ Colin Scarlett (2017) 'The Future of Law: Artificial Intelligence?' *Colliers Knowledge Leader*, April 24 2017 [online]. Available at <https://knowledge-leader.colliers.com/colin-scarlett/future-law-artificial-intelligence/> [Accessed 22 May 2018]; Goodman (2016) *ibid*, pp. 3–4; Dave Gershgorn (2016) 'AI is Now so Complex its Creators Cannot Trust Why it Makes Decisions', *Quartz.com* [online]. Available at <https://qz.com/146753/ai-is-now-so-complex-its-creators-cant-trust-why-it-makes-decisions/> [Accessed 22 May 2018].
- ⁷⁷ Knight (2017) *ibid*.
- ⁷⁸ Rose Eveleth (2015) 'My Robot Bought Illegal Drugs', *Bbc.com*, July 21, 2017 [online]. Available at <http://www.bbc.com/future/story/20150721-my-robot-bought-illegal-drugs> [Accessed 2 May 2018].
- ⁷⁹ See Ellen Hunt (2016) 'Tay, Microsoft's AI Chatbot, Gets a Crash Course in Racism from Twitter', *TheGuardian.com*, March 24, 2016 [online]. Available at <https://www.theguardian.com/technology/2016/mar/24/tay-microsofts-ai-chatbot-gets-a-crash-course-in-racism-from-twitter> [Accessed 2 May 2018]; Garcia (2016) *ibid*, p. 111.
- ⁸⁰ Sarah Perez (2016) 'Microsoft Silences its New A.I. Bot Tay, After Twitter Users Teach It Racism' *TechCrunch*, March 24, 2016 [online]. Available at <https://techcrunch.com/2016/03/24/microsoft-silences-its-new-a-i-bot-tay-after-twitter-users-teach-it-racism/> [Accessed 22 May 2018].
- ⁸¹ Russell Cameron Thomas (2016) 'Tay Twist: @Tayandyou Twitter Account Was Hijacked ...By Bunglinh Microsoft Test Engineers', *Exploring Possibility Space*, March 30, 2016 [online]. Available at <http://exploringpossibilityspace.blogspot.hk/2016/03/tay-twist-tayandyou-twitter-account-was.html> [Accessed 22 May 2018].
- ⁸² Choe Sang-Hun & John Markoff (2016) 'Master of Go Board Game Is Walloped by Google Computer Program' *New York Times*, March 9, 2016. Available at <https://www.nytimes.com/2016/03/10/world/asia/google-alphago-lee-se-dol.html> [Accessed 2 May 2018].
- ⁸³ Knight, (2017) *ibid*.

- ⁸⁴ See Cade Metz (2016) 'In Two Moves AlphaGo and Lee Sedol Redefined the Future', *Wired.com*, March 16, 2016 [online]. Available at <https://www.wired.com/2016/03/two-moves-alpha-go-lee-sedol-redefined-future/> [Accessed 2 May 2018]; Bob van den Hoek (2016) 'Review of Game 2: AlphaGo's New Move and Devastating Aggression' *Deep Learning: Sky's the Limit?* April 23 2016 [online]. Available at <http://deeplearningskysthelimit.blogspot.hk/2016/04/part-7-review-of-game-2-alphagos-new.html> [Accessed 2 May 2018].
- ⁸⁵ See David Silver et al. (2017) 'Mastering the Game of Go without Human Knowledge' *Nature*, 550, pp. 354–9.
- ⁸⁶ Lucas Baker & Fan Hui (2017) 'Innovations of AlphaGo' [online]. Available at <https://deepmind.com/blog/innovations-alpha-go/> [Accessed 2 May 2018].
- ⁸⁷ Matt McFarland (2016) 'What AlphaGo's Sly Move Says About Machine Creativity', *The WashingtonPost.com*, March 15, 2016 [online]. Available at https://www.washingtonpost.com/news/innovations/wp/2016/03/15/what-alphagos-sly-move-says-about-machine-creativity/?utm_term=.4a1758bd8ad9 [Accessed 2 May 2018].
- ⁸⁸ See Scarlett (2017) *ibid.*
- ⁸⁹ George Krasadakis (2018) 'Artificial Intelligence: The Concerns', *Becoming Human*, Jan 26 2018. Available at <https://becoming-human.ai/artificial-intelligence-risks-concerns-2a19ba21cfd9> [Accessed 2 May 2018].
- ⁹⁰ Harvard University (2018) [online]. Available at <https://hls.harvard.edu/academics/curriculum/catalog/default.aspx?o=71157> [Accessed 2 May 2018]; Massachusetts Institute of Technology (2018) [online]. Available at <https://dam-prod.media.mit.edu/x/2018/07/30/Syllabus%20Ethics%20and%20Governance%20of%20AI%20.pdf> [Accessed 15 October 2018].
- ⁹¹ Georgetown University (2016) 'Computer Programming for Lawyers' [online]. Available at <https://cp4l.org/> [Accessed 15 October 2018].
- ⁹² Legg Michael (2017) 'UNSW Mini-Curriculum Review Report on Technology and the Law School Curriculum' [online]. Available at <http://classic.austlii.edu.au/au/journals/UNSWLRS/2017/90.pdf> [Accessed 15 October 2018], p. 11.
- ⁹³ The University of New South Wales (2018) 'Designing Technology Solutions for Access to Justice' [online]. Available at http://www.law.unsw.edu.au/form/Designing_Technology_Solutions_for_Access_to_Justice [Accessed 15 October 2018].
- ⁹⁴ Chris Zhang (2016) 'Cathy O'Neil, author of 'Weapons of Math Destruction,' on the dark side of big data, *Los Angeles Times*, Dec. 30, 2016 [online]. Available at <http://www.latimes.com/books/jacketcopy/la-ca-jc-cathy-oneil-20161229-story.html> [Accessed 23 May 2018]. See Garcia (2016) *ibid.*, p. 116.
- ⁹⁵ See Goodman (2016) *ibid.*, p. 4.
- ⁹⁶ See Palmer, (2012) *ibid.*
- ⁹⁷ See Article 42, 'Regulation 201/679 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data'.
- ⁹⁸ See Goodman (2016) *ibid.*, p. 4.
- ⁹⁹ Garcia (2016) *ibid.*, pp. 115–6.
- ¹⁰⁰ Garcia (2016) *ibid.*, p. 115.
- ¹⁰¹ See Articles 13–15 of the GDPR. In particular, Article 13(2)(f) reads "In addition to the information referred to in paragraph 1, the controller shall, at the time when personal data are obtained, provide the data subject with the following further information necessary to ensure fair and transparent processing:.. the existence of automated decision-making, including profiling, referred to in Article 22(1) and (4) and, at least in those cases, meaningful information about the logic involved, as well as the significance and the envisaged consequences of such processing for the data subject".
- ¹⁰² Garcia (2016) *ibid.*, p. 115.
- ¹⁰³ Goodman (2016) *ibid.*, pp. 3–4.
- ¹⁰⁴ David Gunning (2018) 'Explainable Artificial Intelligence (XAI)', *DARPA*, April 14, 2018 [online]. Available at <https://www.darpa.mil/program/explainable-artificial-intelligence> [Accessed 22 May 2018].
- ¹⁰⁵ Article 14 'Regulation 201/679 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data'.
- ¹⁰⁶ Article 12 'Regulation 201/679 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data'.
- ¹⁰⁷ See Andrew D Selbst Julia Powles (2017) 'Meaningful Information and the Right to Explanation' *International Data Privacy Law*, 7 (4), pp. 233–242.
- ¹⁰⁸ Cliff Kuang (2017) 'Can A.I. Be Taught to Explain Itself?', *The New York Times*, Nov 21 2017 [online]. Available at <https://www.nytimes.com/2017/11/21/magazine/can-ai-be-taught-to-explain-itself.html> [Accessed 22 May 2018].
- ¹⁰⁹ Kuang, (2017) *ibid.*
- ¹¹⁰ Kuang, (2017) *ibid.*
- ¹¹¹ D.H. Park, L. A. Hendrix, et al. (2017) 'Multimodal Explanation: Justifying Decisions and Pointing to Evidence' [online]. Available at <https://arxiv.org/pdf/1802.08129.pdf> [Accessed 22 May 2018].
- ¹¹² See Elliot Wremm (2017) 'Must See Legal Technology to Deliver Better Answers Faster', *Thomson Reuters* [online]. Available at <https://legalsolutions.thomsonreuters.com/law-products/westlaw-legal-research/insights/must-see-cutting-edge-legal-technology-to-deliver-better-answers-faster> [Accessed 11 October 2018].

- ¹¹³ SAS, 'Natural Language Processing' [online]. Available at https://www.sas.com/en_us/insights/analytics/what-is-natural-language-processing-nlp.html [Accessed 22 May 2018].
- ¹¹⁴ See Article 5(3)(k), Directive 2001/29/EC of the European Parliament and of the Council of 22 May 2001 on the harmonisation of certain aspects of copyright and related rights in the information society.
- ¹¹⁵ Manna, Martina (2014) 'The Concept of Parody and its Limits: The Recent Interpretation From the European Court of Justice' [online]. Available at <http://www.martinimanna.com/the-concept-of-parody-and-its-limits-the-recent-interpretation-from-the-european-court-of-justice/> [Accessed 16 October 2018].
- ¹¹⁶ A 2013 paper noted that increasing retrieval latency by search engines (i.e. employing 'slow search') can increase the quality of search results. See: Jaime Teevan, Kevyn Collins-Thompson, Ryan W. White, Susan Dumais, *Communications of the ACM*, Vol. 57 No. 8, pp. 36–8
- ¹¹⁷ Garcia (2016) *ibid*, p. 117.
- ¹¹⁸ Garcia (2016) *ibid*, p. 117.
- ¹¹⁹ Garcia (2016) *ibid*, p. 117.
- ¹²⁰ See Zeynep Tufekci (2016) 'Machine Intelligence Makes Human Morals More Important', TED.com. [online]. Available at <https://www.youtube.com/watch?v=hSSmmlridUM> [Accessed on 2 May 2018].
- ¹²¹ Marlene Jia (2018), *ibid*.

Biographies

Mr. Ronald Yu, Esq. (余克偉), JD, is a Board Member and a Founding Team member of IIPCC (www.iipcc.org), a non-profit international NGO focused on IP commercialization, enjoying the status of WIPO Permanent Observer as well as Advisor and Partner to WIPO Green and WIPO ADR. He is also a U.S. patent agent, IT engineer, digital forensics examiner and an inventor. He has published and lectured on a variety of topics including online learning, digital forensics, intellectual property and information technology.

Corresponding address: IIPCC, Suite 103, 77 Connaught Road, Central, Hong Kong. Email: ron.yu@iipcc.org.

Gabriele Spina Ali holds a PhD from The University of Hong Kong and is a postdoctoral fellow at the EW Barker Centre for Law & Business, National University of Singapore. His research interests lie mainly in the field of intellectual property, where he has published and researched in fields such as copyright, regulatory data protection and artificial intelligence.

Corresponding address: EW Barker Centre for Law & Business, Faculty of Law, National University of Singapore, Eu Tong Sen Building, 469 G Bukit Timah Road, Singapore 259776. Email: law.gsa@nus.edu.sg.