IJAIL

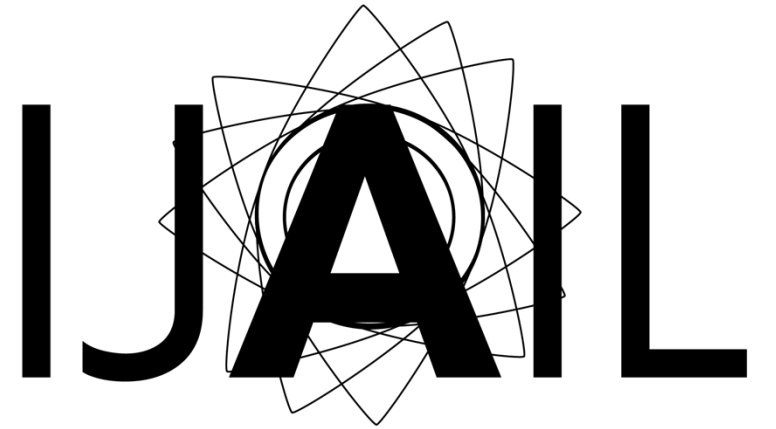
Indian Journal of Artificial Intelligence and Law

Volume 2, Issue 2 (2022)
e-ISSN: 2582-6999
April 15, 2022

ISSN 2582-6999
9 772582 699002

Abhivardhan, Editor-in-Chief
Aditi Sharma, Chief Managing Editor

Supported by
INDIAN SOCIETY OF ARTIFICIAL INTELLIGENCE & LAW



Indian Journal of Artificial Intelligence and Law

e-ISSN: 2582-6999

Volume 2, Issue 2 (April 2022)

e-ISSN: 2582-8398.

Printed and distributed online via **the Indian Society of Artificial Intelligence & Law (isail.in)** in the Republic of India.
Volume: 2
Issue: 2
Date of Publication: April 15, 2022.

Publisher: Abhivardhan C/O Indian Journal of Artificial Intelligence and Law,
8/12, Patrika Marg,
Civil Lines, Allahabad, Uttar Pradesh, India – 211001

For the purpose of citation, please follow the format for the list of references as follows:

2022. Indian Journal of Artificial Intelligence and Law, e-ISSN: 2582-6999. Prayagraj.

You can also cite the book through citethisforme.com (recommended).
For Online Correspondence purposes, please mail us at:
journal@isail.in;

For Physical Correspondence purposes, please send us letters at:
8/12, Patrika Marg,
Civil Lines, Allahabad, Uttar Pradesh, India - 211001

Preface

The Indian Journal of Artificial Intelligence and Law is a biannual law journal covering technology law in a combination of theoretical and practical approaches. It also provides coverage of the relationship between law and artificial intelligence in businesses, education, research and innovation practices. The journal publishes 2 issues per year in due frequency. This journal is supported by Indian Society of Artificial Intelligence and Law.

I would like to express my deepest of gratitude to our esteemed Managing Editors and the Associate Editors for their contribution towards the Journal and its efforts.

Abhivardhan
Editor-in-Chief
Indian Journal of Artificial Intelligence and Law.

Table of Contents

Technical Articles

Interview Transcripts for AI Now in partnership with the Indian Society of Artificial Intelligence and Law for AI Now

Editorial Board

Technical Articles

1 Thaler v Commissioner of Patents [2021] FCA 879: DABUS – An 'Inventor'?

Rahul Kanna & Pallavi Singh

Jindal Global Law School, Sonepat, India

Abstract. Artificial Intelligence is arguably one of the greatest creations of mankind encapsulated in the dawn of the 21st century, however, the growth of Artificial Intelligence has spiked over the past couple of decades owing to the technological revolution and widespread use of internet technology and partly due to the globalization of the world economy as a whole. The first mention of AI has occurred in the nascent period of the internet and technology industry specifically in the Dartmouth Summer Research Project on AI in 19561. Today there has been around 3,40,000 AI associated inventions were filed for patents globally and a further 16 Lakh AI-based publications in scientific directories2 have been subsumed. The growth of AI-based innovations has been gradually increasing and have become a part & parcel of everyday operations of corporations, governments and humans as well both consciously and subconsciously we have been using AI for the ease and betterment of our daily lives. The topic of Artificial Intelligence as a result is being considered by policymakers and global leaders across the world 3whilst innumerable scientific publications and scholarly work is being conducted constantly.

This article focuses on the legal realm of Artificial Intelligence devices and structures in pertinence to the Intellectual Property law whilst focusing on the focal argument of patentability and assigning of the term 'inventor' under the current Intellectual Property law ambit through a comparative jurisprudential analysis of the leading global leaders of AI innovations and IP Laws. The use case for the thesis enumerated above would be Thaler v Commissioner of Patents [2021] FCA 879 and 'DABUS' AI device. Further, the article would cover the jurisprudence of Europe and the United States

along with leading countries with a rich jurisprudence of AI-based legal framework under the universe of patent laws. Finally, the article would embark on the potential practical implications and the advantages for the development of AI systems for a sustainable future.

Introduction

Artificial Intelligence has become a legal matter of contention and particularly with its relationship with Intellectually Property Rights which can be evidenced with WIPO has commenced towards bringing in various stakeholders of AI including the public for a conversation on the interplay and ramifications circumventing AI and IP laws particularly concerning patent laws (World Intellectual Property Organisation, 2019). DABUS ("device for the autonomous bootstrapping of unified sentience") is an AI system or device architected by Dr Stephen Thaler pioneer of AI development and the current issue in contention is a result of the patent filed by Mr Thaler across several jurisdictions in the world including the USA, UK, Australia & South Africa etc. This has intensified the debate into the implications and potential of AI devices or systems to be granted a patent and to be recognized as an 'inventor'. Several jurisdictions have taken up the matter and the analyses would be critically interpreted in the latter part of the article. Thus, in the current case the Australian court had to answer certain issues concerning the contention as to whether an 'inventor' can be a non-human i.e., AI if so can it be granted a patent and whether such patent can be transferred to the owner of the AI device are the most pertinent questions that have been the arterial focus of the matter in contention.

DABUS – Jurisdictional Approach

Australia

Thaler v Commissioner of Patents [2021] FCA 879 is a case involving several facets of intellectual property rights with its focus aimed towards 'patent 'rights. The federal court of Australia in the

current case has given an exhaustive analysis of the case by taking into all the statutory regulations[1] circumventing the issue including adequate references to international treaties and regulations which Australia has ratified and been an active member since its inception. The current case is the solitary jurisprudence on the patentability and inventor states of AI in the current global IP structure. The Honourable Justice Beach in the current case primarily diverged into the background of AI system and the very foundational structure behind the operation of AI in this case concerning DABUS which was alleged to perform its tasks through artificial neural networks similar to the neural network present in the human brains though not as advanced as the biological brain as estimates believe that 'strong AI' equivalent to the performance of the human brain would not be possible for the near future and appallingly the median year for the accomplishment of the same was recorded by some scientist and scholars around the year 2099 (Ford, 2018 p. 528). The artificial neural networks are a subdivision in the field of machine learning and in an advanced manner as they possess these neural links similar to the neurons in the brain except for the fact that these are binary[2].

To understand the legal implication of the DABUS system in relevance to patent laws it would be important to primarily understand certain basics of the system, DABUS it possess a mix of two types of artificial neural networks while the former produces

---

[1] Patents Act 1990 (Cth) ss 2A, 3, 7, 18, 15, 29, 29A, 40, 45, 59, 64, 101B, 101E, 113, 138, 142, 172, 182, 185,
208; sch 1, Intellectual Property Legislation Amendment (Raising the Bar) Regulation 2013 (No. 1) (Cth), Intellectual Property Legislation Amendment (TRIPS Protocol and Other Measures) Regulation 2015 (Cth), Patents Regulations 1991 (Cth) regs 3.1A, 3.2A, 3.2B, 3.2C, 3.18, Patent Cooperation Treaty (Washington, 19 June 1970) arts 4, 9, 27, 5. As mentioned under the legislation division of the original judgement of the federal court.

[2] BEACH J, had used the interpretation he had provided in the case of Otsuka Pharmaceutical Co Ltd v Generic Health Pty Ltd (No 2) (2016) 120 IPR 431; [2016] FCAFC 111 (Para 135-138) where a detailed discourse on neural networks.

output based on disturbance leading to self-stimulation of the network and the second type of these networks classify these signals in a hierarchical basis thereby conduction evaluation of the output and assessing its novelty whilst eliminating the remaining and ultimately resulting in novel creations, much similar to the human brains cerebral cortex and thalamus relationship. Justice Beach asserts that DABUS performs neurocomputing much similar to human brain activity.

Commissioner of Patents contentions

The commissioner iterates that in line with provision Section 15(1)(a)[3] that inventor refers as being human and there is no room for a non-human interpretation taking into account the necessary literature such as the dictionary meaning of the word 'inventor' itself arguably refers to that of a human being. Further, the application was filed through the PCT as a result the commissioner argues that the word 'inventor' as mentioned in the act[4] bears the same meaning as Section 15(1)(a). He further claims that the true intent and object of Section 2A object i.e. "balances over time the interests of producers, owners and users of technology and the public" can be given due effect only if the inventor is said to be a human being and without delving further into legal reasoning stated that the matter as to what is the true intention of the statutes object and for it, to aide the owners of AI the commissioner left the ambiguity to be dealt by the parliament for enhanced clarity on the issue, finally, the commissioner in the concluding remark iterated that Section 15(1)

[3] 15(1) provides: Subject to this Act, a patent for an invention may only be granted to a person who: (a) is the inventor; or (b) would, on the grant of a patent for the invention, be entitled to have the patent assigned to the person; or (c) derives title to the invention from the inventor or a person mentioned in paragraph (b); or (d) is the legal representative of a deceased person mentioned in paragraph (a), (b) or (c).

[4] Reg 3.2 C(2)(aa) Patents Regulations 1001 (Cth) " provide the name of the inventor of the invention to which the application relates".

(a),(b),(c) have no application as DABUS can neither be an inventor nor can grant the patent to the owner Mr Thaler due to lack of legal status of DABUS.

Judgement

The court opined on the fact that there is an acute absence of any provision or regulation mandating that artificial intelligence is beyond the scope of being considered an 'inventor' as per the mandate of the patent rules and regulations applicable in Australia[4]. The court whilst engaging in the interpretation ruled out the commissioner interpretation of the legal meaning of the word to that of the dictionary which the court contradicted stating dictionary meanings by nature are 'exemplary and inclusive'[4]. Further added that dictionaries are updated by usage and over changes in the socio and technological advancements while iterating that dictionary meaning are not to be strictly observed as statutory interpretation in the absence of legal definitions[4]. The argument of 'inventor' mandatorily being a human was rejected as the law gives room for body politic as well as corporation. The court however concurred that the DABUS device was incapable of being granted a patent nor assign or own a patent due to the absence of any legal standing. However, the question of the AI system being an 'inventor' was answered in the affirmative. One of the reasonings for an AI being capable of being an 'inventor' was with regard to the word 'inventor' being an agent noun, thus adding a suffix 'er', 'or' and therefore the noun in effect postulates the agent undertaking the act for instance 'printer', 'computer', 'earthmover' etc. Drawing from this analogy the agent can either be a person or thing, thus removing the compulsion of the 'inventor' being a human and the artificial intelligence, as a result, can be regarded an 'inventor' in this parlance. The court finally delved into the issue after stating that DABUS can be identified as an inventor for the act however it can neither be an applicant for a patent nor can it legally assign the same. The court concluded that however, the patent could be assigned under Section 15(1) (c) as the requirement of communication and the constricted interpretation as

the means of title transfer through an assignment (High Court of Australia, 1923) was rejected as contended by the commissioner. The court gave a broader construction and interpretation whilst holding that the DABUS being under the control and ownership of the applicant Mr Thaler is ought to have communicated the invention to him effectively. Therefore, the court was of the opinion that on a fair and broad reading of Section 15(1) (b) & (c) the patent can be given to a legal person in a situation where the invention is carried forth by an artificial intelligence system or device as the 'inventor'.

Europe (EPC)

In the crusade of arguing the patentability of artificially generated inventions over time reducing the involvement of humans in the process14 and as a matter of fact, it is argued by Prof Abott (Abbott, 2016 p. 1079) that computers have been effecting patentable innovations since the dawn of the 21st century and it would be redundant to overlook the rich jurisprudence deliberated in Europe where the European Patent Convention (EPC)[5] serves as the foundation for the legal framework behind patents across Europe. The EPC excludes mathematical computations and computer programs from being patentable in character except for reasons technical in nature[6] and beyond the general expectations on the interaction between the software and hardware of the system. The European Patent Office in line with recent developments has provided guidelines for the qualifications of patentability in relevance to artificial intelligence[7] The European Parliament needs to be commended for taking up a resolution towards better understanding, awareness and applicability of AI in intellectual property rights

[5] Please refer to: European Patent Convention (Convention on the Grant of European Patents),1973 and subsequently revised in two instances once in 1991(Art 63 EPC) and 2000(Art 29 EPC)

[6] Case T-1173/97 Computer Program Product/IBM, Technical Board of Appeal 1/7/1998. Refer to Article 52(2) & 52 (3) EPC.

[7] Please refer to: The European Patent Office, Guidelines for Examination, G-II 3.3.

(Committee on Legal Affairs, European Parliament, 2020). The EPO does not exclude patentability of 'inventive' outputs registered by an AI and further the process behind such an invention is deemed inapplicable (Blok, 2017 p. 69). Similarly, another angle that needs to be considered is from both an international perspective and domestic stance is synonyms to the dictum that whilst granting of patents needs to be scrutinised in-depth concerning novelty, inventive step, industrial applications & disclosure it must be borne in mind that the same should be devoid of any discrimination towards the fields of technology[8].

DABUS system filed patent applications through Dr Thaler in the EPO much similar to the filings made in Australia and the patent office in Europe had rejected the contentions due to the absence of a human characteristic of the 'inventor' which the EPO argued was in contradiction to the article and rules of the EPO[9]. The EPO contended that as per the auspices of the provisions the details of the inventor needed to be included in the application i.e., address, surname etc., and therefore the addition of a machines name (DABUS) would not be acceptable to the statutory interpretation of the provision (European Patent Office, 2020). It is important to note some of the arguments raised by the EPO office with regard to the definition of an 'Inventor'. It was argued that 'Inventor' is referred to as a natural person (European Patent Office, 2020) and the same is also an internationally applicable standard (European Patent Office, 2020). and as a result, the human character standard applies to a majority of international jurisdictions, domestic courts[10] and other EPC members (European Patent Office, 2019). Lastly, the EPO concluded that none of its members has opined that an 'Inventor' could be non-human and if so particularly that AI can be an 'Inventor'

[8] Please refer to Article 27 TRIPS (The Agreement on Trade-Related Aspects of Intellectual Property Rights) & Article 52 EPC.

[9] Please refer to: Article 81 & Rule 19(1) EPC.

[10] Please refer to: UK Case laws of University of Southampton's Applications [2004] EHWC 2107 (Pat) [39], Yeda Research v Rhone-Poulenc [2007] UKHL 43 [20].

(European Patent Office, 2019). The most important aspect of the decision which was opined in the affirmative by the Australian Federal Court was rejected with regard to assignment and ownership of DABUS, the EPO contends that the applicant Dr Thaler would not be the employer of the DABUS device and ruled out the successor in title to him citing a lack of legal personality of DABUS prevents it from either employment or transfer of title (European Patent Office, 2019). Lastly, the TRIPS agreement for the protection of technology argument put forth by Dr Thaler was rejected as the issue is two-faced one concerning mere administrative details of 'inventor' such as name, address etc. whilst the other with relevance to the practical grant of patent i.e. the patentability of the AI system/device (European Patent Office, 2019).

United Kingdom (UKIPO)

The UKIPO handled similar issues with regard to DABUS on the application filed by Dr Thaler and held that a non-human inventor cannot be accepted in the current laws of the UK[11] and only a natural persona can be accepted as an 'Inventor'. The patent office while giving its decision had conceded that there is limited jurisprudence on the matter and with regard to the current statutory provisions the interpretation would only lead to the fact that an 'Inventor'needs to possess human characteristics and thus ruling out any machines and AI devices (DABUS) from being named an 'inventor' (UK Intellectual Property Office, 2019). The UKIPO also delved into the definition of the term 'Person' as used in Section 7 & 13 holdings that having limited jurisprudence and directions from the courts or legislature it would be redundant to hold that the term can refer to non- humans such as machines & AI (UK Intellectual Property Office, 2019). However, it is interesting to note that the office had stated that the accommodation of the machine or AI inventions would be possible in the future whilst stating that the present architecture of the law and

[11] Section 7 and 13, UK Patents Act,1977.

legal reasoning would not accommodate such a request and it is for the parliament to take up the matter and devise a new legislations for the accommodations of such AI induced innovations (UK Intellectual Property Office, 2019). The High Court of Appeal upheld the contentions of the UKIPO (The High Court of Appeal, 2020). The overall contention of the court can be summarised that the current law limits the application of non-human under the ambit of the definition of 'Inventor', however, the court stated that the issues raised in the current claim were more suitable for legislature and policymakers to delve into the matter and make the necessary modifications or amendments of the provisions under the UKIPO such as Section 7 & 13. It is interesting to note the critic of the judgement as stated by Prof Abott that delay in the overhaul of the IP laws would ultimately result in limitation of AI induced innovations and a situation where humans take credit for the work carried on by AI (English, 2021).

US (USPTO)

Dr. Stephen Thaler's first two applications for products created by DABUS were considered incomplete, and therefore, rejected by the United States Patent and Trademark Office (USPTO) on the grounds that Dr. Thaler failed to name a natural person as the inventor. The USPTO justified their stance by explaining that US legislation repeatedly refers to inventors as natural persons; thus, allowing a broader interpretation of the term 'inventor' would "contradict the plain reading of patent statutes." (Dhaliwal, 2020)

Dr. Thaler appealed the decision to the District Court of Eastern Virginia seeking for the declarations that a patent application for an invention generated by AI that meets the inventorship criteria, should list the AI as inventor; and that a patent application for inventions generated by AI should not be rejected on the basis that a natural person has not been identified as the inventor. The District Court rejected the appeal for reasons largely similar to those of the USPTO. They argued that the words 'inventor' and 'individual' were limited to

natural persons. Dr. Thaler argued that the general purpose of the patent laws would be furthered by granting AI inventor's rights as it would promote innovation by incentivizing the development of AI, and encourage the commercialization and disclosure of AI information. Dr. Thaler added that this would also protect the moral rights of human inventors, which is also part of the general purpose. The Judge however dismissed these arguments stating that Dr. Thaler failed to provide a justifiable reason for the Court to consider policy reasons instead of following the plain meaning of the statute. Therefore, the Court ruled that Dr. Thaler's arguments were not sufficiently strong to convince the Court to ignore the fact that it was abundantly clear that Congress intended to have the definition of the term 'inventor' be limited to natural persons only. The Court added that while there may come a time where AI reaches a level of sophistication that may fulfil the criteria for inventorship and that time has not arrived yet. They explained further that when the time does arrive, it will be up to the US Congress to legislate upon the same and decide whether or not they find the need to expand the scope of patent law to allow AI to be considered inventors (United States District Court for the Eastern District of Virginia, 2021).

South Africa

In July 2021, a patent application that listed DABUS as the inventor was granted by the South African Patent Office; making it the first country in the world to do so. However, South Africa is a non-examining country, that is, South African patent applications are not examined to check whether patent requirements are met. Additionally, South Africa does not require the inventor to divulge any previous art. Due to these reasons, any patent application in South Africa is granted if formal requirements are met, and are subject to third-party objections. Given that South African patents are not thoroughly scrutinised, many believe that the patent being approved was merely an oversight by the Patent Office (Patil, 2021).

Limitations of AI-Inventors

Given that AI-inventors are an extremely new concept, and that it is difficult to ascertain how AI as a field will evolve, legislating upon the same is a mammoth task. Many questions remain unanswered such as the difference between the terms 'automated' and 'autonomous', and even the difference between 'AI-generated' or aided by AI. Unfortunately, these questions will remain unanswered until AI-inventors are acknowledged by the global legal community, which at the time seems to be difficult to say the least.

Further, it has been suggested that AI-inventors may lack the ingenuity that is required to ideate and generate inventive solutions, given that they lack creative thinking like natural-persons. However, a paper by the World Economic Forum makes the claim that 'AI is no longer "just crunching numbers" but is generating works of a sort that have historically been protected as "creative" or as requiring human ingenuity' (Kim, 2020). Apart from this, patent infringement liability by AI is another issue that remains unclear. In cases of patent infringement, it is the infringer who is liable to pay compensation to the owner of the patent. However, in case of an infringement by an AI-inventor, it is unclear who the liability falls upon. As per the European Parliament Resolution of 16 February 2017, if a third party faces damages due to AI, the natural person behind the AI is to be held liable for the same, not the AI itself; which could be what is applied in cases of patent infringements by AI as well (World Economic Forum, 2018). If the natural person responsible for the Ai is to be held liable, it must be ensured that the liability is proportionately placed with respect to the authority that is delegated to the AI.

Conclusion

The legal discourse surrounding AI-inventors is currently in its nascent stage. While many countries seem averse to the idea of expanding the scope of 'inventors' to include AI along with natural

persons, we believe that it is inevitable. Given that AI generated inventions are unpredictable and somewhat boundless and that each country has different capacities and capabilities of adapting to new inventions, at this stage, the legal, economic, and social impact of AI-inventors can only be speculated upon. The current benchmark for patentable inventions ought to be re-evaluated to accommodate for the changing scenario. The question regarding whether inventions that are autonomously created by AI-inventors are to be considered and protected under the gambit of patent law must be answered post scrutinising the possible effects, both positive and negative.

The argument, that certain jurisdictions relied upon to state that AI is not eligible to be an inventor solely due to the fact that the statutes that define inventors describe natural-persons is questionable and short-sighted; given that the statutes were drafted long before AI was invented, and that it is inevitable that soon than later AI-inventors will be significantly impactful, and lawmakers will be forced to create legislations that include the same. In fact, it will only be a competitive advantage for countries that accept AI-inventors as early as possible. On February 19, 2020, the European Commission ("EC") presented its proposal for comprehensive regulation of artificial intelligence ("AI") at European Union ("EU") level: the "White Paper on Artificial Intelligence – A European approach to excellence and trust" ("White Paper"). In our opinion, it is best that the rest of the legal community follow suit with the EC and begin to accept AI- inventors in the current scenario, instead of leaving the revolutionization of AI, AI-inventions, and intellectual property laws to chance.

References

World Intellectual Property Organisation. 2019. WIPO Conversation on Intellectual Property (IP) and Artificial Intelligence (AI), Second Session, DRAFT ISSUES PAPER ON INTELLECTUAL PROPERTY POLICY AND ARTIFICIAL INTELLIGENCE, WIPO/IP/AI/2/GE/20/1. *World Intellectual*

Property Organisation. [Online] December 13, 2019. wipo_ip_ai_2_ge_20_1.pdf.

Ford, Martin. 2018. *Architects of Intelligence: The truth about AI from the people building it.* s.l. : Packt Publishing, 2018.

High Court of Australia. 1923. *Russell v Wilson (1923) 33 CLR 538.* 33, 1923.

Abbott, R. 2016. I Think, Therefore I Invent: Creative Computers and the Future of Patent Law. *Boston College Law Review.* 2016, Vol. 57, 4.

Committee on Legal Affairs, European Parliament. 2020. Report on intellectual property rights for the development of artificial intelligence technologies, (2020/2015(INI)). *European Parliament.* [Online] 2020. www.europarl.europa.eu/doceo/document/A-9-2020-0176_EN.html.

Blok, P. 2017. The inventor's new tool: artificial intelligence - how does it fit in the European patent system? *European Intellectual Property Review.* 2017, Vol. 39, 2.

European Patent Office. 2020. Grounds for the EPO decision of 27 January 2020. *European Patent Office.* [Online] 2020. https://register.epo.org/application?documentId=E4B63SD6219149 8&number=EP18275163&lng=en&npl=fal.

—. 2019. Legal aspects of patenting inventions involving artificial intelligence (AI). *European Patent Office.* [Online] 2019. https://documents.epo.org/projects/babylon/eponet.nsf/0/3918F57 B010A3540C125841900280653/$File/A I_in.

UK Intellectual Property Office. 2019. Patent Decision BL O/741/19 (2019). *UK Intellectual Property Office.* [Online] 2019. https://www.ipo.gov.uk/p- challenge-decision-results/o74119.pdf.

The High Court of Appeal. 2020. *Thaler v Comptroller-General of Patents, Designs and Trade Marks [2020] EWHC 2412.* [2020] EWHC 2412. 2020.

English, R. 2021. Law Pod UK latest episode: Can AI receive patent protection? *UK Human Rights Blog.* [Online] 2021. https://ukhumanrightsblog.com/2021/10/15/law-pod-uk-latest-episode-can-ai-receive-patent-protection/.

Dhaliwal, Inder. 2020. Third strike for AI machines as patent inventors – this time USPTO says no (via Passle). *The Lens, Slaughter and May.* [Online] 2020. https://thelens.slaughterandmay.com/post/102g8b7/third-strike-for-ai- machines-as-patent-inventors-this-time-uspto-says-no.

United States District Court for the Eastern District of Virginia. 2021. *Thaler vs Lancu, et al.* s.l. : United States District Court for the Eastern District of Virginia, 2021.

Patil, Utkarsh. 2021. India: South Africa Grants A Patent With An Artificial Intelligence (AI) System As The Inventor – World's First!! *Mondaq.* [Online] October 19, 2021.

Kim, Dariur. 2020. 'AI-Generated Inventions': Time to Get the Record Straight? *GRUR International.* 2020, Vol. 69, 5.

World Economic Forum. 2018. *White Paper - Artificial Intelligence Collides with Patent Law.* 2018.

2 Crossing the Rubicon: Evaluating the Use of Artificial Intelligence in the Law and Singapore Courts

Tor Ming En

Singapore Management University, Singapore

Abstract. In recent years, Artificial Intelligence ("AI") has challenged many fundamental assumptions of how organisations and industries should operate. The Courts, traditionally seen as a hallowed ground graced by the best of lawyers, still remains as unchartered territory for AI's infiltration. Yet, there is growing evidence which suggest AI may soon cross this frontier to replace important court functions. This article critically assesses the use of AI in law and the courts. Part II will first examine the arguments for and against the adoption of AI in the legal profession. Thereafter, Part III will critically examine whether AI should replace judges in the courts. Based on the analysis, the article provides some detailed recommendations on how AI integration with the courts should be conducted in Singapore. In view of the possible threats against AI applications, Part IV provides a security and safety framework which guides Singapore courts in the adoption of AI. Against the backdrop of this article's recommendations, Part V will then discuss how automated AI judging may be done in the context of property disputes. Finally, Part VI concludes that AI integration should be readily welcome amongst legal practitioners, while AI should support instead of replacing current human judges. The implementation of AI should also be done in a calibrated, gradualist fashion. Unless AI judges can overcome their technical limitations in replicating judgecraft, AI should not be thrust into high responsibility judging roles on their own.

Introduction

AI and the Law: Overview

Before this article examines the use of AI in the courts, it is perhaps timely to evaluate the current state of AI in the law in general. Based

on the article's analysis, it is clear that AI should continue to be embraced for bringing significant benefits to the legal industry.

(a) Improving the efficiency and quality of legal services

The use of AI has allowed lawyers to streamline their workflow, leading to greater efficiency and productivity in the legal profession. Increasingly, lawyers are using AI to review due diligence documents and perform analysis of contracts in bulk (Donahue, 2018). By providing automated support to lawyers, this would reduce bottlenecks which may delay deals (Donahue, 2018).

Also, some AI tools greatly expedite the process of legal research. For instance, ROSS Intelligence streamlines time-consuming functions like writing case summaries, noting up previous precedents, or even finding cases which similarly mention a finely litigated point (Arruda, et al., 2018).

Moreover, the increased automation of legal work will increase the quality of legal services, particularly in terms of their accuracy and reduced costs.

Firstly, the use of AI may result in a higher quality of legal services, especially in the event of an anticipated litigation. For example, advanced AI applications like CaseCruncher Alpha are capable of predicting judicial decisions with a high degree of accuracy (Ashley, 2019 pp. 93, 108-109). With these insights, lawyers will be better able to advise their clients in their litigation strategy.

Secondly, an increased adoption of AI may allow clients to enjoy high quality legal services at a lower price. This is so as the cost savings of AI automation will translate into a reduction of fewer billable hours (Validatum, 2016). Such a phenomenon will be highly beneficial to individuals and smaller businesses, as they can gain greater access to affordable options for legal services (Cohen, 2016).

(b) Automation bias

Detractors of AI contend that AI might result in automation bias within the legal profession. By definition, "automation bias" refers to an impulse to accept a computer's recommendation with excessive faith, without processing the algorithmic output vigilantly (Citron, 2008 pp. 1249, 1271). They further proffer that the reliance of AI would result in a deterioration of legal reasoning in society, reminiscent of how GPS navigation eroded people's ability to find their directions on their own (Michaels, 2020 pp. 1083, 1088).

However, there is currently no strong empirical evidence in the machine-learning context to substantiate the concerns of automation bias (Huq, 2020 pp. 611, 682). Furthermore, it is important to note that legal AI is mostly used to *support* a lawyer's work functions. As such, a lawyer will still have to exercise one's independent legal judgment to decide whether an AI's recommendations should be incorporated.

Hence, the concerns over "automation bias" is not fatal enough to resist the adoption of AI tools in law.

(c) Technological displacement of lawyers

The most vociferous objections against AI adoption lies in the concerns that legal professionals would be displaced by technology. This stems from the fact that some legal practices with more routine job functions are highly susceptible to automation (Lin, 2019).

Despite such well-founded concerns, it is unlikely that AI can completely replace a human lawyer. Other than processing legal information and making predictions, AI machines lack the capability to negotiate with clients (Lohr, 2017) and providing commercially sensible legal advice.

Moreover, empirical evidence suggests that the advent of legal AI merely translates into a future reduction of lawyers' working hours (Lohr, 2017). It does not definitively conclude that AI adoption would necessarily lead to the future unemployment of lawyers.

In any event, the strong concerns over a lawyer's job security should not militate against the adoption of beneficial AI technologies. As the experience of AI adoption in other industries have shown, the onus should lie on lawyers to develop new skills and adapt to the prospective changes accordingly.

AI and the Courts

In many countries, AI has been incorporated to serve auxiliary judicial functions. This includes how U.S employs the COMPAS AI to determine a defendant's risk of recidivism, which would then inform human judges' decisions about bail and sentencing (Tashea, 2017). Meanwhile, Mexico uses the Expertius system to provide advisory opinions on whether an individual is entitled to a financial pension (EXPERTIUS: A Mexican Judicial Decision-Support System in the Field of Family Law, 2008).

Increasingly, countries have begun to challenge such a paradigm by allowing AI to be an arbiter in disputes. Since 2017, China's Hangzhou Internet Court uses AI judges to adjudicate online trade disputes, copyright cases and e-commerce product liability claims (Ito, 2019). This has since been replicated in other states like Beijing and Guangzhou (Ito, 2019). Similarly, Estonia also began employing robot judges to adjudicate small claims (Tangermann, 2019).

Presently, the integration of AI into the courts has not gained traction in Singapore yet. Despite its enumerated risks, Singapore should seek to delegate some important court functions to AI — particularly the adjudicative role of human judges — *as long as there is a robust system of checks in place.*

(1) Evaluation of using AI judges in the Courts

(a) AI's efficiency and ability to reduce human bias

The use of AI would increase the efficiency of the courts. Unlike human judges, an AI's algorithmic decision procedure can be used to adjudicate vast number of cases in a highly *efficient* manner (Solow-Niederman, et al., 2019 pp. 242, 255-256). This may also lead to other spill-over effects, such as increased cost-effectiveness (Volokh, 2019 pp. 1135, 1139) arising from greater digital economies of scale (Wu, 2019 pp. 2001, 2002).

By creating some form of standardisation in the adjudication process, AI judges also aspire to reduce arbitrariness — which may stem from many factors like unconscious assumptions or even decision fatigue (Sourdin, 2018 pp. 1114, 1129) — and some forms of cognitive biases inherent in human judges' intuition (Sourdin, 2018; Kahan, 2013).

While the use of AI ineluctably increases courts' efficiency in resolving their heavy caseloads, it is doubtful whether AI judges can completely eliminate bias from judicial decision-making *in practice*.

In particular, machine learning AI systems inevitably suffer from learning bias due to its path dependent background. Bias will inevitably be normatively introduced by human design and engineering choices at any point (Kelleher, et al.) — whether in performing an evaluative selection of an algorithmic architecture (Huq, 2020 p. 646) (whether a neural network or a decision tree-based model), network topology, (Huq, 2020 p. 647) or in the selection of training data.

One key challenge of AI is the quality of the data upon which AI systems rely. As a product of existing social structures, the training data replicates the biases and blind spots of individuals who have curated it (Huq, 2020). In the U.S., the assessment of risk of recidivism under COMPAS' sentencing algorithms are found to contain internal

biases against black offenders, as the AI was trained on crime data reflecting the preexisting racial bias in law enforcement (Thompson, 2019; Angwin, et al., 2016).

In such situations, it becomes highly important for AI engineers to perform counterfactual fairness testing regularly (Personal Data Protection Commission of Singapore, 2020), so that any inherent bias within the system can be duly identified and addressed. This would prevent the discrimination based on sensitive attributes like race and gender (Kusner, et al., 2017).

Nonetheless, studies have found that such detected bias in decision-making can be reduced by auditing its outcomes and setting appropriate constraints (Kleinberg, et al., 2017 pp. 237, 275-278). This is important as it underscores a key advantage that AI has over human judges, which entails how algorithmic bias can detected and corrected more easily than human bias (Casey, et al., 2019 pp. 333, 352).

Aside from the issue of inherent bias, a machine learning AI system might make inaccurate decisions due to statistical limitations. For instance, AI judges would usually apply a high variance model with many legal parameters (Singh, 2018), which reflects the multiplicity of legal rules. By basing the AI decision-making very closely on a limited set of training data, this would result in over-fitting of available training data when applied to more novel scenarios (Solow-Niederman, et al., 2019 p. 271).

This issue, however, can easily be mitigated in part by integrating rule-based expert systems into the AI's machine learning system. By encoding specific rules as code into the AI algorithm, it can help to fine-tune areas of legal reasoning with lower precision or recall, or even areas which has yet to be successfully trained (Hybrid Approach Combining Machine Learning and a Rule-Based Expert System for Text Categorisation, 2011 p. 328).

(b) Concerns over transparency of AI decisions

At present, machine learning AI systems can produce outcomes without being able to provide a proper explanation for human comprehension.

For machine learning AI systems to attain their desired ends, deep learning functions help to draw mass correlations within data to infer complex statistical patterns. Unlike linear regression models, machine learning models use indeterminable weights for each feature that are used to make a prediction (Hacker, et al., 2020 pp. 415, 417). Therefore, the workings of a machine learning model often lie beyond conventional human understanding (Casey, et al., 2019 p. 355), since it does not involve any logical reasoning or causal inferences (Solow-Niederman, et al., 2019 p. 263; Huq, 2020).

While an AI product can deliver automated *post hoc* rationalisations, they may not reflect the AI's true decision-making process (Solow-Niederman, et al., 2019 p. 261). For example, instead of providing direct reasons for its decision, an AI judge may use counterfactual explanations (Buocz, 2018 pp. 41, 49) like "*if* X had been done, *then* you would have attained outcome Y as you hoped".

It appears that AI critics are chiefly concerned about such a degree of incomprehensibility of AI decisions. They seem to accept how human judges leave room for uncertainty on particular issues for future reinterpretations (Re, 2014 pp. 1861, 1891), but object to how the *process* of obtaining AI outputs is *completely* incomprehensible.

In the critics' view, the use of AI judges may compromise one's procedural right to due process (Citron, 2008 p. 1252) to understand the operation of the law. It may also create intractable technical challenges for the judiciary to maintain oversight or intervene in an AI's decision processes (Solow-Niederman, et al., 2019 p. 266).

However, these concerns over AI's "black box" reasoning is rather misdirected as it is based on a false comparison (Casey, et al., 2019 p.

355). It is pointed out that a human brain also functions like a black box, where judges' written opinions may actually be *ex post* justifications for other underlying factors (Casey, et al., 2019 p. 356).

If anything, it may well be possible that the critics' argument against AI's black box reasoning is merely a way to bring in their emotional preference for human interaction in court processes (Huq, 2020 p. 655). Given that humans generally grow accustomed to the use of AI after initial resistance, the arguments against the AI's incomprehensibility will likely lose their vehement force sometime after its adoption (Wu, 2019; Volokh, 2019).

Therefore, rather than focusing on the transparency of computer reasoning, it is more appropriate for AI to be assessed based on how well the AI system achieves its intended objective — whether in terms of performing rule application or fact finding.

(c) Effects of codified justice on judge's discretion

The proposed use of AI has also been opposed for its potential implications on judicial discretion, which is often viewed as tool to ensure fairness (Lai, 2019).

Critics of AI are of the view that AI adjudication would not only update legal rules, but also inexorably alter underlying *values* of equitable justice. Traditionally, equitable justice aspires to apply consistent principles in a contextual manner, while necessarily allowing for minor deviations from principles in exceptional circumstances (Solow-Niederman, et al., 2019 p. 253). As such, equitable justice allows judges to weigh relevant mitigating factors meticulously before deciding on a just and appropriate sentence for an offender.

A shift towards codified justice would promote the standardisation of all legally relevant variables in advance, in view of promoting efficiency and consistency (Solow-Niederman, et al., 2019 p. 254). As a by-product, this may arguably result in reduced spaces for judicial

discretion to consider other fact-sensitive factors in individualised proceedings, leading to a mechanical application of the law. Critics of AI also contend that it is disempowering for litigants to submit to an external authority (Michaels, 2020 p. 1097) (Solow-Niederman, et al., 2019 p. 276) as they are no longer able to persuade judges to exercise their discretion.

Such concerns are oversimplistic, as AI judges do not necessarily result in a shift away from equitable justice.

By using a rich data set in training the AI model, an AI judge might even be able to identify more granular distinctions and better calibrate decisions to individual case facts (Solow-Niederman, et al., 2019 p. 259). For instance, the algorithm can parse many mitigating and aggravating factors to deliver a sentence that is highly calibrated to the offender's circumstances.

Furthermore, equitable considerations can be encoded to form the AI's intended end objective. Volokh points out that it is possible to ensure AI judges are not only legally correct, but also compassionate. This can be done by including compassion as one of the criteria when a panel of legally trained evaluators tests the AI machine at the preliminary stage (Volokh, 2019 p. 1167).

On a final note, it is also important to note that people may fundamentally disagree on the utility of discretion across different issues. In particular, many Singaporean citizens were aggrieved by the State Courts' recent decision to put a young sexual offender under probation on account of his "potential to excel in life" (Sun, 2019; Zula, 2020). As a judge's discretion might be perceived to be arbitrary at times, the use of AI judges would be timely to ensure discretion, if any, is kept within acceptable bounds.

(d) Perceived inability to apply the law in complex cases

Besides the concerns of codified justice, the ability of AI judges to apply the law has also been questioned.

At first blush, it is inconceivable how AI systems can apply relatively vague legal standards like "reasonableness" to the facts. It is also difficult to envisage how an AI judge can reconcile different pieces of evidence and draw relevant inferences accordingly. However, since AI uses a data-driven paradigm which is completely different from human reasoning, its ability to apply the law should be judged in terms of its *outcome* (Volokh, 2019 pp. 1140-41) rather than based on its reasoning process.

In terms of applying the law, it appears that AI systems actually outperform human judges in issuing accurate predictive decisions (Huq, 2020 p. 654) (Sourdin, 2018 p. 1125). Within the context of pre-trial bail and domestic violence arraignments, empirical evidence suggests that machine learning tools generate fewer false positives and negatives at the population level as compared to most human decision-making (Huq, 2020 pp. 654-655).

However, AI systems face difficulties in applying the law to "hard cases". When a novel case does not share pertinent features with prior analysed cases, the AI's algorithmic logic as inferred from the training set cannot be extrapolated to make accurate predictions (Surden, 2014 pp. 87, 105). It is further contended that "hard cases" necessitate a fresh evaluation of circumstances via a judge's discretion, which lies beyond the capability of machine learning AI systems (Comes, et al., 2018).

Additionally, despite the rise of Natural Language Processing ("**NLP**"), there are still concerns that AI judges may not appreciate different text nuances in different contexts (Wu, 2019 p. 2024). For instance, an AI judge might interpret "I am going to kill X" erroneously as a death threat, even though it might be articulated as a figurative statement in reality (Wu, 2019 p. 2024).

AI machines also face similar difficulties in applying the law when legal rules contradict each other (Wu, 2019 p. 2003). This problem is less fatal, as it can be mitigated by applying complex conflict resolution techniques within the machine learning system (Methods for Rule Conflict Resolution, 2004).

(e) Perceived inability to perform finding of fact

Besides AI's limitations in applying the law to hard cases, critics also contended that AI judges are virtually incapable of replacing a human judge's fact-finding capabilities.
The process of fact-finding often involves decisions which include but are not limited to: whether an evidence is probative of the existence of a fact in issue (Bell, 2013 pp. 519, 521), whether a witness' evidence is reliable and consistent (Bell, 2013 pp. 524-28), what inferences can be drawn on the facts (Bell, 2013 p. 540), and whether the fact in issue is proved to the requisite standard of proof on a whole (Bell, 2013 p. 546).

Some of the questions can be answered in light of current advancements in technology. With the developments in affective technologies, AI may soon be able to interpret human emotions (Comes, et al., 2018 pp. 59, 97) to help assess witnesses' credibility. In addition, an AI algorithm can parse many factors, which can be used by AI to ascertain if a piece of evidence is reliable and thereby admissible under the Evidence Act.

While some types of fact-finding are suited for mechanisation, there are nonetheless aspects of fact-finding which are best left to human judges (Solow-Niederman, et al., 2019 p. 283). Notably, questions like "whether adverse inference should be drawn" involve the considerable use of judicial discretion (Gennaioli, et al., 2008 pp. 1-2), which lies beyond the scope of machine learning (Comes, et al., 2018 p. 100).

(f) Perceived inability to develop the law

Another objection mounted towards AI judges lies in how it stifles the dynamic process where legal rules and standards are updated. This is perhaps the strongest argument which critics of AI can put forth.

Unlike human judges, AI cannot draw upon their foresight and lived experiences in society to adapt existing legal rules. Even if a "programme of discretion" can be regularly updated into the AI to reflect a changed social or legal consensus (Solow-Niederman, et al., 2019 p. 280), an AI system still lacks a human judge's rights-driven moral reasoning (Wu, 2019 p. 2025). This makes it difficult for an AI to develop or reformulate the law in groundbreaking fashion like Singapore Court of Appeal's *Spandeck* decision (Supreme Court of Singapore - Court of Appeal pp. [73]-[86]).

(2) Recommendations for AI integration with Singapore courts

In summary, the enumerated advantages of AI systems should justify its inclusion within the courts, albeit via a carefully calibrated approach. The courts should start off by adopting a hybrid machine-human model, thereby allowing them to streamline their workflow while mitigating all AI-related concerns.

While it is empirically shown that human involvement as backstop to AI algorithm decreases accuracy of prediction (Huq, 2020 pp. 665-67), this is necessary in light of AI's limitations. As mentioned earlier, AI systems are still afflicted with a limited ability to apply the law to "hard cases", perform judicial fact-finding and develop the law when necessary.

Nevertheless, there is an immediate scope for in-house AI to be applied in various *lower-level* support functions. *At every stage*, it is important for the judge to critically evaluate the AI's generated work instead of processing it into a trusted final decision immediately. This will guard the judiciary against concerns over possible automation bias (Citron, 2008 p. 1272).

An AI system can assist a judge's input request to search for requested literature, and expediently return the searches via ranked retrieval results (Buocz, 2018 pp. 40, 50). The AI system may go even further by drawing links between research materials and structure them into a template opinion (Sourdin, 2018 p. 1124). The AI's algorithmic output can then provide an advisory role to support the human judges' deliberation process.

Once the AI system is capable of producing high quality briefs and opinions *consistently*, the Attorney-General's Chambers ("**AGC**") may allow AI to replace low-level judging roles. It may be allowed to perform unassisted adjudications at the State Courts, when the impact of a potentially defective judgment is not as serious.

After the AI is sufficiently tested and calibrated to deal with more complex issues, it may gradually be elevated to a judging role at the High Court.

In line with the finding in **Part III(1)(c)**, AI should be designated to adjudicate on disputes which favour a rigid and mechanical application of the law, such as issues of tax law. For criminal or administrative law cases, AI may nonetheless be used to provide a decision on specific sub-issues as long as it lies within the capabilities of an AI system.

There are two possible ways in which an AI judge can be used at trial. For one, a human judge may be included in the AI's decision-making loop, such that all algorithmic outputs come with a human judge's affirmation. Such a "human-in-the-loop" approach is justified in light of the possibly high severity of an incorrect decision (Personal Data Protection Commission of Singapore, 2020 p. 33).

Alternatively, both the human judge and AI system can be included together in an expanded coram of two judges. In both circumstances, the AI machine's decision or fact-finding may be reviewed and even overturned on appeal to the apex court.

In any event, the entry of AI systems into the courts is likely to be a gradual process. Given that AI mostly excel in their individual tasks separately, it would take time before AI can truly replicate the entire range of skills possessed by a human judge (Buocz, 2018 pp. 40, 46) (Solow-Niederman, et al., 2019 pp. 285-287).

AI Safety and Security Measures

Before automated judging is implemented, the AI system should incorporate robust safety and security countermeasures. By adopting a multilayered system of protection, this ensures that the AI's network and data security is better protected against many types of attack vectors.

(a) Threat modelling, penetration testing and attack response plans

The *overarching* security framework will first require AI engineers to perform threat modelling. During this process, AI engineers must identify the potential threats to the AI system, and possible structural vulnerabilities at each point where data is stored and transferred (Myagar, et al., 2005). Thereafter, the AI system should be fitted with necessary security mechanisms to mitigate the identified threats.

In addition, AI engineers should proceed to conduct penetration testing, which simulates an attack by an unauthorised user on the AI system. The tests would indicate whether the implemented security measures have been effective (Bacudio, et al., 2011 pp. 19, 20). If the security infrastructure was ineffective in combating an attack vector, it should be fine-tuned and improved accordingly.

Beyond these measures, it is also important to formulate attack response plans (Comiter, 2019). This ensures that parties can react quickly to control or mitigate any damage in the event of a malicious attack.

(b) Transport Layer Security protocol for online information retrieval

One of the proposed security measures involves the use of cryptographic encryption protocols like Transport Layer Security ("**TLS**"). As the AI system might need to retrieve foreign literature from online legal databases like Lexis, the TLS system would ensure the security of the data *in transit*. This is necessary to ensure the integrity and confidentiality of the AI's retrieved data (Parmar, et al., 2015 pp. 35, 36).

For instance, the AI judge would first send a hello message to the Lexis server, before Lexis sends over its public key certificate (Das, et al., 2014 pp. 68, 70). The AI system uses Lexis' digital certificate for server authentication based on its chain of trust, so as to ensure that the AI system is not interacting with a spoof website (Grigorik). Once done, the session keys would be generated using a key exchange algorithm (based on the RSA (Rivest Shamir Adleman) or Diffie-Hellman paradigm), which would be used to encrypt and decrypt the transmitted data.

With the TLS protocol's encryption and authentication functions in place, the TLS protocol also employs a "Record" sub-protocol to protect the data from being tampered with. For each TLS record, this is done by generating and appending a pseudo-random message authentication code (MAC), (Das, et al., 2014 pp. 85, 70) which is essentially a one-way cryptographic hash function. The MAC checksum can be used to verify the *integrity* of the transferred data, such as the foreign literature sent from the Lexis server to the AI system.

In the event that there is unauthorized alteration of transmitted data by third parties, the discrepancy in the hash values would be flagged out via a fatal error message to the AI system (Sarikaya, et al., 2011). The session would accordingly be terminated to halt the transmission of tampered data.

Although cryptography offers high level of security for data transmission, the AI system should take additional steps to guard against evolving cryptographic attacks. As seen in **Figure 1**, the Man-In-the-Middle attack is still capable of targeting specific TLS vulnerabilities, which allows an attacker to eavesdrop on the data transmission elusively (Mallik, et al., 2019 pp. 77, 78) or even impersonate (Mallik, et al., 2019 p. 81) the Lexis server.

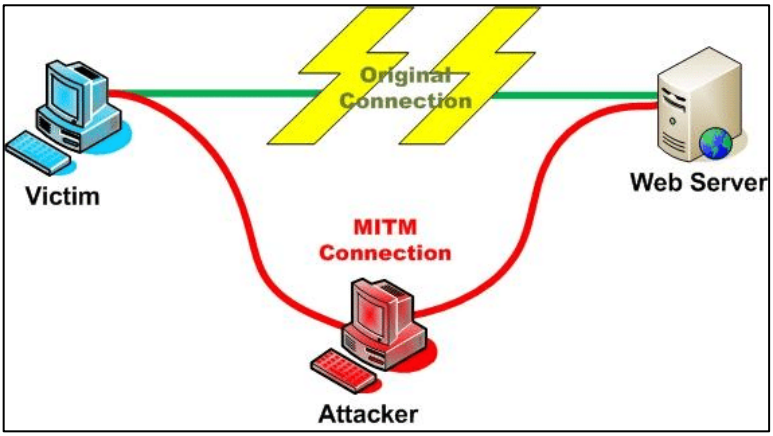


Figure 1: Diagram of a Man-In-the-Middle (MITM) attack on a data channel (Parmar, et al., 2015 p. 37)

Therefore, the AI system may fortify its defences to make it more difficult for attackers to intercept and decrypt messages without the appropriate keys. This includes the adoption of Geffe generation to yield highly random binary sequences (Kahder, et al., 2015). Alternatively, less secure key generation methods, such as the RSA model, should be replaced by more robust paradigms like the Elliptic Curve Diffie-Hellman key exchange model (Hema, et al., 2018) (Ronen, et al., 2019 p. 13).

(c) Firewall protections for AI's cloud database

During their operation, it is highly likely that multiple AI judges will be connected to a cloud database, which would store a repository of statutes, case precedents and literature. To prevent the integrity of data *within the database* from being compromised, strong firewall protections should be established within the internal cloud network.

As shown by **Figure 2**, a firewall system only allows authorised traffic to access the cloud database, by protecting the points of entry into the network (Abie, 2000). This is achieved by applying a set of rules to inspect the contents of moving network packets, which entail an examination into its protocol type, destination IP address and its source IP address (Pundar, et al., 2014 pp. 841, 842).

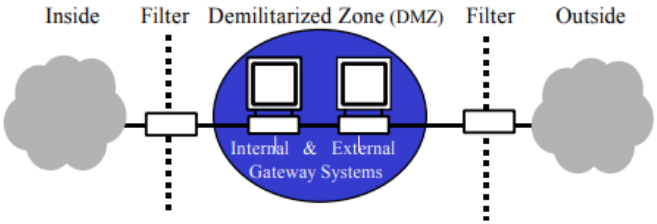


Figure 2: Diagrammatic representation of a firewall (Abie, 2000 p. 2)

The firewall infrastructure must be supported by robust Intrusion Detection Systems ("**IDS**") and Intrusion Prevention Systems ("**IPS**").

The IDS and IPS tools can identify cybersecurity threats (like malware and port scanners) by comparing their network activities to a known threat signature database, and thereby deny network traffic to such network packages (Rao, et al., 2014). To increase their effectiveness in filtering out cybersecurity threats, the threat signature database should also be updated regularly from other anti-virus labs and security providers (Rao, et al., 2014 p. 230). Hence, by strengthening the cloud database's first line of defence, these tools can allow for enhanced threat management.

Nonetheless, it is important to note that a firewall is a form of perimeter defence (Abie, 2000 p. 4). Hence, a firewall is unable to prevent abuse of authorised access from within, especially by a disgruntled AGC employee with access privileges.

Such a problem may be mitigated by the use of network segmentation (Jackson, 2018) within the internal cloud database. As such, employees will only possess access privileges to sub-networks which are relevant to their job responsibilities. By restricting the impact of any network intrusions or abuse to an isolated sub-network, this ensures that the AI's database network is protected from any third-party tampering.

(d) Guarding against data poisoning attacks in a machine learning model

As AI systems are expected to manage many high-impact cases, hackers would ineluctably be keen to capitalise on AI vulnerabilities to inflict maximum damage on the courts (Volokh, 2019 p. 1174). This can be achieved by targeting either the AI's training dataset, learning algorithm or the AI model itself (Comiter, 2019 p. 30).

In particular, attackers would alter the training data set to prevent the AI model from learning specific patterns (Comiter, 2019 p. 13). Such data poisoning — which is often pernicious and hard to detect (Comiter, 2019 p. 9) — would cause the training algorithm to lose the ability to discern noise and anomalies from high-confidence data (Marshall, et al., 2018). As a result, this would prevent the sound application of targeted legal rules, which may lead to unjust outcomes and decisions. Furthermore, attackers may even install secret backdoors, which may be repeatedly accessed in future to trick the AI system (Comiter, 2019 p. 13).

Hence, it is critical for machine learning AI systems to identify and reject maliciously introduced training data that negatively affects their algorithmic output (Marshall, et al., 2018 p. 106).

This may include internal security mechanisms like the Reject on Negative Impact ("**RONI**") defence. This function measures the empirical effect of each training instance and eliminates data sets which have a substantial negative impact on an AI's classification accuracy (Barreno, et al., 2010 pp. 121, 137).

Other variants of the RONI defence may involve an audit on the accuracy of reconstructed models based on the user's contribution of training data (Ying, et al., 2020) (Collinge, et al., 2019). The system will then reject inaccurate training data which are likely provided by attackers.

Besides the use of internal defence mechanisms, the AI security infrastructure can be enhanced by using other available technologies. Proprietary systems like Darktrace can monitor the behavioural patterns of an AI system. In the event of an attack, it can detect any anomalies to the AI's "normal" state of affairs in real-time (Darktrace, 2019). Immediately after, the cyber defence system would automatically block the suspicious connections to return the AI system back to its normal state (Darktrace, 2017).

In order to guard against data poisoning attacks, all data within the AI's database— whether at rest or in motion — should also be encrypted. Without the securely protected decryption key, the attackers would be unable to reverse engineer the model with the data, thereby making it difficult for them to launch data poisoning attacks (Comiter, 2019 p. 66).

(e) Resolution of AI's internal weaknesses and glitches

Aside from enacting countermeasures against unknown attack vectors, the inherent weaknesses and glitches within the AI system should be duly addressed.

The AI system might exhibit undesirable "emergent properties" in the course of learning from the available training data (Volokh, 2019 p.

1172). Volokh cites an example of how an AI judge may end up relying on irrelevant facts, just because it uses particular words in a fact pattern that are somehow correlated to successful litigation outcomes (Volokh, 2019).

Since it is difficult for AI to identify bias in its data sets, it is important that AI systems can be reviewed and audited by humans (Satell, et al., 2019) (Marshall, et al., 2018 p. 106). This will facilitate the process of debugging and fine-tuning glitches found within the AI model.

As such, the AI system should include an in-built forensics and security logging function. Firstly, it enables human experts to trace specific events, which are recorded as non-repudiable evidence in the algorithm metadata (Marshall, et al., 2018 p. 106). Also, it allows for a close examination into the state of specific classifiers which contributed to an inaccurate decision (Marshall, et al., 2018). Beyond the detection of glitches, the algorithm can also spot existing hacks and hidden backdoors within the AI system (Volokh, 2019 p. 1175) (in relation to **Part IV(d)**).

After the faulty nodes within the AI system are detected and isolated, the AI system can then be fine-tuned accordingly through targeted updates.

(f) Development of AI system by the statutory agencies

On a final note, the development, testing and maintenance of the AI judge should be spearheaded by a public agency like GovTech Singapore.

The main concern of relying on a private company stems from their possible conflicts of interest. There are strong concerns that a company may create a backdoor within the AI system, and seek to sell such an access for profit (Volokh, 2019 p. 1172). Furthermore, if the private company is involved in high-stakes litigation, it is feared that

they might manipulate the AI algorithm for a favourable outcome (Volokh, 2019).

Potential Areas of Application: Property Disputes

Despite the benefits of adopting AI judges, **Part III(2)** earlier recognised that AI's technical limitations may militate against its adoption until the distant future.

Against this background, this article seeks to explore how automated judging might then be applied to future disputes, specifically in the context of property disputes.

(1) Broad structure of AI adjudication as a background

Before exploring how automated judging can apply to property disputes, this section lays out the general process of AI adjudication as a background.

During the process of adjudication, an AI machine will process the evidence from both litigants to establish the facts of the dispute. In cases involving complex fact-finding processes, the courts should instead assign a human judge to perform more sophisticated tasks — like reconciling conflicting version of the facts or drawing inferences from relevant circumstantial evidence.

The fact finder will then funnel its algorithmic output to a separate AI judge, which would also be present in the proceedings. The main function of the separate AI judge is solely to apply the law to relevant facts. During the final adjournment of proceedings, the AI judge would then produce a plain-text judgment for perusal by the human judge in its decision-making loop.

Thereafter, the human judge will review the AI's written decision and verify if there are any misapplication of the law. At this time, a human

judge may also perform substantive editing to reformulate or develop a legal rule, especially if the case is deemed to be a novel case.

Upon the completion of this review, the edited text shall represent the final decision on that issue, which would be sent to the litigants' counsels. Using a combination of natural language processing and advanced text-to-speech models like Wavenet (Rothman, 2017), the AI judge can also generate and read out a summary of the decision in the final hearing.

Through a bifurcation of responsibilities, each AI machine can perform its specialised function more efficiently. The division of labour across multiple AI machines also facilitates a quicker troubleshooting process to resolve unexpected problems during their operation.

Moreover, incorporating a "human-in-the-loop" also greatly reduces the possibility of causing a miscarriage of justice. This is especially important if the AI judge's algorithmic output yields a false positive or negative outcome initially.

(2) Application of multi-factorial tests to reach a conclusion

For many property disputes, the legal test in question involves multiple conjunctive factors, where the conclusion is determined by the *presence or absence of clear factors*. Some examples include: tests to determine whether a contract for sale of land is enforceable, whether an object is a fixture, or whether formalities are satisfied (such as for a caveat application or lease of registered land).

As rule-based expert systems can be integrated to optimise an AI's artificial neural network, (Hybrid Approach Combining Machine Learning and a Rule-Based Expert System for Text Categorisation, 2011 p. 328) (Youngcho, et al., 1994 pp. 497, 499-501) it is possible for the legal conditions to be encoded into the AI's algorithm via "IF-THEN" conditions.

The AI judge can then use a straightforward forward chaining method (Grosan, et al., 2011 p. 158) to test if the relevant factors are satisfied before coming to a conclusion (**Figure 3**).

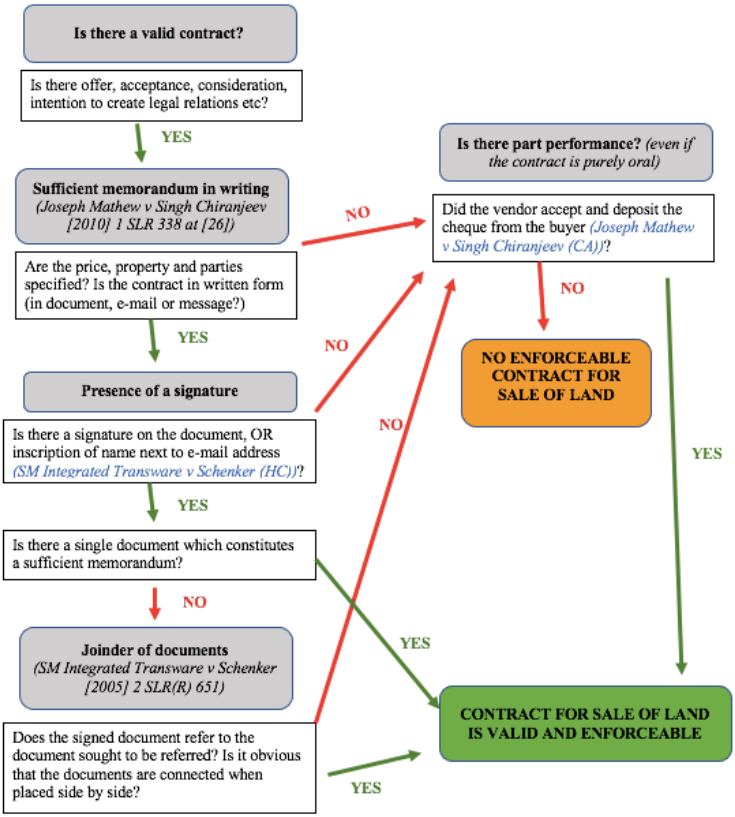


Figure 3: Rule-based logic for AI judge to decide if a contract for sale of land is enforceable

To decide if ownership in an object is passed with the land, courts often have to deliberate if the object is a fixture or a mere chattel. By following the logic in **Figure 4**, a rule-based expert system within an

artificial neural network can allow the AI judge to reach a conclusion on that issue expediently.

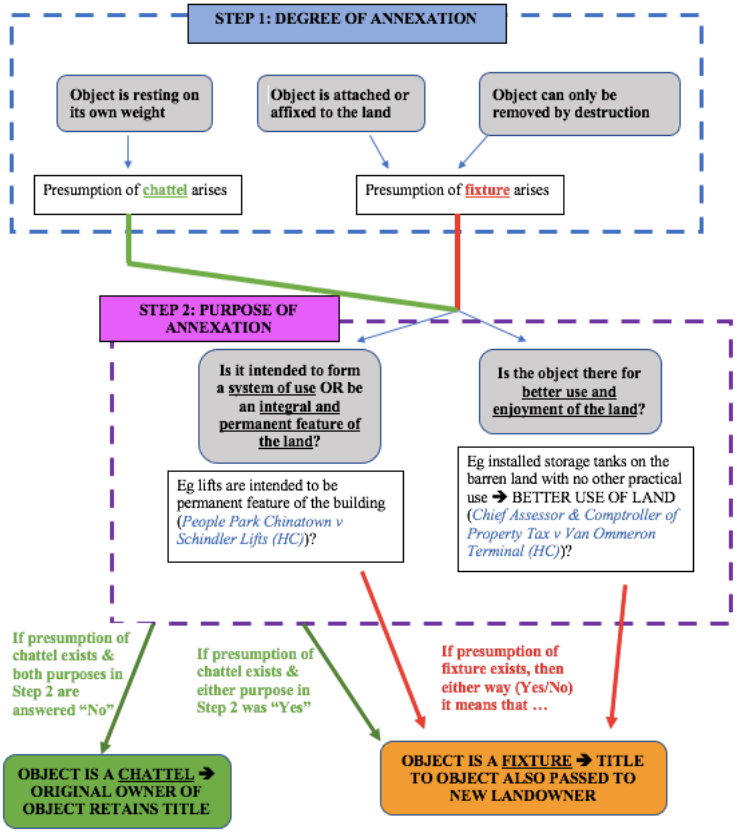


Figure 4: Rule-based logic for AI judge to decide if title in object is passed to new landowner

However, in other areas of property law, AI judges must reach a conclusion based on the *degree* in which each factor is satisfied in the test. Instead of having factors which can be answered in a "Yes/No" fashion, many tests have factors which are more *open-ended* in nature. For instance, whether parties share a close relationship (Supreme

Court of Singapore - Court of Appeal, 2008 pp. [140]-[141]) involves a level of subjectivity which goes beyond a simple Boolean inquiry, especially since there are different gradations of closeness in a relationship.

Nonetheless, AI judges can achieve a conclusion with the inclusion of fuzzy logic atop rule-based expert systems to support the use of interpolative reasoning (Grosan, et al., 2011 p. 424). By having probabilistic figures to determine if the elements in the test is met (see **Figure 5**), this better accords with how human judges answer more open-ended questions of law. At the same time, it provides a mathematical justification to the proof of the fact in issue on a balance of probabilities.

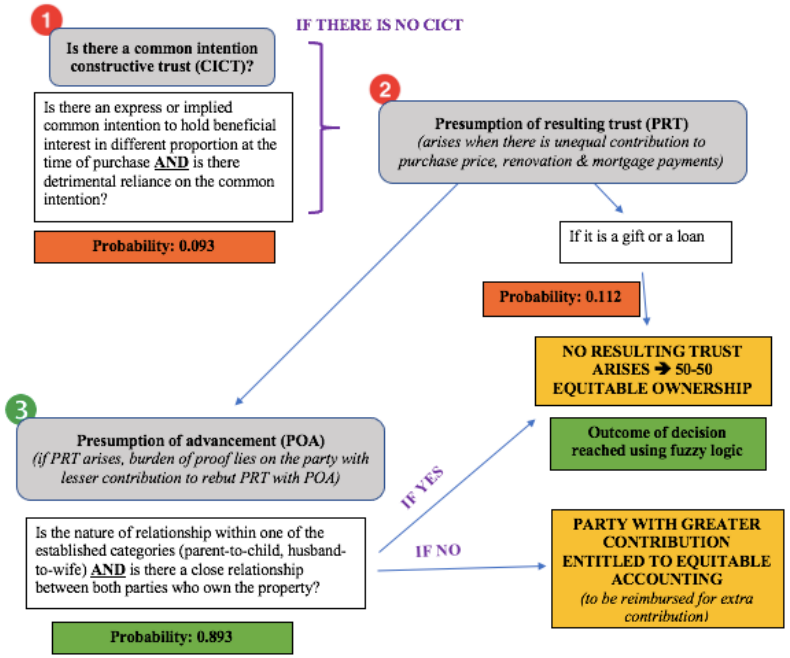


Figure 5: Diagram of how Lau Siew Kim (Supreme Court of Singapore - Court of Appeal, 2008) *might be decided by AI judge with fuzzy logic*

Once done, the fuzzy rule-based expert system's output can serve as a pre-processor (Grosan, et al., 2011 p. 426) to the AI's artificial neural network for the determination of other complex issues.

(3) Application of vague legal standards to reach a conclusion

In some instances, the AI judge must determine if vague legal standards — whether in common law or statute — are satisfied on the facts. For instance, Section 46(2)(a) LTA (Land Titles Act of 2004 in Singapore) provides that the title of a proprietor of registered land is defeasible on grounds of fraud.

As a result of the machine learning model, an AI judge should be able to determine if an action falls under the common law definition of fraud. With NLP functions, the neural network can be trained on different precedents. Through this process, the AI system will learn to classify whether a registered proprietor's mental state crosses the threshold of "fraud" on a graduated scale.

In particular, it should identify Torrens fraud only in cases of extreme dishonesty like *Loke Yew* (Privy Council, 1913). Through the training data set, the AI model should be able to make granular distinctions, by noting that "fraud" must necessarily result in the transfer of title per se. Any subsequent unconscionability after the transfer of title does not constitute Torrens fraud by itself. Additionally, following *Bebe* (Supreme Court of Singapore - Court of Appeal, 2006 p. [22]) and *Malayan Banking*, (High Court of Singapore, 2008 p. [40]) the AI judge should also note that common law negligence does not amount to Torrens fraud under Section 46(2)(a) LTA.

Hence, as long there are enough foreign precedents to train the model, the AI model should possess a higher classification accuracy and lower variance (Junghwan, et al., 2016). Accordingly, it can then make a fairly accurate decision on whether a vague legal standard is engaged.

(4) Determination of remedies

An AI judge is also equipped with the ability to determine the appropriate remedies across different types of property disputes. For instance, it should be able to easily determine the amount of compensation to the other joint owner after equitable accounting, akin to the Court of Appeal's decision in *Su Emmanuel* (Supreme Court of Singapore - Court of Appeal, 2016).

Arguably, AI machines can also determine whether specific performance should be granted in a property dispute. For specific performance to be granted instead of damages, it involves the weighing of several considerations like: whether specific performance might cause undue hardship to one party (High Court of Singapore, 2010 p. [33]) or a third party (High Court of Singapore, 2011 p. [114]), and whether damages might be an adequate remedy (High Court of Singapore, 2011 p. [111]).

As **Part III(1)(c)** alluded, these factors can also be parsed as "factors" into the AI's algorithm. Depending on the factual matrix, each of the "factors" is then assigned a unique weighted value (Grosan, et al., 2011 p. 287). Collectively, these weighted values may lead the AI judge to grant either an order for specific performance or damages.

Conclusion

More than ever, AI technology should be embraced by the law and courts in light of its enumerated benefits. Such a process should be done in a gradual and experimental fashion, so as to ensure minimum disruption to the judiciary.

As a start, AI can provide dedicated support to human judges by assisting in legal research, or even in crafting draft opinions. In the distant future, AI may play a more active role in the adjudication of disputes. However, due to the high severity of an erroneous judgment,

human judges should nonetheless be involved in an AI's decision-making loop as the last line of defence.

On a final note, it is unclear whether AI can ever fully replicate a human judge's ability to apply the law, develop the law or perform fact-finding. Despite this, AI still has a big role to play in spearheading the future trajectory of the courts. In light of this reality, it is high time for the legal fraternity to stop resisting change and start welcoming AI adoption with open arms.

References

Donahue, Lauri. 2018. A Primer on Using Artificial Intelligence in the Legal Profession. *Harvard Journal of Law & Technology Digest.* [Online] 2018. https://jolt.law.harvard.edu/digest/a-primer-on-using-artificial-intelligence-in-the-legal-profession.

Arruda, Andrew and Scherer, Matt. 2018. Artificial Intelligence: Will It Replace Lawyers? *ACLEA 54th Annual Meeting.* 2018.

Ashley, Kevin. 2019. A Brief History of the Changing Roles of Case Prediction in AI and Law. *Law in Context.* 2019, Vol. 36.

Validatum. 2016. The Perfect Pricing Storm: Artificial Intelligence and Hourly Billing. *Validatum.* [Online] 2016. https://validatum.com/articles/the-perfect-pricing-storm-artificial-intelligence-and-hourly-billing.

Cohen, Mark. 2016. How Artificial Intelligence Will Transform the Delivery of Legal Services. *Forbes.* [Online] 2016. https://www.forbes.com/sites/markcohen1/2016/09/06/artificial-intelligence-and-legal-delivery/?sh=68ea402922cd.

Citron, Danielle. 2008. Technological Due Process. *Washington University Law Review.* 2008, Vol. 85.

Michaels, Andrew. 2020. Artificial Intelligence, Legal Change and Separation of Powers. *University of Cincinnati Law Review* . 2020, Vol. 88.

Huq, Aziz. 2020. A Right to a Human Decision. *Virginia Law Review.* 2020, Vol. 106.

Lin, Suling. 2019. Technology can displace lawyers, warns Chief Justice as he urges profession to adapt to a new reality. *Channel News Asia.* [Online] 2019. https://www.channelnewsasia.com/news/singapore/lawyers-legal-industry-tech-disruption-chief-justice-menon-11846.

Lohr, Steve. 2017. AI is Doing Legal Work. But It Won't Replace Lawyers Yet. *The New York Times.* [Online] 2017. https://www.nytimes.com/2017/03/19/technology/lawyers-artificial-intelligence.html.

Tashea, Jason. 2017. Courts Are Using AI to Sentence Criminals. That Must Stop Now. *Wired.* [Online] 2017. https://www.wired.com/2017/04/courts-using-ai-sentence-criminals-must-stop-now/.

EXPERTIUS: A Mexican Judicial Decision-Support System in the Field of Family Law. **Caceres, Enrique. 2008.** s.l. : JURIX 2008: The 21st Annual Conference on Legal Knowledge and Information Systems, 2008.

Ito, Sam. 2019. In brave new world of China's digital courts, judges are AI and verdicts come via chat app. *The Japan Times.* [Online] 2019.

Tangermann, Victor. 2019. Estonia is Building a Robot Judge to Help Clear Legal Backlog. *Futurism.* [Online] 2019. https://futurism.com/the-byte/estonia-robot-judge.

Solow-Niederman, Alicia and & Re, Richard M. 2019. Developing Artificially Intelligent Justice. *Stanford Technology Law Review.* 2019, Vol. 22.

Volokh, Eugene. 2019. Chief Justice Robots. *Duke Law Journal.* 2019, Vol. 68.

Wu, Tim. 2019. Will Artificial Intelligence Eat the Law? The Rise of Hybrid Social-Ordering Systems. *Columbia Law Review* . 2019, Vol. 119.

Sourdin, Tania. 2018. Judge v Robot? Artificial Intelligence and Judicial Decision-Making. *UNSW Law Journal.* 2018, Vol. 41.

Kahan, Dan. 2013. Ideology, Motivated Reasoning and Cognitive Reflection. *Judgment & Decision Making.* 2013, Vol. 8.

Kelleher, John and Tierney, Brendan. *Data Science (The MIT Press Essential Knowledge Series).* 2018 : s.n.

Thompson, Derek. 2019. Should We Be Afraid of AI in the Criminal-Justice System? *The Atlantic.* [Online] June 2019. https://www.theatlantic.com/ideas/archive/2019/06/should-we-be-afraid-of-ai-in-the-criminal-justice-system/592084/.

Angwin, Julia, et al. 2016. Machine Bias. *ProPublica.* [Online] 2016. https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing.

Kusner, Matt, et al. 2017. Counterfactual Fairness. *31st Conference on Neural Information Processing Systems.* 2017.

Kleinberg, Jon, et al. 2017. Human Decisions and Machine Predictions. *The Quarterly Journal of Economics.* 2017, Vol. 133.

Casey, Anthony and Niblett, Anthony. 2019. A Framework for the New Personalisation of Law. *University of Chicago Law Review.* 2019, Vol. 86.

Singh, Seema. 2018. Understanding the Bias-Variance Tradeoff. *Towards Data Science.* [Online] 2018. https://towardsdatascience.com/understanding-the-bias-variance-tradeoff-165e6942b229.

Hybrid Approach Combining Machine Learning and a Rule-Based Expert System for Text Categorisation. **Villena-Roman, Julio, et al. 2011.** s.l. : Proceedings of the 24th International Florida Artificial Intelligence Research Society Conference, 2011.

Hacker, Phillip, et al. 2020. Explainable AI under Contract and Tort Law: Legal Incentives and Technical Challenges. *Artificial Intelligence and Law* . 2020, Vol. 28.

Buocz, Thomas Julius. 2018. Artificial Intelligence in Court: Legitimacy Problems of AI Assistance in the Judiciary. *Copenhagen Journal of Legal Studies.* 2018, Vol. 2.

Re, Richard M. 2014. Narrowing Precedent in the Supreme Court. *Columbia Law Review.* 2014, Vol. 114.

Lai, Ho Hock. 2019. The Fair Trial Rationale for Excluding Wrongfully Obtained Evidence. *Ius Gentium: Comparative Perspectives of Law and Justice (Volume 7).* s.l. : Springer, 2019.

Sun, David. 2019. Potential to excel in life: NUS undergrad who molested woman gets probation for 'minor intrusion' offence. *The Straits Times.* [Online] 2019.

Zula. 2020. How Justice was Served for 3 NUS Offenders Sends a Disappointing Message to Women in Singapore. *Zula.* [Online] 2020.

Surden, Harry. 2014. Machine Learning and Law. *Washington Law Review.* 2014, Vol. 89.

Comes, Richard and Sourdin, Tania. 2018. Do Judges Need to be Human? The Implications of Technology for Responsive Judging. *The Responsive Judge: International Perspectives.* s.l. : Springer, 2018.

Methods for Rule Conflict Resolution. **Lindgren, Tony. 2004.** s.l. : European Conference on Machine Learning, 2004.

Bell, Evan. 2013. An Introduction to Judicial Fact-Finding. *Commonwealth Law Bulletin.* 2013, Vol. 39.

Gennaioli, Nicola and Shleifer, Andrei. 2008. Judicial Fact Discretion. *Journal of Legal Studies.* 2008, Vol. 37.

Supreme Court of Singapore - Court of Appeal. *Spandeck Engineering (S) Pte Ltd v Defence Science & Technology Agency.* [2007] 4 SLR(R) 100 (SGCA). s.l. : 2007.

Personal Data Protection Commission of Singapore. 2020. *Model Artificial Intelligence Governance Framework: Second Edition.* s.l. : Personal Data Protection Commission of Singapore, 2020.

Myagar, Suvda, Lee, Adam and Yurcik, William. 2005. *Threat Modelling as a Basis for Security Requirements.* s.l. : IEEE Symposium on Requirements Engineering for Information Security, 2005.

Bacudio, Aileen, et al. 2011. An Overview of Penetration Testing. *International Journal of Network Security and Its Applications.* 2011, Vol. 3.

Comiter, Marcus. 2019. *Attacking Artificial Intelligence: AI's Security Vulnerability and What Policymakers Can Do About It.* s.l. : Harvard Kennedy School Belfer Centre Paper, 2019.

Parmar, Hiren and Gonsai, Atul. 2015. Analysis and Study of Network Security at Transport Layer. *International Journal of Computer Applications.* 2015, Vol. 121.

Das, Manik Lal and Samdaria, Navkar. 2014. On the Security of SSL/TLS-Enabled Applications. *Applied Computing and Informatics.* 2014, Vol. 10.

Grigorik, Ilya. Chapter 4: Transport Layer Security. *HPBN*. [Online] https://hpbn.co/transport-layer-security-tls/.

Sarikaya, Kazim and Can, Ahmet Burak. 2011. Password-based Client Authentication for SSL/TLS using ElGamal and Chebyshev polynomials. *5th International Conference on Application of Information and Communication Technologies*. 2011.

Mallik, Avijit, et al. 2019. Man-In-The-Middle Attack: Understanding in Simple Words. *International Journal of Data and Network Science*. 2019.

Kahder, Aqeel Sahi and Lai, David. 2015. Preventing Man-In-The-Middle Attack in Diffie-Hellman Key Exchange Protocol. *22nd International Conference on Telecommunications*. 2015.

Hema, Thunjira, Raj, Michael and Rabara, Albert. 2018. An Analytic Method of Elliptic Curve Cryptographic Security. *International Journal of Scientific Research in Computer Science Applications and Management Studies*. 2018, Vol. 7, 1.

Ronen, Eyan, et al. 2019. The 9 Lives of Bleichenbacher's CAT: New Cache Attacks on TLS Implementation. *IEEE Symposium on Security and Privacy*. 2019.

Abie, Habtamu. 2000. *An Overview of Firewall Technologies*. s.l. : Norwegian Computing Centre, 2000.

Pundar, Shwetambari and Bamnote, G. R. 2014. Analysis of Firewall Technology in Computer Network Security. *International Journal of Computer Science and Mobile Computing*. 2014, Vol. 3.

Rao, Umesh Hodaghatta and Nayak, Umesha. 2014. Chapter 11: Intrusion Detection and Prevention Systems. *The InfoSec Handbook: An Introduction to Information Security*. s.l. : Apress Open, 2014.

Jackson, Nic. 2018. Network Segmentation in Modern Environments. *HashiCorp*. [Online] 2018.

https://www.hashicorp.com/blog/network-segmentation-in-modern-environments.

Marshall, Andrew, Rojas, Raul and Brinkman, Donald. 2018. Securing the Future of Artificial Intelligence and Machine Learning at Microsoft. *Microsoft*. [Online] 2018.

Barreno, Marco, et al. 2010. The Security of Machine Learning. *Machine Learning*. 2010, Vol. 81.

Ying, Zhao, et al. 2020. PDGAN: A Novel Poisoning Defence Method in Federated Learning Using Generative Adversarial Network . *ICA3PP 2019: Algorithms and Architectures for Parallel Processing*. 2020.

Collinge, Greg, Lupu, Emil and Munoz-Gonzalez, Luis. 2019. Defending against Poisoning Attacks in Online Learning Settings. *European Symposium of Artificial Neural Networks*. 2019.

Darktrace. 2019. *Darktrace White Paper, Machine Learning in the Age of Cyber AI: A Review of Machine Learning Approaches for Cyber Security and Darktrace's Underlying Technology*. s.l. : Darktrace, 2019.

—. 2017. Darktrace Stops Ransomware Attack at UK Construction Company. *Darktrace News*. [Online] 2017. https://www.darktrace.com/en/press/2017/197/ .

Satell, Greg and Sutton, Josh. 2019. We Need AI That Is Explainable, Auditable and Transparent. *Harvard Business Review*. [Online] 2019. https://hbr.org/2019/10/we-need-ai-that-is-explainable-auditable-and-transparent.

Rothman, Daniel. 2017. The Promise of AI in Audio Processing. *Towards Data Science*. [Online] 2017. https://towardsdatascience.com/the-promise-of-ai-in-audio-processing-a7e4996eb2ca.

Youngcho, Yoon, Guimaraes, Tor and Swales, George. 1994. Integrating Artificial Neural Networks with Rule-Based Expert Systems. *Decision Support Systems.* 1994, Vol. 11.

Grosan, Crina and Abraham, Ajith. 2011. *Intelligent Systems: A Modern Approach.* s.l. : Springer, 2011.

Supreme Court of Singapore - Court of Appeal. 2008. *Lau Siew Kim v Terence Yeo Guan Chye.* [2008] 2 SLR(R) 108 (SGCA). s.l. : Supreme Court of Singapore - Court of Appeal, 2008.

Privy Council. 1913. *Loke Yew v Port Swettenham Rubber Co Ltd .* [1913] AC 491. s.l. : Privy Council, 1913.

Supreme Court of Singapore - Court of Appeal. 2006. *United Overseas Bank Ltd v Bebe bte Mohammad.* [2006] 4 SLR(R) 884 (SGCA). s.l. : Supreme Court of Singapore - Court of Appeal, 2006.

High Court of Singapore. 2008. *Malayan Banking Berhad v Sivakolunthu Thirunavukarasu.* [2008] 1 SLR(R) 149 (SGHC). s.l. : High Court of Singapore, 2008.

Junghwan, Cho, et al. 2016. How Much Data is Needed to Train a Medical Image Deep Learning System to Achieve Necessary High Accuracy? *4th International Conference on Learning Representations.* 2016.

Supreme Court of Singapore - Court of Appeal. 2016. *Su Emmanuel v Emmanuel Priya Ethel Anne.* [2016] 3 SLR 1222 (SGCA). s.l. : Supreme Court of Singapore - Court of Appeal, 2016.

High Court of Singapore. 2010. *Wong Chee Siong v Tan Boon Hwa.* [2010] SGHC 222 at [33]. s.l. : High Court of Singapore, 2010.

—. 2011. *E C Investment Holding Pte Ltd v Ridout Residence Pte Ltd.* [2011] 2 SLR 232 (SGHC). s.l. : High Court of Singapore, 2011.

3 Artificial Intelligence and Terrorism Prevention: Finding a balance between Privacy Right and National Security

Prince Samuel Amadi & Herbert Best Eti

Solicitors, Supreme Court of Nigeria

Abstract. For more than two decades now terrorism has rampage the world, causing diverse damming consequences. These include kidnapping, torture, maiming and killing of innocent people globally by terrorist. The worst victims of terrorism are women and children as it is always the case in situation of unrest. Following the incessant terrorist acts which before now were concentrated mostly in the Middle-East, albeit, now affecting almost every corner of the world, the United Nations and law enforcement agencies of various countries declared that the best way to win the fight against terrorism is through the use of artificial intelligence. In this choice of the use of AI in the fight against terrorism the major concern remains the question of abuse of the use of AI. This concern is especially so considering the antecedent of the government and law enforcement in flagrant abuse of rights of the citizens. Given that security is a duty incumbent on the government- to secure it national borders and be free from threat(s) or actual attack(s) either from within or from without on the one hand, and the protection of the rights of individuals against unwarranted government incursion, this paper thus seeks a balance between these competing interests. It therefore espouses the concept of AI and its use in the in terrorism prevention, investigation and ultimate prosecution. While it observes the advantages of the use of AI in the fight against terrorism, this article also seeks to balance its use against the competing right of privacy of the individual. It begins by the discussion of general concept of AI. Secondly, it delves into the various uses of AI. The subsequent section discusses the challenges of the use of AI in terrorism detection and prevention. Lastly, it concludes with a proposal towards a balance use of AI and the protection of individual privacy right.

Introduction

As technology develops, governments, through its law enforcement agencies tend to take advantage of it in the fight against crime. The use of technology has aided in the past and present in the prevention, detection, investigation and prosecution of crimes. Ostensibly, the use of technology for investigation and prosecution of crimes cannot be overemphasized. However, what has always and remains the challenge, is the extent to which government may use technology in its criminal investigation and prosecution while performing its security duty and achieve peace within its borders without undue or unwarranted violation of the right of the individuals within its jurisdiction.

The discussion on the issue of security and privacy rights is somewhat delicate in that, it lies in the labyrinth or intersection of two competing important interests. The first concern is the duty of government to protect itself from threat and attacks, on one hand; and the right of individual to be secured against unwarranted intrusion by the state (Slove, et al., 2021). Therefore, this article seeks to address the question of the use of artificial intelligence (AI) by government in the fight against terrorism. It further examines the extent of the use of AI in combating terrorism and the rights to privacy.

As a matter of introduction, it is argued here that the overall objective of the state is the welfare of its people (Heyman p. 515). These include its entire population: citizens and non-citizens alike. Since one of the cardinal functions of the state is security, it means that the state must continue to deploy every reasonable means to ensure that this objective is achieved. In doing that, the state takes advantage of development in science and technology (Some, 2018). In the recent times, AI has taken a centre stage as an emerging technology that can be used for efficient policing and crime fighting within the society (Miracola p. 18). Since the use of AI has proven effective and relatively efficient in crime detection, prevention and investigation, it has been argued that law enforcement should utilize AI in the fight

against terrorism. As observed in the foregoing sentence, this article reiterates the argument that AI is necessary for the fight against global rise in terrorism if the state is to achieve its security objective.

AI, like every technology or instrument of policing, has its shortcomings. For instance, in China, the extreme surveillance of the populace raises privacy concerns and the question of democracy in the whole system of AI controlled society, which has been described as heading to "dystopian" regime, "where every aspect of life is under constant scrutiny" (Champbell, 2019). Still, the benefits of the utility of AI are enormous for the overall interest of the state.

These benefits as seen from countries which have adopted AI in the fight against terrorism stand out (Miracola). Thus, without undermining the inherent legal challenges in the use of AI, it is highly recommended here that AI should be adopted and deployed for effective terrorism detection and prevention. The Chinese government has taken the lead in the use of nearly all available means of AI, ranging from facial to voice recognition tracking; all encompassed in what has been termed the "surveillance capitalism" in China's fight against terrorism today (Miracola). As a means of fighting terrorism, the Chinese government "created integrated Joint Operations Platform (IJO) that uses AI to monitor checkpoint in and around its cities" (Miracola). Through this AI riven technology, the "social profile and facial attributes" of persons are captured, thereby making it easier for data analytics and identification of potential terrorist gathering within the Republic of China.

Apart from the Chinese government's deployment of AI in crime detection, prevention and investigation, other countries, such as the UK and the US, also utilizes AI for predictive policing. The common use of AI programmed technology in UK and US, just like China, is the surveillance (Malik, 2018). The only difference is that, while the Chinese government is open about the ultimate use of AI; including facial recognition and behavioural surveillance of its citizens. The UK, US and most countries conduct surveillance majorly through data

gathering and analysis, which also involve facial and voice recognition. And more recently, it has been shown that UK's deployment of AI has behavioural software aimed at "preventing crime before it occurs". Irrespective of individual perception, the point being made is that, AI has been deployed by many countries globally for crime detection, prevention and investigation. Thus, this paper canvasses for the "right deployment" of AI for the national security.

Artificial Intelligence: An Overview

In order to have a full grasp of this segment, it is imperative to first understand the idea or concept of AI. As such, in this session, the meaning of AI will be espoused. While the meaning of AI is evolving and inconclusive, adoption of a close definition that is relevant to this paper will be the best approach in presenting a meaning of what constitute AI. In that sequence, this segment will again consider the use of AI in criminal investigation, specifically tailored towards terrorism prediction, prevention and prosecution. In considering the above issues, this session will again raise the question of the benefits of AI. And in contrast to the benefits, this paper will place the benefits side by side with the potential problems arising from the use of AI. In addition, it will examine how AI is used in targeting terrorism. Finally, it will consider whether the approach or method of the use of AI in counterterrorism is the best approach and at the same time consider ethical issues arising from the use of AI.

What is artificial intelligence?

Artificial intelligence is a broad concept that cannot be completely captured in a single definition. In other words, its definition continues to expand as the concept expands. At some point, AI was defined within the perimeter of "science and engineering" capable of producing "intelligent machines" (Rigano). Again, AI is described as the "ability of a machine to perceive and respond to its environment independently", performing tasks natural to human without aid or support of human intelligence (Rigano). To Forester, AI consists of

the "embryonic method of programming", which allows "software to mimic human thought" (Forester, 1985). While referencing the American Association for the Advancement of Artificial Intelligence, Dupont, quoting Robert Atkinson, describes AI "as the scientific understanding of the mechanisms underlying thought and intelligent behaviour and their embodiment in machines" (Dupont, 2019).

Commenting on the definition, Dupont described it as "a very broad net" (Dupont, 2019). He noted further that the definition is broad because "it includes any intelligent seeming behaviour a machine can perform" (Dupont, 2019). In justifying his argument that the definition of AI offered by the American Association for the Advancement of Artificial Intelligence is broad, he did so using a programming chat interface which merely asks questions that allows yes or no to equally exhibits some degree of intelligence (Dupont, 2019). The second illustration of Dupont showing the broad definition of AI, as noted above, is "electric drier that stops when it senses that cloths are dry". These devices, argued Dupont, exhibit intelligence of some sort. Given that several machines exhibit certain level of intelligence by the engineering design, it is confusing to categorize such machines as possessing artificial intelligence within the context of "intelligence".

Having perhaps struggled with the comprehensive definition of AI, Dupont considered a rather descriptive approach to understanding the meaning of AI. In that light, he considered AI in categories and concludes that, "AI is generally split into two categories: General Artificial Intelligence and Narrow Artificial Intelligence". To the general AI, which he considers to be stronger, he asserts that it is "thought to be a computer system exhibiting human or superior intelligence in all fields. It would be able to take knowledge from one field and transfer it to another" (Dupont, 2019). This type of AI, in his opinion, could have enormous impact on the human society and perhaps replace all human labour (Dupont, 2019). He however deferred the possibilities of having an AI with such capability to the future.

On the contrary, he views narrow AI as limited and concludes that "all human achievements" so far made in the field of AI falls within the "category of narrow AI" (Dupont, 2019). In further differentiating general AI from narrow AI, Dupont opined that narrow AI is concerned "with solving a predefined problem" such as board-game, image identification and car driving control (Dupont, 2019). Although Dupont recognized the significance of his narrow classification of AI, he however noted that "it is not concerned with a fully conscious, human-level intelligence" (Dupont, 2019).

No definition of AI will suit the expanse of the whole concept, as the concept of AI is broad and supersede any solitary definition or explanation. Noting this conundrum, Allen observed that "AI is an extremely broad field, one that covers not only the breakthroughs of the past few years, but also the achievements of the first electronic computers dating back to the 1940s" (Allen). The observation of Allen is not far from the reality surrounding AI. Beyond any definition, AI encompasses those actions which are ordinarily human oriented. These actions which before now could only be carried out with the use of human intelligence include language-understanding natural language, images- facial classification and sounds- speech recognition and motion detection. With the development of technology and machine learning programming, AI is now designed to "respond in useful ways to language, images, and sounds" (Boucher, 2020). The above summarises the broader capabilities of AI in crime detection and prevention in law enforcement.

The uses and benefits of AI in detection and prevention of crime

As noted above, law enforcement leverage on available technology in its fight against crime. As a technological instrument, law enforcement has in the past few years rely more on AI in detection, prevention and prosecution of crime. Bearing that in mind, this segment discusses the use of AI in crime detection, prevention and prosecution. In general, it accesses the capabilities of AI in relation to overall criminal justice.

The capabilities of AI are therefore classified into the following: object classification; object recognition; speech recognition; gunshot detection; DNA analysis and digital forensics. These AI use capabilities will be discussed in the details below.

In this discussion, the first capability of AI in crime detection and prevention to be discussed here is the object classification. Using object classification AI group some elements found in images and video and go on to provide labelling for the elements. This AI does unilaterally and independently without human effort. Using this algorithmic software, AI is able to categorise elements of images and videos and make a decision based on the categorization. Therefore, applying the object classification of AI, access is gained into the image surrounding commission of crime (Dupont, 2019). Through the images or video, location is identified. This is most recently identified using the Google program which, as observed, relies "on conventional neutral networks for its geolocation" (Dupont, 2019). The use of object classification of AI is equally significant in that, it is deployed to "detect… possible existence of criminal activities depicted within an image" (Dupont, 2019). Since the aim of the algorithm is to identify, classify and select, it becomes a ready tool for detection of a classified image or video with possible or likely criminal activity. An illustrative example of how AI uses image or video classification in detection of crime is the PhotoDNA. Commenting further on the use of AI in DNA, the author observed that:

> [The tool] primarily aims to detect a child's pornography and works by a) creating a digital signature (known as a 'hash') associated with the image to prevent image alterations, and b) converts the image to black and white, resizes it, breaks into a grid, and quantifies its shading. It then compares an image's hash against a database of images that have been identified as illegal, and matches can be manually reviewed by humans. Other examples of technology that seek to detect the commission of a crime within imagery include the European P-REACT Project, the loss-prevention product

offered by the US-based company StopLift, and the Chinese software SenseTime (Dupont, 2019).

Lastly, Dupont noted that the object classification capability of AI using image and video identification is useful to law enforcement in corroborating "findings of criminal activity" (Dupont, 2019). Since the burden of establishing crime rest on the prosecution, the law requires credible evidence to be presented by the prosecution in proof of any alleged criminal offence. In deploying object classification mechanics of AI, it aid in corroborating existing findings of criminal activities by law enforcement. The AI does the extraction of images and videos for recognition, which eventually is utilized by the law enforcement in corroborating any criminal activity.

The second use of AI in crime detection and prevention is object recognition. At the centre of object recognition use of AI is face recognition. Facial recognition has generated heated debate amongst scholars, writers, policy makers and opinion holders from several quarters, including the European Union Parliament. However, discussion on the argument surrounding facial recognition will be deferred to the discussion on the challenges of the use of AI. Here, attempt will only be made as to how law enforcement uses AI object (face) recognition in detection and prevention of crimes. By object recognition, it connotes the ability of AI enabled machine to recognise certain categories of specific targeted subjects with the aid of algorithmic programming (Deepomatic). The underpinning aim of object recognition capability of AI is the ability to understand images and videos for purpose of classifying targeted subjects. It is the abilities of machines powered by AI programming "to recognize… things and entities". Through this algorithmic pattern, AI is engineered to locate objects in images and videos with a considerable degree of certainty. As noted earlier, this aspect of AI use is surrounded by heated controversy, which revolves around its intrusive nature to the privacy of the individual. Still, governments around the world have already deployed this AI machine in their security architecture. As mentioned at the beginning, while more nations are adopting the use of facial

recognition in their fight against crime in general and terrorism in particular, China remains a classic example of this trend (Miracola).

According to The U.S. National Security Commission on Artificial Intelligence report, the Chinese government has developed AI oriented machine, which "allows machines to exhibit characteristics associated with human" which are "applied to areas of facial and speech recognition, natural language processing and automated reasoning" (Hsu, 2021). Also, while it is not clear whether the Russian government would make use of facial recognition in the same manner as China, the Russian government has set up plan on expanding the application of AI in all facets of life in Russia (The Future of Life Institute, 2018). The main aim of the Russian government program of action on AI "include improving the availability and quality of data, increasing the availability of hardware and creating appropriate standards and a regulatory system that guarantees public safety and development of AI technologies" (The Future of Life Institute, 2018). Another classic example of countries with the use of facial recognition is South and North Korea. According to the new project launched by the government of South Korean, the AI facial recognition developed by it would be used to track and "monitor the activities and movements of some 800,000 citizens" (Pesek, 2019). Like the South, North Korea also has its system of AI facial recognition machine which, supposedly is to be deployed in monitoring the activities of its citizens in "its home market" (The Biometric Update, 2021). Commenting on the privacy implication in the application of this pattern of security check, it is observed that the use of AI facial recognition is akin to the totalitarian regime painted and warned by George Orwell (Pesek, 2019).

On its part, the European Union adopted a regulated system for the application of AI within its union. By the tenor of the EU regulatory instrument (General Data Protection Regulation (GDPR), AI in general and facial recognition technology (FRT) will only be allowed within the borders of its member states where such technology provider exhibits proof of a system built "tailored approach to risk

management and quality processes" (Louradour, 2021). Since the EU considers the use of AI as "high-level risk system", it limited the use of AI or "real-time remote biometric identification systems in publicly accessible spaces for the purpose of law enforcement" to only "exceptions related to public safety such as the targeted search of missing persons or the prevention of imminent terrorist threats". The EU regulation also provides for the modus of granting authorization for the use of AI facial recognition technology upon 'evaluation' "to law enforcement agencies" (Louradour, 2021). Authorization is to be granted on individual basis "by a judicial authority or by an independent administrative authority of Member States, unless it is operated in duly justified situation of urgency" (Louradour, 2021).

Indeed, the adopted regulation of the EU is a better approach to the use of AI. It balances the use of AI and the protection of right to a greater degree possible. While still prioritizing the nation's security, the EU regulations ensure respect for the dignity of the human person by not arbitrarily eroding into private affairs of the citizens. In fact, when in the extreme circumstance such as, searching for a missing person or imminent occasion of terrorist attack for the application of AI, it still maintains standard procedure to ensure that the civil rights, especially, the privacy right is guaranteed and protected. One of the shortcomings to the EU regulation is that it fails to indicate what constitute expediency for the grant of authorization by the judicial or administrative body. In other words, it fails to specify in clear terms situation of urgency that application of authorization would be required.

In undergoing object recognition task, AI undertakes four categories of tasks, which are: "classification, tagging, detection, and segmentation" (Deepomatic). Through these means, specific targeted subjects within images and videos are classified, identified and segmented for decision making (Lisowski, 2021). An example of use of image recognition is the identification of vehicle numbers. By applying image recognition, AI launches series of activities which include data gathering- that is, the classification of the image, data

tagging and ultimately, data segmentation. Through these levels of activities, AI achieves a reliable and dependable chat of object recognition for decision making. This method of AI use is readily deployed by the law enforcement in identification of targeted objects like handwriting, vehicle plate numbers, fingerprints, and human face in detection, prevention and eventual prosecution of crime.

In addition to the above, AI also has the ability to recognise speech. This is called the recognition component of AI. But then, what is speech recognition? Like the other components, there is no adequate definition of speech recognition. Instead, it has been defined or described according to the background of the definer. Thus, speech recognition has been described as the process initiated with the computer system with the ability to recognise speech and repeating the same thing in real time. It is further described as a process that "functions as a pipeline that converts Pulse Code Modulation (PCM), digital audio from a second card into recognized speech". From another point of view, speech recognition is taken to mean a self-coordinated "transcription of spoken language into a readable text" (Rayan). Speech recognition is also described as the ability of "technology to identify speech patterns" with the overall goal of identifying the speaker (Dupont, 2019).

Notwithstanding the manner of description of speech recognition, the underpinning point is the ability of AI enabled machine using sound waves and frequency pattern of the spoken word to recognize the identity of the speaker. Beyond identifying the speaker, the algorithm by its process also function to classify sound to determine the exact spoken word (Dupont, 2019). The voice recognition today is a viable tool of operation in the hands of the law enforcement in the detection and prevention of crime. According to a review of Interpol Project on the use of speech or voice recognition, "the combination of speaker identification with other biometric technologies such as fingerprints and facial recognition…enhance investigative capabilities". Thus, law enforcement agencies use speech recognition for improved policing.

Speech recognition, like other AI capabilities, aids law enforcement in conducting efficient and effective policing (Tombul, et al., 2021).

Indeed, the use and benefits of AI in crime detection and prevention is enormous. Various machines with AI programming have been deployed by the law enforcement to carter for the security need of the society. Examples of such AI enable machines or technology used for the above discussed capabilities of AI for detection and prevention of crimes include close circuit television (CCTV), gunshot detection software, DNA analysis machines, and digital forensics software.

Targeting terrorism through the use of AI

This segment discusses the role of AI in targeting terrorism and how AI has been used to combat terrorism. But first, an understanding of terrorism is significant for proper review of how AI is used to combat it. In a bid to define terrorism, some authors as categorized terrorism "based on victim, the method and the purpose. Activity of violence targeting a group on a large scale" (Muntha, et al., 2020). The authors went on to observe that "the world has been victim of such unlawful activity since a long time" and "for steady growth of the world and civilization, it is very important to fight back with all we got and try to extinct such unlawful activities" (Muntha, et al., 2020).

Since the 9/11 terrorist attack on the US, terrorist activities have continued to gain global prominence. Several organized terrorists have continued to emerge over the years. Terrorist groups are getting bigger and more powerful following the merger of smaller terrorist groups to form larger terrorist and crime syndicates. Terrorist groups performs other form of organized crimes such as: internet fraud, money laundering, kidnapping, oil bunkering, piracy, drug peddling, human trafficking, etc. as a means of financing the advancement of terrorism.

The global upsurge in terrorism has equally been matched with global legislative action. This has taken the form of international counter-

terrorism legislations at the global, regional and bilateral level. At the global level, apart from the several declarations made by the United Nations General Assembly, there are over ten United Nations counter-terrorism Conventions, amongst which are: International Convention for the Suppression of Terrorist Bombings 1997; Convention for the Suppression of Terrorist Financing 1999; International Convention for the Suppression of Acts of Nuclear Terrorism 2005 etc.

At the regional level, the African Union have established the Algiers Convention on the Prevention and Combating of Terrorism 1999. European response to terrorism is marked by the European Convention on the Suppression of Terrorism 1977. In 1998, the League of Arab States created the Convention for the Suppression of Terrorism. The counter terrorism instrument of the Organization of American States, the Inter-American Convention against Terrorism 2002 also stands out among the regional instruments for suppression of terrorism. An instance of bilateral counter-terrorist treaty is the Agreement between the United States and Cuba on the Suppression of Certain Terrorist Acts 1973. Several states have also signed extradition, information sharing, technical assistance, etc. counter-terrorism treaties amongst themselves.

The objective of most of these conventions is to facilitate the cooperation of State Parties into adopting relevant measures that would ensure the prevention, punishment and elimination of various expression of terrorism as per Article 1 of the Inter-American Convention against Terrorism, 2002. A common denominator amongst the international counter-terrorism instruments is that they all urge state parties to those various conventions to adopt measures towards dealing with terrorism threat. It is pursuant to this provision that countries like US and Nigeria created the PATRIOT Act and Terrorism (Prevention) Act 2011, respectively. Part of the measures which states can adopt to this effect would be to leverage on technology, particularly AI, to match the advancing sophistication of terrorist activities.

Terrorist groups have leveraged on the use of internet for the purpose of quick dissemination of toxic messages, fund raising, communicating with syndicate groups, purchase of weapons, recruitment and indoctrination of members and obtain other forms of support. Although most of their communications are cryptic, they are usually in open sources, mostly social media (Voyager Labs, 2021). This puts these terrorist groups in a vulnerable position, as there are chances of their communications being detected and decoded by proactive investigation. This is where AI can be leveraged upon. AI, can, in quick succession, decode and analyse loads of database that would take experts several days to decode. AI can be used to capture the "content of terrorists at a faster speed, within some minutes from social media" (Voyager Labs, 2021). It can easily identify terrorist associated patterns and parameters and create early warning signs that criminal investigators can quickly act upon to foil potential terrorist threat and attacks before they materialize. The following section shall consider the various format or expression of AI in combating terrorism.

2.1.1 Image or identification/face recognition

One of the most potent uses of AI today is image identification/facial recognition. While image identification is slightly differently different from facial recognition, they relatively perform similar functions. Facial recognition as "AI-powered object recognition" is a technology which enables the analysis of video footage to identify targets (UNICRI, 2021). As an increasing technology, facial recognition has been adopted for use by law enforcement across the globe (UNICRI, 2021). For instance, the German authority "in 2017 and 2018 piloted" the use of "facial recognition-enabled CCTV" within its city for the sole aim of identifying terrorist and offenders in public places (UNICRI, 2021).

Nigeria offers another example where AI facial recognition is employed for counter terrorism. Faced with years of terrorist attacks, Nigeria declared the adoption of AI-object and facial recognition in its fight against the Boko Haram and other terrorist attack in the country

(News Agency of Nigeria, 2019). "INTERPOL also operates a facial recognition system", which has in store facial images received from countries" around the world (UNICRI, 2021). The AI powered algorithm is also used "for forensic investigation" through analysing of video evidence collected online to determine perpetrators of crime (UNICRI, 2021).

2.1.2 Speech recognition

As crime, especially terrorism becomes the major concern of the world, government and law enforcement, as observed; continue to use every available technology to ensure the security of lives and properties within its territory. The contemporary fight against crime-terrorism, being such that is driven using AI, speech recognition by AI-powered technology, is deployed by law enforcement target terrorist. With AI, "speech and gesture as natural means of communication by humans" is being carried out by machine to select and interpret human speech.

As noted earlier, speech recognition as a component of AI "seeks to identify elements of speech patterns" with the goal of identifying the maker and the words spoken (Dupont, 2019). Through speech recognition, law enforcement targets terrorists to detect and prevent terrorist attacks. Through speech recognition, the sound is measured, wave and frequency patterns of speech signal are detected (Dupont, 2019). Upon recognition of the targeted speech, the AI-powered software classifies data collected, segment and identify what was or is being said from the speech extracted.

2.1.3 Tracking

Tracking is an AI enabled device used in monitoring and tracing object. It entails "estimating the state of the target object present in the" scene from prior given information (Shah, 2020). This AI system is deployed by law enforcement agencies in tracking particular objects, by analysing "videos to identify the object belonging to" particular

"categories, such as pedestrians, cars, animals and inanimate objects, without any prior knowledge about the appearance and the number of targets" (Shah, 2020).

Through this AI device, the law enforcement tracks objects used by terrorist, either before or after a commission of terrorist act. Using this tool, objects such as vehicle plate number is tracked by the law enforcement agency, as it is able to "access precise, accurate, and timely information… essential, especially in dealing with crimes" (Mubarak, A. U. et al.;, 2021). While tracking device in targeting terrorism is useful, however, it is laden with certain setback which can call to question its accuracy. Some of the factors which may affect an AI-powered device have been identified to be in "uncontrolled environment, weather or un-tactical positioning of cameras" (Mubarak, A. U. et al.;, 2021).

2.1.4 Data analysis

Data collection is one of the ways AI is used in the fight against terrorism. "This task is done by trawling the Internet, Deepweb and Darknet for specific information about terrorism chats, webs and forums where information for activities can be found, as well as open data sources that could be linked to the case." (Vallis-Prieto) Since contemporary society depends on the use of internet in virtually every sphere of endeavour, it becomes expedient to acquire data "necessary to process natural language in order to extract information that can be processed by a machine" for detection and prevention of terrorism.

Through data collection and analysis terrorist activities is monitored (Muntha, et al., 2020). As one of the ways of combating terrorism, data is largely collected and narrowed to a particular target. The essence of this operation is to use AI in analysing communication and finances used in terrorism. By this method of AI, large data is collected, sorted, and analysed for prediction and prevention before the eventual occurrence of a particular terrorist activity. As observed somewhere, data gathering or collection "focuses not on crime, but on the

possibility that a crime might be committed at some future date" (Bjelopera, 2014). The foregoing demonstrates how AI can be put to effective use in the fight against crime in general and terrorism in particular.

Potential legal problems and ethical issues arising from the use of AI

Despite the promises AI holds in the fight against crime and terrorism, several concerns or challenges have arisen from its deployment. This segment of this paper will concentrate on these concerns or challenges with a view of balancing them, while still maintaining the use of AI in the fight against terrorism. The problems arising from the use of AI will be discussed below under two broad heading of ethics of AI and the privacy right issues arising from the use of AI. At the end of this discussion, an alternative argument will be suggested for the use of AI, with a view of avoiding the ethical and legal issues.

Ethical issues

Ethics is one of the major concerns of the use of AI. With AI in the hands of users or operators comes great opportunity or power which, if left unchecked, may result and will continue to result in misuse and abuses. AI no doubt, without proper guidelines for its use "offers those in its possession" with great abilities which may not only be used for good cause but an instrument of oppression or unlawful actions in the hands of its operators (Olech, et al., 2021). Without guiding code AI can be used in many ways to hurt innocent citizens. For instance, China has for many years implemented surveillance of its citizens (Das, 2020). Apart from government surveillance of the people without their consent, other ethical issues exist in the areas of DeepFakes. DeepFakes as an AI wired programming system allows for creation of speech and actions for an individual who neither said the words nor performs the actions (Das, 2020). This aspect of AI poses a grave danger in that, without laid down principles to guide the

operation of AI it may lead to wrongful indictment of innocent persons.

Since AI is a machine induced concept, another ethical question becomes the question of accountability (Olech, et al., 2021). How accountable are the operators of AI? This concern remains central to the debate over the use of AI which has driven some nations, especially, the European nations into adopting policies aimed at ensuring ethical use of AI. In addition, the use of AI raises the concern of accuracy. Since AI is engineered with the aid of machine, the question of accuracy becomes a challenging factor. The argument here is simple. Since AI is programmed to operate in a certain way, it is bound to misrepresent facts. It therefore means that accuracy and accountability will be almost difficult to charge as judgment made by AI cannot exactly be accurate. AI, not operating with the being human brain is certain to lag in accountability and accuracy identify images and objects with exactitude. Take for instance, in the case of voice, where speech identification is the aim, object identification, where object is in issue, and image, where the contention is facial recognition. In all of these instances, it is doubtful for AI to make accurate and highly dependable judgment. This certainly, poses a limitation to use of AI in combating terrorism, especially with the use of DeepFakes where speech and actions can be generated independent of the individual believed to have made them.

Further to the above, the need for ethical AI is driven by the consciousness that without guiding principles, individuals will be subject of bias and discrimination. The justification for this is that, since the operators of AI "gain insights from the existing structures and dynamics of the societies they analyse, data-driven technologies can be reproduced, reinforced and amplify the patterns of marginalisation, inequality, and discrimination" already in existence in the society (Leslie, 2019). Again, the issues of "unreliable, unsafe, or poor-quality outcomes" raise concern in the use of AI (Leslie, 2019). Without reasonable implementation of a system where stakeholders in the field of AI, can be responsible in the system they create, there is

every possibility for "irresponsible data management, negligent design and production processes, and questionable deployment practices", which may result in "implementation of AI systems that produces unreliable, unsafe, or poor-quality outcomes". These outcomes, as noted in the instance of DeepFakes can do great harm to the individual in particular and the society in general (Leslie, 2019).

Legal issues

Another problem in the use of AI is its inherent legal issues. Since AI functions with machine, several legal issues arise. The system of AI functions in such random mechanics, whether in image and facial recognition or object tracking, data is randomly extracted from individuals mostly without knowledge or consent. This situation therefore present questions as to what degree is privacy right protected. Since the individuals are either tracked or placed on surveillance without notice, the individual's right under the 4th Amendment and 14th Amendment (under the US Constitution) are usually implicated. Further, the issue of establishing intent in crime is implicated here, since AI feeds on data for criminal prediction, investigation and prosecution. This segment will consider these issues and offer an alternative argument as to the use of AI without violation of the privacy right protected under the 4th Amendment. One the cases in which the court would have to decide whether the 4th Amendments limits the use of AI by the law enforcement is the case of *Carpenter v. United States* (Supreme Court of the US, 2018).

In *Carpenter*, the Supreme Court came to the conclusion that the law enforcement cannot search or access the content of citizens' cellphone without obtaining a search warrant. While the gamut of the case revolves around the question of whether information held by a third party was subject to privacy right provision of the 4th Amendment, it clearly and intrinsically affects modern policing with the use of modern AI powered technology. The judgment of the court was from the overall issue presented to the court, which is: whether the agent of the Federal Bureau of Investigation (FBI) acted within the confines of

the constitution in its bid to retrieve data of a suspected serial robber's cellphone to establish in evidence that he was near the scene of crime where a theft incident had occurred. The court noted that, although technology is useful for policing in the hand of the law enforcement, it pointed that "its use raises the risk of the kind of government 'encroachment' on personal liberty that the framers of the Constitution sought to prevent" (Ware, 2018). This decision like others poses great challenge to the use of AI by law enforcement in crime prediction and prevention. As observed by the dissenting opinion of Justice Anthony Kennedy, the court's decision has the tendency to "unduly interfere with law enforcement's legitimate efforts to investigate and counter serious crimes" (Ware, 2018).

From the court's vantage in *Carpenter*, deference was to be given to human policing which requires almost strict compliance with the warrant condition for any search to be conducted on any citizen in accordance with the provision of the 4th Amendment to the US Constitution. While rights of the citizens is sacrosanct, especially the private right of the citizens, the whole right would be meaningless if government fail in its duty to protect the lives and property of the citizens resulting from bottleneck interpretation of the 4th Amendment provision relating to search and seizure. It then means that for any meaning to be given to the lives of the citizens, while taking every reasonable precaution, a liberal interpretation must be given in order to allow for the application of AI in modern policing with modern means of crime commission.

Another example is the case of *In re Ashley Madison Customer Data Sec. Breach Litigation* (United States District Court Eastern District of Missouri , 2016). The claims were "data breach which resulted in mass dissemination of user information and allegations that the defendants were engaging in deceptive and fraudulent conduct by creating fake 'host' or 'bots', which were programmed to generate and send messages to male members under the guise that they were real women, and inducing users to make purchases on the website". In determining the liability of the defendants, the court held "that the use

of a computer program to simulate human interaction could give rise to liability for fraud" (United States District Court Eastern District of Missouri , 2016).

The challenges of application of AI available literature continue to broaden several factors inhibiting the use of AI. The decision of the court in *United States v. Athlone Industries Inc.*, (United States Court of Appeals, Third Circuit, 1984) is rather instructive on this point (United States Court of Appeals, Third Circuit, 1984). In this case, the court concluded "that "robots cannot be sued" and discussed instead how the manufacturer of a defective robotic pitching machine is liable for civil penalties for the machine's defects". This classically shows the attitude of the court in acceptance of AI as an integral part of today's society.

Evidentiary issue - establishing intent from data

In this discussion, the first identified issue is the evidentiary question of establishing intent from data collected with the aid of AI for the fight against terrorism. As a general rule, in criminal law, crime is established where the two elements of intent and physical action are present (in exception of strict liability offences where the mere establishment of the physical element is sufficient to ground conviction). This elements must co-exist and not one without the other. Unlike other means of crime investigation, AI uses data which is incapable of establishing criminal intent standing alone. Data alone cannot establish crime without proving the intent of a crime suspect with exactitude. By using AI in determining criminal action, the shortcoming arises in the conclusion of the intent of a suspect before it is proved.

Privacy right issue

Commenting on the importance of privacy, an author observed that "Privacy is a fundamental human right, enshrined in numerous international human rights instruments". It is described as "central to

the protection of human dignity and forms the basis of any democratic society". Presupposing that individual be free from unwarranted or unnecessary government intrusion in his private space, it has an impact on other protected rights, such as right to free speech, information and association. Since privacy denotes autonomy of an individual in matters which are entirely private to the individual, without interference by the government or other third party, this then, becomes an issue with the use of AI, especially in surveillance of individuals. Accordingly, AI initiated activities such as surveillance through facial recognition, tracking and use of object recognitions are all argued to interfere with reasonable expectation of privacy of the individuals, thereby implicating the provisions of the 4th, 5th and 14th Amendment primarily. This, therefore, presents privacy violation as a major setback in the use of AI in criminal investigation, especially with the fight against terrorism.

Of all the setbacks to the use of AI in criminal investigation, violation of privacy right is the major concern commentators have spotted out. This is basically because of the threat AI poses on privacy of individual. Since AI evolves within the circle of data configuration and processing, it implicates the usage of personal data. As herein noted earlier, the data are mostly obtained without prior knowledge or consent of persons concerned.

In *United States v. United States District Court (The Keith Case)*, the court considered the question of national security in relation to the surveillance of a citizen of America, on the one hand, and the question of privacy of the individual, on the other hand. In determining whether the 4th Amendment is implicated, where the agency sought surveillance without the procedure of obtaining warrant. The court noted that:

> As the Fourth Amendment is not absolute in its terms, our task is to examine and balance the basic values at stake in this case: the duty of Government to protect the domestic security, and the potential danger posed by

> unreasonable surveillance to individual privacy and expression. If the legitimate need of Government to safeguard domestic security requires the use of electronic surveillance, the question is whether the needs of citizens for privacy and free expression may not be better protected by requiring a warrant before such surveillance is undertaken. We must also ask whether a warrant requirement would unduly frustrate the efforts of Government to protect itself from acts of subversion and directed against it.

The court did not stop there. It went on to declare that:

> [C]ontention in behalf of a complete exemption from the warrant requirement, when urged on behalf of the President and the national security in its domestic implications, merit the most careful consideration. We certainly do not reject them lightly, especially at a time of worldwide ferment and when civil disorders in this country are more prevalent than in the less turbulent periods of our history. [W]e do not think a case has been made for the requested departure from Fourth Amendment standards.

In summary, the court concluded that the surveillance of citizens without following the warrant requirement does not offend the 4th Amendment provision on reasonable expectation of privacy of the individual. This is just one of the instances where the law enforcement deploy AI aided machine with aim of combating terrorism which raises the question of interference with the privacy right of citizens.

Conclusion

In reviewing the use of artificial intelligence by the law enforcement agencies of the government for the fight against terrorism on the one hand and the protection of individual right to privacy on the other

hand, it is important to note that the point of convergence still remains difficult in reconciling the competing interests. As observed at the beginning of this paper, the government reserves the duty to protect its territory from attack of any kind. At the wake of this duty is the constitutional right of the individual which is guaranteed by the provision of the constitution through various provisions, especially the 4th, 5th and 14th Amendment. The use of AI, apart from the due process right of the above referenced provision of the constitution also implicates the provision of the 4th Amendment, dealing with reasonable searches and seizure. More so, the 1st Amendment is also mostly implicated where AI is used indiscriminately in the fight against terrorism by the law enforcement agencies.

Admittedly, the private right of the individual may be implicated under the foregoing referenced provisions of the constitution. However, there exists the moral debate over the time allowed- especially regarding the obtaining of warrant for conducting of searches and seizure of individual properties. The unanswered question remains whether the government should be able to use artificial intelligence in conducting searches of individual which bypasses the warrant requirement of the constitution. On the other hand, if strict adherence is advocated on the use of warrant, would that not with all its consequences jeopardizes the security architecture of the country. Having these competing interests in mind, the conclusion here remains a reasonable application of the use of AI for counter-terrorism measures. In other words, while artificial intelligence should be encouraged by the state for prediction, detection and investigation of terrorism, reasonable legal and ethical guidance should be employed to avoid overt intrusion or undue disturbances of individual right to privacy. Put differently, as much as the circumstances permit, artificial intelligence should be employed in a way that will not arbitrarily violate individual privacy protected by extant provisions of the constitution.

References

Slove, D. J. and Schwartz, P. M. 2021. *Information privacy law.* New York : Wolters Kluwer, New York, 2021.

Heyman, S.T. The first duty of government: protection, liberty and the fourteenth amendment. [Online] https://scholarship.law.duke.edu/cgi/viewcontent.cgi?article=3172&context=dlj.

Some, Kamalika. 2018. Detecting Crime Through Artificial Intelligence. *Analytics Insight.* [Online] September 8, 2018. https://www.analyticsinsight.net/detecting-crime-through-artificial-intelligence/ .

Miracola, S. How China uses A.I. to control society Italian Institute for International Political Studies Retrieved from . *ISPI Online.* [Online] https://www.ispionline.it/sites/default/files/pubblicazioni/isp_commentary_miracola_04.06.2019.pdf.

Champbell, C. 2019. The entire system is designed to suppress us! What the Chinese surveillance state means for the rest of the world. *TIME.* [Online] 2019. https://time.com/5735411/china-surveillance-privacy-issues/.

Malik, Nikita. 2018. How can we use artificial intelligence to prevent crime? . *Forbes.* [Online] November 26, 2018. https://www.forbes.com/sites/nikitamalik/2018/11/26/how-can-we-use-artificial-intelligence-to-prevent-crime/?sh=6fc056e3498c.

Rigano, C. Using artificial intelligence to address criminal justice needs. *NIJ.ojp.gov.* [Online]

Forester, T. 1985. *The information technology revolution.* Cambridge, Massachusetts : MIT Press, 1985.

Dupont, B. 2019. Artificial intelligence in the context of crime and criminal justice: A report of the Korean Institute of Criminology. *Korean Institute of Criminology.* [Online] 2019. https://www.cicc-

iccc.org/public/media/files/prod/publication_files/Artificial-Intelligence-in-the-Context-of-Crime-and-Criminal-Justice_KICICCC_2019.pdf.

Allen, G. Understanding AI technology. *AI.MIL*. [Online] https://www.ai.mil/docs/Understanding%20AI%20Technology.pdf.

Boucher, P. 2020. Artificial intelligence: how does it work, why does it matter, and what can we do about it? (Scientific Foresight Unit (STOA, 2020), within the Directorate-General for Parliamentary Research Services. *European Parliament*. [Online] 2020.

Deepomatic. Object recognition. a new concept. *Deepomatic*. [Online] https://deepomatic.com/en/what-is-object-recognition-and-how-you-can-use-it.

Hsu, Sara. 2021. China and Artificial Intelligence. *The Diplomat*. [Online] April 19, 2021. https://thediplomat.com/2021/04/china-and-artificial-intelligence/.

The Future of Life Institute. 2018. AI Policy - Russia. *The Future of Life Institute*. [Online] July 12, 2018. https://futureoflife.org/2018/07/12/ai-policy-russia/?cn-reloaded=1 .

Pesek, A. 2019. South Korea experiments with AI to track citizens with COVID. *The New American*. [Online] 2019. https://thenewamerican.com/south-korea-experiments-with-ai-to-track-citizens-with-covid/.

The Biometric Update. 2021. North Korea Introducing Facial Recognition Tech Developed in Country-to-Domestic Market . *The Biometric Update*. [Online] 2021. https://www.biometricupdate.com/201911/north-korea-introducing-facial-recognition-tech-developed-in-country-to-domestic-market .

Louradour, Sébastien. 2021. What to know about the EU's facial recognition regulation – and how to comply. *World Economic Forum.* [Online] April 22, 2021. .

Lisowski, Edwin. 2021. Using artificial intelligence (AI) for image recognition . *Addepto.* [Online] 2021.

Rayan, Y. What is automatic speech recognition. *Academia.edu.* [Online] https://www.academia.edu/8033956/What_is_Automatic_Speech_Recognition.

Tombul, F. and Caker, B. 2021. Police use of technology to fight against crime. *Paperity.* [Online] 2021. https://paperity.org/p/61680511/police-use-of-technology-to-fight-against-crime.

Muntha, Y. and Singh, G. D. 2020. A study of leadership and decision making towards growth of artificial intelligence. *Journal of Advances and Scholarly Researches in Allied Education.* 2020, Vol. 17, 2.

Voyager Labs. 2021. Leveraging artificial intelligence to counter terrorism. *Voyager Labs.* [Online] September 9, 2021. https://www.voyager-labs.com/leveraging-artificial-intelligence-to-counter-terrorism/.

UNICRI. 2021. *Countering terrorism online with artificial intelligence: An overview for law enforcement and counter-terrorism agencies in South Asia and South-East Asia, A joint report by UNICRI and UNCCT.* 2021.

News Agency of Nigeria. 2019. Army says AI crucial in war against insurgency. *Pulse.ng.* [Online] June 18, 2019. https://www.pulse.ng/news/local/army-says-ai-crucial-in-war-against-insurgency/bvdsrt.

Shah, D. 2020. The surveillance phenomenon you must know about: Multi Object Tracking. *Vision Wizard.* [Online] April 26, 2020. https://medium.com/visionwizard/object-tracking-675d7a33e687.

Mubarak, A. U. et al.;. 2021. Fighting crimes and insecurity in Nigeria: an intelligent approach. *International Journal of Computer Engineering in Research Trends* . 2021, Vol. 8, 5.

Vallis-Prieto, J. Combating terrorist financing with artificial intelligence systems. *CORE.AC.UK.* [Online] https://core.ac.uk/download/pdf/286429792.pdf.

Bjelopera, J. P. 2014. The Federal Bureau of Investigation and terrorism investigations. *digital.library.unt.edu.* [Online] February 19, 2014. https://digital.library.unt.edu/ark:/67531/metadc284453/m1/1/high_res_d/R41780_2014Feb19.pdf.

Olech, A. K. and Lis, A. 2021. Technology and terrorism: artificial intelligence in the time of contemporary terrorist threats . *ResearchGate.* [Online] January 2021. https://www.researchgate.net/publication/348916912_Technology_and_terrorism_Artificial_Intelligence_in_the_time_of_contemporary_terrorist_threats.

Das, S. 2020. The social impact of artificial intelligence and data privacy issues . *RedGate.* [Online] 2020. https://www.red-gate.com/simple-talk/development/data-science-development/the-social-impact-of-artificial-intelligence-and-data-privacy-issues/.

Leslie, D. 2019. Understanding artificial intelligence ethics and safety: a guide for the responsible design and implementation of AI systems in the public sector. [Online] 2019. https://www.turing.ac.uk/sites/default/files/201906/understanding_artificial_intelligence_ethics_and_safety.pdf.

Supreme Court of the US. 2018. *Carpenter v. United States 819 F. 3d 880.* s.l. : Supreme Court of the US, 2018.

Ware, J. G. 2018. Does the Fourth Amendment block Cops from using articial intelligence? *The Crime Report.* [Online] November 6, 2018.

United States District Court Eastern District of Missouri . 2016. *re Ashley Madison Customer Data Sec. Breach Litigation.* 148 F. Supp. 3rd 1378, 1380. s.l. : United States District Court Eastern District of Missouri , 2016.

United States Court of Appeals, Third Circuit. 1984. *United States v. Athlone Industries Inc.* 746 F.2d 977, 979 (3rd Cir. 1984). s.l. : United States Court of Appeals, Third Circuit, 1984.

Interview Transcripts for AI Now in partnership with the Indian Society of Artificial Intelligence and Law for AI Now

4 On AI and Web3, featuring an Intellectual Property Angle with Ankit Sahni, IP Law Practitioner

Aditi Sharma

Chief Managing Editor

Abstract. This is an interview conducted by Abhivardhan, the Editor-in-Chief, which has been transcribed and paraphrased. He interviews Ankit Sahni, an IP Law Expert.

How should we relate with AI vis-à-vis world of IP Laws simply?

The reason why the question of AI has become relevant in today's world when we talk about Intellectual Property Rights is because – IPRs being generated by Artificial Intelligence is reality & no more stuff of literature or art that we have seen or read, maybe 20 years or 50 years ago. Not from yesterday, but for quite a few years this has been a reality, we have seen programmes generating or completing in their own expression – an idea that was brought into existence by renowned musicians such as Mozart who were not actually able to complete their work for unfortunate reasons by learning from the existing datasets comprising his other works or several other Western Classical Musicians & their datasets – essentially reimaging how Mozart would have finished a particular work or if a painter or a particular artist was alive, reimagining how

he or she would have completed their work. The question of giving protection to IP that gets generated by AI has become more relevant than ever before, while on the other hand, several jurisdictions have legislations with regards to IP which date back to 90s or in India's case that dates back to the 50s – 60s which are somewhat lagging behind in terms of the innovations at the forefront today. So, the policy makers & leaders everywhere need to realize the significance of AI platforms & programmes taking the centre stage in creation & generation of new material.

As the 2000s have passed & the Euro-Crisis has gone, how do you see the role of AI globally in terms of IPR especially & how can India learn from the already existing issues?

Nobody realized but it was almost time for somebody to come forward & bring up this issue, & of course in this case, the credit goes to the project team – the Artificial Inventor project comprising of many imminent personalities.

How the pandemic has played a role cannot be said for certain, but the understanding of the past few months – things have become increasingly digital, perhaps if this pandemic was not to come our way, the scale & pace of us going digital would have ended up being much slower than what it is today, so, in many ways, the Pandemic has acted as an accelerator to the digital journey & has brought a lot of things on to the digital platform cross industries, & has influenced growth everywhere. All kinds of technologies, not just AI but all tech. fields that have improved rapidly so that everchanging or suddenly changed needs of the people around the world would be catered to. Question arises as to whether, if one person – the inventor or developer of an AI programme with talents not up to the mark as a painter or an innovator has his programme develop something, then should that person be awarded with the Intellectual Property Rights relevant to the innovation or invention

at all? & Whether, the creation of one person who is better capable talent wise of generating a creation, is entitled to the same IPRs as a person who relied solely on his creation – the AI/ML Programme to generate a work of art or literature or not?

As far as digital content is concerned vis-à-vis creativity, people usually have issues of two kinds – economics & who to sue. Since the Indo-Pacific is a significant emerging region, how should the approach towards IPRs be framed in the region?

All of it has to deal at a granular level with the perception of others' rights & that of our rights. We are at a juncture where we are talking of the metaverse as being the future, in some time we might be sitting across the world & not just in a 2-dimensional manner. This obviously, complicates issues to a very large extent, because once w3e get out of the conventionalities, everywhere there's something of interest & of significant & every action would have a consequent action & reaction on someone else. The underlining flavour of the era of the setting up of international organizations was exploitation of human rights, in our age & going forward the underlying issues will be privacy, data protection, integrity, IP protection & enforcement, disparagement & reformation, etc in the digital world. A parallel body or a parallel world would need to be set up, as we did when the aim was to have an international body for countries to be able to get together on an international platform.

5 On the Council of Europe's approach to AI Ethics with Gregor Strojin, President, CAHAI

Akash Manwani

Special Associate Editor

Abstract. This is an interview conducted by Abhivardhan, the Editor-in-Chief, which has been transcribed and paraphrased. He interviews Gregor Strojin, the former Chair of CAHAI, Council of Europe.

What does the Committee do, what were the opinions of the stakeholders & how satisfied was everyone with the output?

The CAHAI has managed to fulfil the tasks assigned to the CAHAI (Ad Hoc Committee on Artificial Intelligence). A two year mandate was existent wherein a study of design, investment & application of AI according to the standards of the Council of Europe, human rights, democracy & the Rule of Law was required to be done – where also important was an elaboration & preparation of a compilation of potential elements for future legal frameworks(s) that would address the issues, despite of the challenging circumstances, such as the Covid situation which prevented physical interaction but at the same time allowed technological interaction, a road map was made for the project work - & for many of the members it was the only roadmap that they had had a chance to look at for the past 2 years. The roadmap had 5 plenary meetings scheduled while, in the end 6 were held. A document of recommendations was

added to feed into the future work of the Council of Europe – while negotiations are to start by May this year. The document is still restricted & subjected to vetting by the main policy body of the Council of Europe, the Council of Ministers. If they adopt it or take note of it, the document would be released to the general population. However, it is a relatively short document with 464 paras while each para is an elaboration of what kind of content should be included in a future legal document. So, it gives a certain recipe as to how to prepare a Treaty dealing with AI application, design, development in a way preventing negative impacts that might befall the Rule of Law, Democracy & Human Rights.

The ecosystem in which the work developed – is that the Council of Europe is the oldest & largest human rights oriented intergovernmental organization of Europe, consisting of 47 member states which is all of the countries in Europe except for Belarus & most of the countries from the former USSR including Russia, Azerbeijan, etc. Overall, the organization represents over 830 million people, along with European Union having 27 member States all of whom are a part of the Council of Europe. The instruments being prepared by the EU & the Council will hopefully become complementary, as has been seen with regards to various past documents as well, e.g. convention 108 on automated data processing (1980). It is precisely because of the combination of the instruments promulgated by the organizations which can potentially serve as global standards. Because the treaties adopted by the Council can be acceded to by Non-Member States, but Non-Member States cannot accede to the Instruments prepared by the EU. So, there's a combination that necessitates the production of complementary instruments.

The task of the team was fulfilled, ambitions of different stakeholders differ. There have been a lot of preservations on the side of the business community for example, a lot of ambition on the side of the civil society, a lot of often conflicting positions from member states, because different ministries & different members might have

different ambitions. A wide representation was a very fortunate factor aiding in coming up with solutions which are based on consensus & compromises allowing for further work addressing all the concerns.

How much similarity & commonality do you think exists amidst the recommendation by the UNESCO & the draft made by the CAHAI?

It is a very wide & complex subject which can reflect throughout the length of the process & it can reflect the difficulties in getting an effective instrument in the long run. Small & Medium sized businesses welcome regulations while the BigTech is waiting for the moment when to engage & it is a matter of incentives – who has a stronger position now. One of the risks that need to be taken into consideration while preparing such instruments is to prevent incumbent infringement & there is real danger that various lobbying attempts will be towards that direction, & there are concerns that the proposal by the European Commission which is a bit more concrete than what was prepared by the Council of Europe is already elaborating the obligations of AI users or developers, while putting some compliance burdens on developers which might be easier to achieve if you are a big player. But, to go back a bit, there were conventions or treaties adopted in a matter of months in the Council, especially when they related to cooperation in the need to combat terrorism. It all depends on will & complexity of the subject. When it comes to the Council, the main task is to frame the abilities, necessities & priorities appropriately, so that we do not go too far or too wide. It is up to the member states to dictate the mandates of the organizations & the instruments. The biggest task now will be to elaborate by the convention(s) what are the actual negative & positive obligations, where do we set the potential bands, & how compliance can be supervised by independent authorities. It can be seen that this

will happen simultaneously with the attempts of the Council & the Union, there's a chance of finding common solutions through risk assessment & impact assessment, this is one the biggest criticisms that the ECs' proposal received that – it did not have any methodology, when it comes to risk assessment.

What could be a preparatory model?

A step-by-step approach can be preferred, & step 1 for everybody would be transparency & not in terms of XAI Explainability or interpretability but also in terms of what tools are being used & are proposed to be used, e.g. if Police force used live facial recognition, it should be transparently explained that it's doing so & it should also preferably known what type of tool is being used & it should be verified by independent researchers & institutions. Because, we do not even know what inequalities are growing at a scale level, & when there's transparency there needs to model to assess the effectiveness of the tools & this should be something that is welcomed by institutions & private sectors (knowing in the long run that it is not selling snake oil – that it is selling something that it is actually effective & will remain so in the long run). The need to regulate AI is not an imperative but what is an imperative is the need to regulate the ecosystem in which it is being developed.

6 On the Pendency of Cases in the Supreme Court of India, with Ayan Chandra & Shubham Pandey, IIT Kharagpur

Mridutpal Bhattacharyya

Deputy Managing Editor

Abstract. This is an interview conducted by Abhivardhan, the Editor-in-Chief, which has been transcribed and paraphrased. He interviews Ayan Chandra and Shubham Pandey from IIT Kharagpur.

How has pendency in litigations in the Hon'ble SC of India been assessed algorithmically?

In the Indian Judiciary – comprising 3 tier hierarchical structure with the SC at the apex, followed by High Courts in States & then the subordinate courts. If a glance it afforded to the number of cases pending at these 3 levels, the figure stands at a whopping 4.3 Crores as of now. These cases, most of them are pending for 5+ years & in some cases the pendency is of 20+ years, & that includes Criminal cases as well. The identification of the bottlenecks & the addressing of them with technology is important.

Initially, all the prior research work on the pendency area were of majorly two categories – one was people trying to understand whether there is a claim or measured claim, whether there's a premise where they are attacking or supporting each other (Argument Structure), while the other was a rhetorical structure. Whereas, none of the structures categorically addressed the underlying structures of

the judgments. This could be seen most prominently with respect to the Supreme Court Judgments while a huge level of pendency is prevalent. It was identified that – finding a relevant precedent for other cases (Legal Research) is something that practitioners want to experience in a low latency setting, giving it as little time as possible. People want to go through the entire judgment with a certain precipice or context in mind, & identify the same in the judgement as quickly as possible. After due discussions with legal minds of eminence, a particular structure was defined & laid out & a pipeline method was suggested – as a combination of rules & different NLP tools, in a matter of saying. The pipeline worked quite well, & initially the environmental pollution (air pollution) cases were taken, wherein, about 28 cases against the Central or respective State Governments were taken, while 24 were analyzed & annotated via the pipeline & the same were evaluated on the basis of the remaining 4 cases, it was found that the pipeline worked well.

The distinction between enumerated & unenumerated facts needs to be made clear & explicit while on the other hand the unenumerated facts need to be analyzed as well. What can be seen in the pollution cases datasets is that the Central or State Governments were parties to them, & therefore that made the cases stakeholder heavy. Now a machine would need to know to take the contentions of all the respondents & the applicants into account. In this particular case, as an initial step forward, a logic was used that – an appellant token was used, but later on when someone wants to take it forward, they can differentiate as Appellant 1, 2, etc, that option is available.

What are the future plans?

The system should not be constrained inside a particular rule set. The influence upon the system from different policies & rules &

vice versa should be explored as well. The tool can serve as a game-changer wherever technically – computationally possible. The strategic identification of the cases is also important. How long one can let the pendency go on is a crucial question, because pendency will forever be there but the floodgates need to be closed at some time & the previous pile needs to be cleared.

The system would pick & choose as per parameters all the required information including precedents cited/relied upon, arguments, issues, etc. It is a form of dealing with cases in the Court – a module merely, but maybe someday it will change the system. Who knows!