

Unpacking the Black Box: Addressing the ‘Social’ to Make Construction of AI-Powered Legal Technologies More Transparent and Unbiased

Siddharth Peter de Souza¹

I. Introduction

Over the past few years, there has been much debate about the impact of Artificial Intelligence (AI) on the legal profession. Much has been imputed to how changes in technology will impact the way legal research is conducted, how judges decide cases, how firms strategise for big value deals, and clients engage with legal information.

AI is defined as the capacity of machines to perform functions that are normally attributed to humans, such as the ability to find reason, make connections, generalise and adapt from experience.² Due to factors such as an increase in access to capital for companies to invest in research and development,³ the increase in incisive and effective algorithms,⁴ the availability of large data sets wherein product development can be tested and scaled,⁵ and the increase in computing power,⁶ there has recently been an increased impact of AI in law despite the conversation beginning over thirty years ago.⁷

While much of the conversation around AI and law has been about the technological advancements that such products bring, the economies of scale that they deliver and the challenges that automation will cause for jobs in the profession, this essay will focus on one particular aspect of AI technologies, the darker side, that of the ‘black box’ of these technologies – wherein there is much less understanding of how the systems work, and what constitutes their functioning.

The second section of this essay will describe the kind of products that are influencing and transforming the functioning of the legal profession to provide a

¹ Doctoral Candidate, Faculty of Law, Humboldt University of Berlin.

² ‘Artificial Intelligence |Encyclopedia Britannica’ <<https://www.britannica.com/technology/artificial-intelligence>> accessed 5 November 2017.

³ ‘Artificial Intelligence Explodes: New Deal Activity Record For AI’ <<https://www.cbinsights.com/research/artificial-intelligence-funding-trends/>> accessed 5 November 2017.

⁴ Ibid.

⁵ ‘What’s Driving the Machine Learning Explosion?’ <<https://hbr.org/2017/07/whats-driving-the-machine-learning-explosion>> accessed 5 November 2017.

⁶ Mark Purdy and Paul Daugherty, ‘Accenture: Why AI Is the Future of Growth’ <https://www.accenture.com/lv-en/_acnmedia/PDF-33/Accenture-Why-AI-is-the-Future-of-Growth.pdf>.

⁷ Siddharth de Souza, ‘Transforming the Legal Profession: The Impact and Challenges of Artificial Intelligence’ (*Digital Policy Portal*, 16 November 2017) <<http://www.digitalpolicy.org/transforming-legal-profession-impact-challenges-artificial-intelligence/>> accessed 16 January 2018.

context for the emerging changes. The third section will introduce an ethical challenge, using the more ambiguous construct of the 'black box' to highlight the implications that the opaqueness of the internal functioning of these technologies can have for the profession. The fourth section will propose how the use of social science techniques can offer a framework to provide more transparency in an otherwise cloudy framework.

II. AI and the legal profession: trends and possibilities

The field of AI is distinguished by different technologies, many of which have found application in the law. These include machine learning systems which are designed to mine large data sets and produce patterns without definite instructions;⁸ natural language processing systems which are able to evaluate texts to generate content and answer specific questions from the user;⁹ expert systems which provide solutions through decision-making as if they were human experts;¹⁰ speech recognition which helps to convert audio to text and vice versa; and vision systems which analyse images.¹¹

Doing a scan of the different AI-driven legal products provides an insight into the domains of the legal profession that are facing disruption. These include the areas of legal research, document review, e-discovery and predictive analysis. In terms of legal research, ROSS intelligence, which has been marketed as the first robot lawyer, enables lawyers to ask questions on particular legal issues, after which it analyses its database and provides concise answers and hypotheses through identifying patterns in the text.¹² In addition to legal research, Open Text, another AI driven platform, uses machine learning to uncover relationships and patterns in documents and facts of particular cases.¹³ Other platforms such as Kira perform contract review through analysing different clauses and concepts in contracts with the purpose of improving the capacity of lawyers to conduct due diligence accurately.¹⁴ In the field of predictive analysis, platforms such as "Lex Machina" have been developed with the purpose of providing lawyers with data driven insights into the behavior of lawyers, judges, and parties based on scans of litigation databases.¹⁵ These platforms are marketed as

⁸ 'Demystifying Artificial Intelligence' (*DU Press*) <<https://dupress.deloitte.com/dup-us-en/focus/cognitive-technologies/what-is-cognitive-technology.html>> accessed 5 November 2017.

⁹ *Ibid*

¹⁰ Michael Mills, 'Artificial Intelligence: The State of Play 2016' <<https://www.neotalogic.com/wp-content/uploads/2016/04/Artificial-Intelligence-in-Law-The-State-of-Play-2016.pdf>>.

¹¹ *Ibid*.

¹² 'ROSS Intelligence' <<http://rossintelligence.com/>> accessed 5 November 2017.

¹³ 'Accelerate EDiscovery & Investigations Solutions - Recommind' <<https://www.recommind.com/accelerate-ediscovery-product-page/>> accessed 5 November 2017.

¹⁴ 'Kira Systems | Machine Learning Contract Search, Review and Analysis' <<https://kirasystems.com/>> accessed 5 November 2017.

¹⁵ 'Lex Machina™ | LexisNexis' <<http://intl.lexisnexisip.com/products-services/intellectual-property-solutions/lexisnexis-lexmachina>> accessed 5 November 2017.

being tools to augment the capacity of lawyers to forecast the manner in which litigation plays out. Each of these platforms are designed to improve accuracy in legal research, reduce uncertainty and risks in terms of strategic decisions and save time and costs by enabling lawyers to spend more time on strategic tasks.¹⁶

The development of these different products has led to much anxiety in the legal profession as to whether it will result in a loss of jobs due to the automation of the profession. McKinsey has reported that over 22 percent of a lawyer's job and over 35 percent of a law clerk's job is at risk of being redundant due to such technologies.¹⁷ This suggests that there are going to be changes in terms of how the legal profession is structured, how legal education is delivered and how lawyers continue to make themselves relevant by using technology to augment their performance.¹⁸

While technology is shaking up the daily business of being a lawyer, and introducing new standards of economy and efficiency, it also opens up complexities and regulatory challenges that arise in the construction and development of these legal products.

III. The ethical challenge of the black box: questions of transparency and bias

In a study by Pro Publica of an algorithm called COMPAS (Correctional Offender Management Profiling for Alternative Sanctions) used by judges in the USA to determine the recidivism of a criminal defendant,¹⁹ it was found that black defendants were incorrectly judged to be at a higher risk of recidivism than they actually were, whereas white defendants were found to be less at risk of recidivism than they actually were.²⁰ It was also found that that black defendants were twice as likely to be misclassified as white defendants.²¹ This challenge of bias in AI systems is seen as a threat to justice, because there is a potential that without a clear understanding of the methods that have been used to construct these systems or the data used, it is likely that there will be human prejudices that are hidden in the system, both in terms of the data used to train it and in the construction of the product itself.²² The fact that technologies reflect the biases of their developers is

¹⁶ de Souza (n. 7).

¹⁷ Erin Winick, 'Lawyer-Bots Are Shaking up Jobs' (*MIT Technology Review*) <<https://www.technologyreview.com/s/609556/lawyer-bots-are-shaking-up-jobs/>> accessed 16 February 2018.

¹⁸ de Souza (n. 7).

¹⁹ Jeff Larson Julia Angwin, 'Machine Bias' (*ProPublica*, 23 May 2016) <<https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing>> accessed 16 February 2018.

²⁰ Julia Angwin Jeff Larson, 'How We Analyzed the COMPAS Recidivism Algorithm' (*ProPublica*, 23 May 2016) <<https://www.propublica.org/article/how-we-analyzed-the-compas-recidivism-algorithm>> accessed 16 February 2018.

²¹ *ibid.*

²² Will Knight, 'Google's AI Chief Says Forget Elon Musk's Killer Robots, and Worry about Bias in AI Systems Instead' (*MIT Technology Review*)

also seen in the fact that many virtual assistants like Apple's Siri or Amazon's Alexa, which are usually used for domestic work, are presented as women, whereas lawyer bots are represented as men — reinforcing gender stereotypes.²³ Only after being called out on the ways in which these virtual assistants were reinforcing sexism and in addition were not fighting back after being called derogatory things such as “bitch” and “slut” by its users, did Amazon partially resolve the problem, and program Alexa to disengage and shut down if the user used demeaning language.²⁴

In addition to the challenge of bias, there is also the question of transparency. Many of the products that use AI technologies like machine learning or natural language processing end up generating layers of complexity and patterns that often become difficult even for their creators to understand.²⁵ These systems take the information and process it through commands and webs of networks akin to the neural network in the human brain, which allows machines to solve problems on their own — a process called deep learning.²⁶ As these programs assume an independent internal development, they cannot be easily unmasked, and thus assume the form of a ‘black box’. Such opaqueness in technology, when applied to fields such as law, as demonstrated in the COMPAS case, becomes problematic. It is therefore imperative that a method is developed that requires these systems to explain how they arrive at particular conclusions, such that using them does not require ‘a leap of faith’ for both the creator and the user to whom they should be accountable.²⁷

IV. Unpacking the black box

The technical aspects of the development of algorithms and technologies are typically shrouded in secrecy. However, in order to unmask how these systems function, it is useful to examine the social relationships and processes that result in the construction of these products.²⁸ Introducing sociological insights can help to reveal the character of AI-driven legal products as well as provide a context for the

<<https://www.technologyreview.com/s/608986/forget-killer-robotsbias-is-the-real-ai-danger/>> accessed 16 February 2018.

²³ Natasha Mitchell for Science Friction, ‘Alexa, Siri, Cortana: Our Virtual Assistants Say a Lot about Sexism’ (*ABC News*, 11 August 2017) <<http://www.abc.net.au/news/2017-08-11/why-are-all-virtual-assistants-female-and-are-they-discriminatory/8784588>> accessed 16 February 2018.

²⁴ Ian Bogost, ‘Sorry, Alexa Is Not a Feminist’ [2018] *The Atlantic* <<https://www.theatlantic.com/technology/archive/2018/01/sorry-alexa-is-not-a-feminist/551291/>> accessed 16 February 2018.

²⁵ Will Knight, ‘The Dark Secret at the Heart of AI - MIT Technology Review’ <<https://www.technologyreview.com/s/604087/the-dark-secret-at-the-heart-of-ai/>> accessed 5 November 2017.

²⁶ Ariel Bleicher and Ariel Bleicher, ‘Demystifying the Black Box That Is AI’ (*Scientific American*) <<https://www.scientificamerican.com/article/demystifying-the-black-box-that-is-ai/>> accessed 14 February 2018.

²⁷ Knight (n. 25).

²⁸ Reza Banakar and Max Travers, ‘Introduction to Theory and Method in Socio-Legal Research’ (Social Science Research Network 2005) SSRN Scholarly Paper ID 1511112 <<https://papers.ssrn.com/abstract=1511112>> accessed 18 February 2018.

political, economic or cultural influences and decisions that determine the development of these technologies.²⁹ Exploring the 'social' in the evolution of these technologies can also offer an insight into the network of power structures of people, finances, processes and technologies that influence these products.

In many ways, technologies reflect the worldview of the people who develop them.³⁰ These typically tend to be scientists or engineers who introduce a technocratic approach to the development of the product.³¹ An argument can be made that by diversifying the pool of developers to include other disciplines, such as sociologists, designers, historians, and psychologists, a multiplicity of views will be brought to the table.³² While this may at first seem obvious, it is also critical because many of these legal products are actively being used to offer solutions for how judges decide cases or how offender risk is assessed — attributes which go beyond just textual interpretations of the law to instead study the context in which particular cases have emerged. Introducing a plurality of views would ensure a more balanced outlook on the use, development and management of data and methods that are being used to build the AI-driven legal products.

A Thomson Reuters report found that in 2016, 579 patents were filed in legal services technologies, compared to 99 in 2012, which amounted to a rise of over 484 percent in the four-year period.³³ This trend is being driven due to businesses looking at new avenues for legal advice and alternative legal providers entering into the market due to changes in regulatory practices, particularly in the UK, USA and Australia.³⁴ The rising demand from business as well as the increase in equity funding of over 1.5 billion dollars in 2016, with major investors in AI (broadly) including Google, Intel and Khosla Ventures, has spurred an increase in development.³⁵ An analysis of the scale, purpose and the diversity of investments in AI will provide an insight into the strategies of funding across different industries and how they relate to particular kind of products being developed for the legal profession. With firms, and traditional legal providers being compelled to respond to demands from clients and alternative legal providers, an analysis of drivers such as regulatory changes could allow for explanations as to why particular aspects of the profession are being automated as opposed to others.

²⁹ Roger Cotterrell, 'Why Must Legal Ideas Be Interpreted Sociologically?' (1998) 25 *Journal of Law and Society* 171.

³⁰ Kriti Sharma, 'Can We Keep Our Biases from Creeping into AI?' (*Harvard Business Review*, 9 February 2018) <<https://hbr.org/2018/02/can-we-keep-our-biases-from-creeping-into-ai>> accessed 16 February 2018.

³¹ 'Why AI Needs the Humanities' <<http://blog.nextit.com/popular/why-ai-needs-the-humanities/>> accessed 23 May 2018.

³² Sharma (n. 30).

³³ 'Thomson Reuters Analysis Reveals 484% Increase in New Legal Services Patents Globally' (*Thomson Reuters*, 16 August 2017) <<https://www.thomsonreuters.com/content/thomsonreuters/en/press-releases/2017/august/thomson-reuters-analysis-reveals-484-percent-increase-in-new-legal-services-patents-globally.html>> accessed 5 November 2017.

³⁴ *ibid.*

³⁵ 'Artificial Intelligence Explodes: New Deal Activity Record For AI' (n. 3).

Linked to the criticism of the opacity of the processes and technologies through which AI products arrive at decisions, there have been arguments for introducing an 'ethical black box', which would establish a process for discovering how and why a robot acted in a particular way, similar to the way in which a flight data recorder tracks and transmits internal data.³⁶ The purpose of this intervention is that robots will be making decisions that often require a moral compass, and introducing such a framework would allow for accountability and transparency in their functioning, in addition to public trust in their processes.³⁷ This is especially relevant in the legal field where, as the example of the case of machine bias in recidivism technologies demonstrates, the development of these legal products cannot just be carried out in a technocratic manner, but should instead be conducted cognisant of the social and cultural implications of AI-assisted decisions. As technologies increasingly adopt processes that supplant evaluation done by humans, they should also be held to criteria of predictability, inspection and accountability as would any human official in a similar position.³⁸ Each of these aspects require that the framework and algorithms that go into designing the processes and technologies of AI products adopt elements of social, ethical and moral reasoning, because the implications of the decisions of many of these products are entering into spheres that consist of assessment, appraisal and judgement, with profound implications for humans.

V. Conclusion

This essay has sought to explore the advancement of AI in the legal profession by considering some of the innovations that are disrupting the profession. It has focused particularly on some of the ethical implications in the legal sphere – those of transparency and bias in terms of how these products are constructed. The essay has suggested that addressing the social will allow for a more holistic consideration of the increasingly critical functions performed by technologies in the legal domain. By scrutinising the circumstances and actors involved in the construction and deployment of these technologies (particularly three stakeholders – people, finance and processes), they can be opened up to scrutiny and review. Unpacking the “black box” of these technologies can make them more trustworthy, understandable, and accountable.

³⁶ AF Winfield and others, 'The Case for an Ethical Black Box' in Yang Gao and Yang Gao (eds), *Towards Autonomous Robot Systems* (Springer 2017) <<http://eprints.uwe.ac.uk/31760/>> accessed 18 February 2018.

³⁷ *ibid.*

³⁸ Jim Torresen, 'A Review of Future and Ethical Perspectives of Robotics and AI' (2018) 4 *Frontiers in Robotics and AI* <<https://www.frontiersin.org/articles/10.3389/frobt.2017.00075/full#B13>> accessed 18 February 2018.