Volume 2, Number 1, Spring 2018

Artificial Intelligence in Court

Legitimacy Problems of AI Assistance in the Judiciary

Thomas Julius Boucz*

The emergence of Artificial Intelligence (AI) is changing many aspects of our lives. The judiciary is not excluded from this development. However, the use of AI today and in the foreseeable future is limited to specific tasks, whereas the work of a judge requires a broad range of different skills. Therefore, it is unlikely that the use of AI will completely replace the work of human judges from one day to another. Introducing AI assistance to the judiciary will most likely be a gradual and slow process that starts with the parallel existence of AI assistance and human judges. This article investigates the danger of hidden relinquishment of decision-making power during this phase. For this purpose, I present three different ways how AI assistance could be incorporated into the judiciary. Analysing these three scenarios, I conclude that it is separation rather than cooperation that prevents hidden relinquishment of decision-making power.

1. Introduction[1]

The emerging technology Artificial Intelligence (AI)[2] has become a highly relevant topic for many parts of our lives and has been popular in the media for years. Some

* Law Student, Institute of Law, University of Natural Resources and Life Sciences, Vienna

[1] This article is an adapted version of a paper I wrote during my exchange semester at the University of Copenhagen in Spring 2017. I would like to thank Theresa Upperton, Lisa Müllner and Iris Eisenberger for their valuable inputs.

[2] Following the terminology in Ian Kerr and Carissima Mathen, 'Chief Justice John Roberts is a Robot' [2013] < http://robots.law.miami.edu/2014/wp-content/uploads/2013/06/Chief-Justice-John-Roberts-is-a-Robot-March-13-.pdf > accessed 25 March 2018, I use the abbreviation 'AI' to refer to the field of artificial

of the most prominent examples associated with AI are inventions like IBM's *Watson* (known for winning the US quiz show Jeopardy)[3] or DeepMind's *AlphaGo* (known for winning a game of Go against world champion Lee Sedol).[4]

1.1 Clear Terminology

Due to the popularity of the topic, the term 'artificial intelligence' is widely known and heavily loaded with different preconceptions and assumptions. Therefore, it is important to provide clear terminology from the beginning when discussing aspects of AI: For the sake of this article, artificial intelligence is a non-biological autonomous entity; autonomy being the ability to give rules (ancient greek: *nomos*) to oneself (ancient greek: *auto*).[5] Furthermore, when dealing with the parallel existence of AI assistance and human judges – following Jason Millar's and Ian Kerr's definition – I use the term 'Co-Robotics' to refer to any situation in which human and AI experts are working alongside one another.[6]

---

intelligence, 'an AI' or 'the AI' to refer to a particular entity and 'AIs' to refer to the plural.

[3] < www.ibm.com/watson/ > accessed 25 March 2018.

[4] Recently, DeepMind's newest program *AlphaGo Zero* has beaten *AlphaGo* 100 times in 100 games. In contrast to *AlphaGo*, *AlphaGo Zero* was not trained by supervised learning from human expert moves, but only by playing against itself; cf David Silver et al, 'Mastering the game of Go without human knowledge' [2017] 550 Nature 354.

[5] For other definitions see: Shane Legg and Marcus Hutter, 'A Collection of Definitions of Intelligence' [2007] 157 Frontiers in Artificial Intelligence and Applications 17. See also: Ugo Pagallo, *The Laws of Robots: Crimes, Contracts, and Torts* (Law, Governance and Technology Series 10, Springer 2013) 2ff; Neil M. Richards and William D. Smart, 'How Should the Law Think About Robots?' in Ryan Calo, A. Michael Froomkin, Ian Kerr (eds), *Robot Law* (Edward Elgar Publishing 2016) 5-7.

[6] cf Jason Millar and Ian Kerr, 'Delegation, relinquishment, and responsibility: The prospect of expert robots' in Ryan Calo, A. Michael Froomkin and Ian Kerr (eds), *Robot Law* (Edward Elgar Publishing 2016) 104.

1.2 AI and the Law

That AI gets a lot of attention should not be a surprise, since it is changing so many aspects of our lives already. Entire branches such as transportation, health care, education and entertainment are subject to revolutionary changes due to the emergence of AI.[7]

The law is by no means excluded from this development. A great deal of legal work consists of monotonous activities such as sifting of documents, searching for irregularities in large amounts of data and analysing numerous cases. Law firms apply e-discovery software[8] to cases that involve many documents to be screened. Moreover, Casecrunch's[9] *CaseCruncher Alpha* is already predicting judicial decisions with high accuracy.[10]

Technology-assisted legal review is also starting to receive judicial stamps of approval.[11] During a visit at Rensselaer Polytechnic Institute in April 2017, John Roberts, Chief Justice of the Supreme Court of the United States, was asked whether he could foresee a day, when AIs would assist with courtroom fact-finding or, more

---

[7] For an overview see: Stone P et al, 'Artificial Intelligence and Life in 2030.' One Hundred Year Study on Artificial Intelligence: Report of the 2015-2016 Study Panel' (Stanford University, September 2016).

[8] See: John Markof, 'Armies of Expensive Lawyers, replaced by Cheaper Software' *New York Times* (New York, 4 March 2011) <www.nytimes.com/2011/03/05/science/05legal.html>; Hug Son, 'JPMorgan software does in seconds what took lawyers 360,000 hours' *The Independent* (London, 28 February 2017) <www.independent.co.uk/news/business/news/jp-morgan-software-lawyers-coin-contract-intelligence-parsing-financial-deals-seconds-legal-working-a7603256.html> all accessed 25 March 2018.

[9] See: <www.case-crunch.com> accessed 25 March 2018.

[10] In October 2017, *Casecruncher Alpha* won a week-long competition against human commercial lawyers with an accuracy of 86.6% of the predictions made. The human lawyers achieved an accuracy of only 62.3%, see: *Robotics Law Journal,* 'CaseCrunch Lawyer Challenge: Results' (London, 1 November 2017) <www.roboticslawjournal.com/news/casecrunch-lawyer-challenge-results-85492553> accessed 25 March 2018.

[11] Dean L Dalke, 'Can Computers Replace Lawyers, Mediators and Judges?' (2013) 9 The Advocate, 703, 704 with further references.

controversially even, judicial decision-making.[12] 'It's a day that's here,' he said, 'and it's putting a significant strain on how the judiciary goes about doing things.'[13]
A study[14] titled 'Extraneous factors in judicial decisions', conducted in Israel in 2010, has shown the following:

> [J]udges were much more likely to accept prisoners' requests for parole at the beginning of the day than at the end. Moreover, a prisoner's chances of receiving parole more than doubled if his case was heard at the beginning of one of the three sessions, rather than later on in the session. (...) As a case study, one of the judges started in the morning by granting parole to about 65 percent of the prisoners; that percentage dropped to near zero by the end of the first session, then rebounded to about 65 percent after the snack break. The same pattern repeated in the second and third sessions.[15]

The use of AI in the judiciary might help with minimising the influence of extraneous factors such as weariness and emotional instability. However, AI decision-making might reveal different human-made, structural biases that originate from the legal system, the AI's training data or the AI's programming itself.[16] For

---

[12] Adam Liptak, 'Sent to Prison by a Software Program's Secret Algorithms' *New York Times* (New York, 1 May 2017) <www.nytimes.com/2017/05/01/us/politics/sent-to-prison-by-a-software-programs-secret-algorithms.html?_r=0> accessed 23 March 2018.
[13] Ibid.
[14] Shai Danzigera et al., 'Extraneous factors in judicial decisions' [2011] 108/17 PNAS 6889 <www.pnas.org/content/108/17/6889.full.pdf> accessed 25 March 2018.
[15]Kate Shaw, 'To Get Parole, Have Your Case Heard Right After Lunch' *Wired* (Ars Technica, 4 November 2011) <www.wired.com/2011/04/judges-mental-fatigue/> accessed 25 March 2018.
[16]Sandra Wachter et al, 'Transparent, explainable, and accountable AI for robotics' [2017] 2/6 Science Robotics eaan6080. See also: Christian Joachim Gruber and Iris Eisenberger, 'Wenn Fahrzeuge selbst lernen: Verkehrstechnische und rechtliche Herausforderungen durch Deep Learning?' in Iris Eisenberger et al. (eds), *Autonomes Fahren und Recht* [2017] 51, 57ff; Niki Kilbertus et al., 'Avoiding Discrimination through Causal Reasoning' [2017] 1ff <https://arxiv.org/pdf/1706.02744.pdf>; Brent Mittelstadt et al, 'The ethics of algorithms: Mapping the debate' (2016) July-Dec, Big Data & Society 7; Sue

example, the risk assessment software *Correctional Offender Management Profiling for Alternative Sanctions – COMPAS,* that was used for predicting the likelihood of defendants committing a future crime, was racially biased against African American defendants.[17]

2. Problem Outline

This article does not address the question whether to put an AI on the judge's bench. This might be a question of the future. When this time comes, society will want to decide whether to allow AIs to deliver judgements.[18] Far-reaching consequences of such choices should be discussed publicly and subjected to a democratic decision-making process. But what if it never comes to this? What if we will never have that public discourse? What if we assume the time to decide on that matter has not come yet, when in reality we have already chosen a possibly irreversible path much earlier without ever properly addressing it?

Newell and Marco Marabelli, 'Strategic opportunities (and challenges) of algorithmic decision-making: A call for action on the long-term societal effects of "datification"' [2015] 24 Journal of Strategic Information Systems 3, 6 all accessed 25 March 2018.

[17] See: Julia Angwin et al, 'Machine Bias. There is software that is used across the county to predict future criminals. And it is biased against blacks' *ProPublica* (New York, 23 May 2016) <www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing>; Anthony W Flores et al, 'False Positives, False Negatives, and False Analyses: A Rejoinder to "Machine Bias: There's Software Used across the Country to Predict Future Criminals. And It's Biased against Blacks"' (2016) 80 Federal Probation 38. See also: Max Ehrenfreund, 'The machines that could rid courtrooms of racism' *The Washington Post* (Washington, DC, 18 August 2016) <www.washingtonpost.com/news/wonk/wp/2016/08/18/why-a-computer-program-that-judges-rely-on-around-the-country-was-accused-of-racism/?utm_term=.be45d438bb02> all accessed 25 March 2018.

[18] For further discussion of increasing societal acceptance of innovative technology through democratisation see: Iris Eisenberger, *Innovation im Recht* (Verlag Österreich 2016) 284ff.

2.1 'Special-purpose human beings'

Technology is not yet ready to produce an AI with a skill-set that is broad enough for the work of a judge.[19] Not because the necessary intellectual tasks are too complex to be processed by an AI, but rather because the work of a good judge consists of a mix of skills including research, language, logic, creative problem solving and social skills.[20] AIs of the near future excel in some of these individual tasks separately, but not yet in combining them.[21] Jack M Balkin describes AIs in this phase as 'special-purpose human beings':

> A key feature of robotic substitution is that it is partial. Robots and AI entities take on particular aspects and capacities of persons (…). This is what I mean when I say that robots and AI agents operate as "special-purpose human beings"; they are agents for a particular reason or function, straddling the line between selves and tools, or persons and instruments.[22]

That is why AI's entry into the judiciary will most likely be a gradual and slow process that starts with the parallel existence of AI assistance and human judges. This article

---

[19] See: Robin Hanson, *The Age of Em. Work, Love, and Life when Robots Rule the Earth* (Oxford University Press 2016) 271ff.

[20] Clemens Jabloner, 'Richter im Zwiespalt' in Brigitte Schenk et al (eds), *Festschrift für Irmgard Griss* (Jan Sramek Verlag 2011) 315, 316ff; Richard A Posner, *How Judges Think* (First Harvard University Press 2010) 19ff. See also: Iris Eisenberger, 'Richter und Rechtspfleger' in Michael Holoubek and Michael Lang (eds) *Grundfragen der Verwaltungs- und Finanzgerichtsbarkeit* (Linde Verlag 2017) 45, 48; Klaus Rennert, 'Funktionen und Legitimation des Richters' [2015] 5ff < www.bverwg.de/medien/pdf/rede_20151127_hu_berlin_rennert.pdf > accessed 25 march 2018; Jeremy Waldron, 'Judges as moral reasoners' [2009] 7/1 International Journal of Constitutional Law 2 all.

[21] For example: IBM's *Watson* playing Jeopardy (n 3); DeepMind's *AlphaGo Zero* playing Go (n 4); Joshua Browder's *The World's First Robot Lawyer* drafting very specific types of legal documents, see: <www.donotpay.co.uk> accessed 25 March 2018; *Casecruncher Alpha* predicting judicial decisions (n 9).

[22] Jack M Balkin, 'The Path of Robotics Law' [2015] 6 California Law Review 45, 59. In this context see also: Kerr and Mathen (n 2) 7: 'What will happen is that, one day soon, an AI will be developed that regularly and reliably outperforms a professional at some complex task traditionally requiring human expertise.'

investigates the danger of hidden relinquishment of decision-making power during this phase.[23]

Hidden relinquishment of decision-making power is particularly problematic in the judiciary: When an elected representative appoints a judge, the representative equips the judge with the necessary democratic legitimacy to rule on cases. The judge may not delegate the judgment to someone else. If they choose to delegate a specific administrative or legal task to an assistant, the judge needs to stay in the loop by ensuring functioning control and communication between them and their assistants, no matter if a human or an AI provides the assistance.

2.2 Co-Robotics in the judiciary

The key challenge of Co-Robotics in the judiciary is facilitating functioning communication between human and machine. Clemens Jabloner provides a useful metaphor to illustrate the underlying problem:

> Imagine a tinsmith's warehouse, containing thousands of different kinds of nails and screws. In order to find specific sizes and kinds of them, the human warehouse worker needs a storage system. For example: Screws go on the left shelf, nails on the right.
>
> Identical screws and nails are collected in little boxes and placed in order of their size, starting with the smallest type. The machine, however, being able to assign a specific location to each and every single screw and nail, will distribute them in a system that ensures a minimum of distance to be covered based on the following (thousand, for example) requests. Therefore, one finds the system in an ever-changing state. It appears to be pure chaos in the eye of the human observer and yet the AI delivers the right kind of nail or screw every time and much faster than the human ever could.[24]

One can ask the human worker what line of thought led them to their result, initiating a conversation within a common way of thinking. Understanding the

---

[23] For further discussion of relinquishment of decision making power in the context of technology law see: Eisenberger, *Innovation* (n 18).

[24] Clemens Jabloner, University of Vienna, used a similar metaphor in one of his lectures I attended in autumn semester 2016/17. I would like to thank him for the inspiration that he gave me throughout this lecture.

reasoning of an AI, on the other hand, necessitates a considerably bigger effort.[25] That might not seem problematic regarding the quest for screws and nails, especially if the AI brings you the right one. However, the 'products' of the judiciary include judgements that order consequences such as large fines, prison sentences and granting or denial of parole[26] and the object of the judgement is human behaviour. In this regard, understanding the AI's reasoning can be crucial for the acceptance of the technology.[27]

There are two approaches to the challenge of Co-Robotics in the judiciary. In order to retain human control, one has to either enable functioning communication between human and machine (addressing the Co-Robotics problem) or strictly separate them from each other (avoiding the Co-Robotics problem).

2.3 Addressing the Co-Robotics problem: Explanation

When addressing the problem of Co-Robotics, it is important to distinguish between transparency regarding the general functioning of a machine and an explanation of a specific machine-generated decision.[28]

Transparency can only be achieved at the expense of performance. Applying this thought to the warehouse metaphor, this means that instead of establishing a wholly different kind of storage system that maximizes efficiency, the AI is asked to find a compromise that optimises the 'human' system, but stays within its logical

---

[25] Derek Doran et al., 'What Does Explainable AI Really Mean? A New Conceptualization of Perspectives' [2017] <https://arxiv.org/abs/1710.00794>; Finale Doshi-Velez et al, 'Accountability of AI Under the Law: The Role of Explanation' [2017] <https://arxiv.org/abs/1711.01134>; Will Knight, 'The Dark Secret at the Heart of AI' *MIT Technology Review* (Boston, 11 April 2017); Sandra Wachter, et al 'Counterfactual Explanations Without Opening the Black Box: Automated Decisions and the GDPR' (2017) Harvard Journal of Law & Technology, Forthcoming <https://ssrn.com/abstract=3063289> all accessed 25 March 2018.

[26] See n 17.

[27] Eisenberger, *Innovation* (n 18) 287; Wachter et al, 'Counterfactual Explanations' (n 25) 4.

[28] Doshi-Velez et al (n 25) 6; Wachter et al, 'Counterfactual Explanations' (n 25) 25f.

dimension and its methodology at the expense of performance. This system provides a machine version of the *human* warehouse worker rather than a machine version of the *ideal* warehouse worker.

However, we do not require transparency regarding the flow of signals through neurons in a judge's brain.[29] Instead, we require a specific explanation for a decision. Explanation is about answering how certain factors were used to come to the outcome in a specific situation.[30] According to Finale Doshi-Velez et al, 'it is technically feasible to extract the kinds of explanations that are currently required of humans from AI systems'.[31] Facilitating such explanations might require an explanation system distinct from the AI system.[32] Furthermore, Wachter et al point out the importance of counterfactuals explanations in this.[33]

2.4 Avoiding the Co-Robotics problem: Separation

The separation of AI assistance and human judge avoids problems of Co-Robotics, but raises questions about the fate of purely machine-generated decisions. Returning to the metaphor one more time: There are now two warehouses. A human worker administers one, an AI the other one. Do customers have a choice between those two warehouses? What happens if the human worker and the AI deliver different results? Who has the last word? This metaphor allows the best aspects of both options to co-exist, but raises further questions.

---

[29] Doshi-Velez et al (n 25) 6f.

[30] Ibid 7.

[31] Ibid 7.

[32] Ibid 7.

[33] Wachter et al, 'Counterfactual Explanations' (n 25) 5f: 'Counterfactual explanations take a similar form to the statement: "You were denied a loan because your annual income was £30,000. If your income had been £45,000, you would have been offered a loan." Here the statement of decision is followed by a counterfactual, or statement of how the world would have to be different for a desirable outcome to occur.'

3. Selected Designs of AI Assistance in the Judiciary

When discussing problems of the future, such as AI assistance in the judiciary, it is necessary to illustrate the central thoughts sufficiently. Neil M Richards and William D Smart emphasise the importance of metaphors at the nexus between law and technology:

The experience of cyberlaw and other areas of technology-influenced jurisprudence has revealed one particularly important lesson for technologically sophisticated applications of law – when it comes to new technologies, *applying the right metaphor for the new technology is especially important.* How we regulate robots will depend on the metaphors we use to think about them.[34]

In the light of the importance of metaphors in this context, I present a selection of three different assistive roles that an AI could assume in the judiciary. This chapter discusses an AI (1) as a Librarian, (2) as an Advocate General and (3) as an Official with Limited Judicial Powers. The three roles gradually increase both in the required technological complexity of the AI and the degree of institutional separation.

3.1 Librarian

The Librarian's job is to find requested literature. Judges will regularly ask for literature relevant to a specific case. The nature of this task can differ, depending on the input's level of abstraction. It may vary between relatively simple commands, like 'find document x by author y', and complex requests, such as 'find all relevant literature for case z and list them by relevance'. An AI that operates on 'mixed initiative'[35] (receiving high-level commands and operating autonomously only on a lower level) poses different questions than an AI that requires nothing more than the facts of a case to do legal research completely autonomously.

---

[34] Neil M Richards and William D Smart, 'How Should the Law Think About Robots?' in Ryan Calo, A Michael Froomkin and Ian Kerr (eds), *Robot Law* (Edward Elgar Publishing 2016) 13; emphasis added. See also: ibid. 16-18.

[35] Richards and Smart (n 34) 8 use the Mars Exploration Rovers as example for 'mixed initiative'.

3.1.1 Legal research and decision-making

Legal research cannot be reduced to merely collecting somehow related cases or articles. Expertise in legal research resides in the connections and links drawn between the individual pieces of information.[36] The way information is connected and structured itself influences further research. In other words, finding information and transforming it into legal expertise is a reciprocal process. Therefore, in most cases, delegating legal research causes relinquishment of at least some decision-making authority.

However, this is not particularly new. There has always been relinquishment of decision-making authority in legal research; what has changed – and is currently changing again – is the entity that judges delegate the act of research to: the 'gate-keepers of information'.[37]

3.1.2 The trend of increased performance at the expense of transparency

Before judges used the Internet for their legal research, they provided a human Librarian with information, who then came up with suggested results. The judges did not know how the Librarians came to their results, but they had the opportunity to ask them and engage in a professional discussion.

Today, most legal research does not happen in dim and dusty law libraries but on various legal databases online, searchable through sophisticated algorithms.[38] An algorithm is 'a set of step by step instructions, to be carried out quite mechanically,

[36] cf Millar and Kerr (n 6) 109ff with further reference to Harry Collins, *Changing Order: Replication and Induction in Scientific Practice* (University of Chicago Press 1992).

[37] Iris Eisenberger, 'Die Macht der Algorithmen' [2011] 4 Juridikum 517, 521.

[38] Examples: Beck Online <https://beck-online.beck.de/Home>; HeinOnline <https://home.heinonline.org>; LexisNexis <www.lexisnexis.com/en-us/home.page>, Westlaw <http://legalsolutions.thomsonreuters.com/law-products/westlaw-legal-research/>; all accessed 25 March 2018).

so as to achieve some desired result'.[39] Although an algorithm's parameters are always the same, constantly changing sets of data can make it very complex and almost unpredictable.[40] It is more effective in terms of capacity and speed, but less transparent than a human librarian. Additionally, search algorithms of legal databases are very often trade secrets and therefore lack transparency on purpose.[41]

AI is the next stage in this development and amplifies the trend of increased performance at the expense of transparency: An entity capable of altering its own search parameters can carry out legal research even more efficiently. On the other hand, it is very challenging to provide the user with transparency regarding its functioning. As the librarian entity evolves from human through algorithm to AI, it gets harder and harder to ensure functioning communication between judge and librarian. Therefore, facilitating explanations of an algorithm's or an AI's reasoning is crucial in this regard.

I have argued that legal research and decision-making are interdependent. If the input reaches a certain level of abstraction, the AI has no other choice but to search for the requested literature by simultaneously drafting a decision. One example would be the following input: 'Find the most relevant sources for case x and list them in order of their appearance in the (future) decision draft.' In this case, the AI Librarian cannot list the sources in the right order without anticipating the structure of the judgement itself. It would be required to form a path of bread crumbs, leading the judge to a pre-determined, machine-generated result.[42]

In conclusion, assigning legal research to an AI results in an intransparent form of delegation of decision-making authority whose extent is hard to predict.

---

[39] Jean-Luc Chabert, 'Introduction' in Jean-Luc Chabert (ed), *A History Of Algorithms: From The Pebble To The Microchip* [1999] 1. See also: Thomas H Cormen, *Algorithms Unlocked* (MIT Press 2013) 1ff.

[40] Wachter et al 'Counterfactual Explanations' (n 25) 6.

[41] Eisenberger, *Innovation* (n 18) 287.

[42] I would like to thank Hin-Yan Liu, University of Copenhagen, for this idea and many fruitful discussions in class.

3.2 Advocate General

Article 49 of the Statute of the Court of Justice of the European Union[43] reads as follows:

> It shall be the duty of the Advocate General, acting with complete impartiality and independence, to make, in open court, reasoned submissions on certain cases brought before the General Court in order to assist the General Court in the performance of its task.

The Advocates General assist the CJEU. Their role is to propose a legal solution to the cases for which they are responsible:

> He or she analyses in detail the legal aspects of the case and suggests completely independently to the Court of Justice the response which he or she considers should be given to the problem raised. If it is decided that the case raises no new question of law, the Court may decide, after hearing the Advocate General, to give judgment without an Opinion [by the Advocate General in writing].[44]

3.2.1 The advantage of transparency

For this scenario, I apply the concept of the Advocate General of the CJEU to the judiciary in general. Just like the AI Librarian, the scenario of an AI Advocate General addresses the problem of Co-Robotics, because the AI and the judge are working on the same case. Contrary to the metaphor of the AI Librarian, it provides transparency to the general public about the extent of AI usage.

The judge can either incorporate the AI's reasoned submission (or part of it) by citing it or choose to ignore or dissent from the AI's opinion. The nature of the submission is auxiliary and optional. It provides supporting knowledge and counterfactual

---

[43] Consolidated Version of the Protocol (No 3) on the Statute of the Court of Justice of the European Union, annexed to the treaties [2016] OJ C202/01.

[44] Court of Justice of the European Union, 'Presentation of the Court' <https://curia.europa.eu/jcms/jcms/Jo2_7024/en/#composition> accessed 25 March 2018.

explanations[45] that the judge may or may not take into account. Kerr and Mathen emphasis the necessity of designing the AI as assistive tool:

> Undoubtedly, if and when future lawyers or (perhaps one day) judges actually begin to delegate significant legal tasks or decision making to AIs, the profession will require that these AIs are utilized *merely as assistive tools* that help lawyers or judges carry out their responsibilities, *not as replacement* for them.[46]

3.2.2 Evidence-based reasoning as a slippery slope

But is it even possible to utilise an AI Advocate General merely as assistive tool? It is imaginable that social dynamics would guide the judge's use of the AI's submissions. It would show what influence AI usage has on the general public's perception of the judge's judicial ability. For example, on the one hand, it can be a sign of self-confidence if a judge cites the better part of the AI's submission in order to increase efficiency. On the other hand, it is possible that judges who tend to resort to the AI's assistance more often than others will be thought of as persons with inferior competence.

> However, the human judge is likely to incorporate the AI's opinion rather than dissent from it, because statistics will suggest that decisions with a higher percentage of machine-generated content are less likely to be overturned by a higher instance court.[47] Jason Miller and Ian Kerr explain this logic of evidence-based reasoning:

The underlying rationale for adopting evidence-based practice has to do with the normative pull of evidence that, for the most part, is self-evident: if the best available evidence suggests that option x is the most likely to produce desirable outcomes, then one ought to pursue option x. (…) Once there are expert robots, it will be easier to argue in some instances that they ought to be used to their full potential, because the *evidence will suggest that* in those instances *they will, on average, deliver better*

---

[45] See: n 33.

[46] Kerr and Mathen (n 2) 7; emphasis added.

[47] Kerr and Mathen (n 2) 8. For further discussion of the human tendency to quickly rely on automation see: Kate Goddard et al, 'Automation bias: a systematic review of frequency, effect mediators, and mitigators' (2012) 19 JAMIA 121.

results than human experts. It will likewise be harder to argue that they ought not to be used to their full potential.[48]

Eventually, the judge could become nothing more than a conduit for delivering purely machine-generated decisions as the percentage of machine-generated content in their decisions increases constantly.[49]

In conclusion, an AI Advocate General provides more transparency than the AI Librarian. Nevertheless, the slippery slope of decision-making relinquishment due to evidence-based reasoning remains.

3.3 Official with Limited Judicial Powers

The scenario of an AI Official with Limited Judicial Powers[50] separates the human judge's work and the AI's work from each other. Applicants would be required to bring a case before the AI official before appealing to the human judge. The AI would have the judicial power to autonomously decide cases. However, lodging an appeal against its decision would automatically invalidate it. The human judge would then decide anew.

3.3.1 Institutional separation

Contrary to the AI Advocate General the strict institutional separation gives applicants the possibility to opt-out of effective AI involvement in their case. Furthermore, this scenario avoids the slippery slope of a human judge relinquishing decision-making power in close cooperation with an AI due to evidence-based reasoning. Furthermore, it avoids the described problems of the incompatibility of human thinking and machine reasoning, thus presenting an alternative to facilitating understandable explanations of an AI's reasoning. In this scenario, the AI functions as a black box that turns the facts of the case (input) into a legal consequence

[48] Millar and Kerr (n 6) 116-117; emphasis added.

[49] Kerr and Mathen (n 2) 8.

[50] Eisenberger, 'Rechtspfleger' (n 20) 50ff.

(output).[51] Explanations (for example provided by a distinct explanation system)[52] are replaced with the option of requesting a subsequent, purely human-generated decision on the same case.

3.3.2 Plain Cases and Hard Cases

If the human judge and the AI Official disagree over a case, it can be only due to one of three reasons: The AI made a mistake,[53] the human judge made a mistake, or the case allows more than one correct interpretation.

At this point, one has to distinguish between plain cases and hard cases. A judge is confronted with plain cases, 'where the general terms seem to need no interpretation and where the recognition of instances seems unproblematic or "automatic" (…) where there is general agreement in judgements as to the applicability of the classifying terms'.[54] Hard cases, on the other hand, 'arise when the answer demanded by law is unclear, unknown or deeply disconcerting. Perhaps the relevant rules pull in opposite directions or the case is the first of its kind.'[55]

Evidence-based reasoning has a different – even constructive – effect in this regard: To save legal costs, appellants will eventually bring only hard cases before a human judge, because evidence will suggest that AI decisions over plain cases get overturned only very few times. After all, the AIs are being built to outperform human beings in applying clear rules.

---

[51] Doshi-Velez et al (n 25) 7. See also: Jenna Burrell, 'How the machine ″thinks‴: Understanding opacity in machine learning algorithms' Jan-June [2016] Big Data & Society 1ff; Gruber and Eisenberger, 'Deep Learning' (n 16) 57ff; Wachter et al., 'Counterfactual Explanations' (n 25) 11.

[52] Doshi-Velez et al (n 25) 7

[53] See: Stephen Baker, 'How Could IBM's Watson Think That Toronto Is a U.S. City?' *Huffington Post* (New York, 16 February 2011) <www.huffingtonpost.com/stephen-baker/how-could-ibms-watson-thi_b_823867.html> accessed 25 March 2018.

[54] HLA Hart, *The Concept of Law* (2nd edn, Clarendon Press 1994) 123.

[55] Kerr and Mathen (n 2) 31. See also: Ibid. (fn 137).

In the few cases where an AI makes an obvious mistake in deciding a plain case, the human judge acts as a safety net that corrects the error, which otherwise could possibly result in public rejection of AI in the judiciary.[56]

This scenario of an AI Official with Limited Judicial Powers is compatible with evidence-based reasoning. The AI is able to decide on plain cases much faster and more accurately than a human being ever could, thereby increasing judicial efficiency.

However, by automatically invalidating the AI's decision when challenged, this scenario has so far avoided the following question: Who is to be followed when there is disagreement over a hard case between a human judge and an AI?

It is a defining feature of hard cases that they do not provide only one correct solution. One could of course interpret a decision being overturned by a higher instance court as wrong decision in retrospect, but this argument is only valid until one turns to the highest court and its decisions itself. Therefore, evidence-based reasoning might not apply to hard cases.

On the other hand, both the human judge and the AI could both make incorrect decisions. Again, statistics will suggest that it is more likely that the human judge gets it wrong than the AI. However, in contrast to plain cases, there is no way of proving the AI wrong, because it cannot be ruled out that the AI has found another correct answer.

In this scenario, the AI functions as a black box that does not provide an explanation for its decision. However, the explanation is exactly what matters most when turning to (possibly) equally correct results. Therefore, when it comes to hard cases, an understandable justification is more important than a higher probability of correctness. Consequently, the superiority of AI adjudication in this scenario ends where hard cases arise – the limiting factor not being technology but humanity.

[56] See: n 53.

4. Conclusion

The parallel existence of AI assistance and a human judge poses the danger of hidden relinquishment of decision-making power.[57] In order to prevent this, it is important to discuss in time how AIs assistance could be incorporated into the judiciary. The crucial question in this discussion is whether to let the AI and the human judge work alongside one another or to separate them institutionally. In this regard, one of the key issues is facilitating functioning communications between human and AI.

The metaphor of the *AI Librarian* shows that the work of a judge consists of intertwined tasks that cannot be detached from each other completely. It is therefore not possible to simply split the profession of a judge into decision-making and non-decision-making parts. Following this line of thought, including an AI in the process of adjudication is always associated with relinquishment of decision-making authority to some extent.

On this basis, the scenario of the *AI Advocate General* teaches that transparency about the division of responsibility between an AI and a human judge does not prevent loss of human control. Due to evidence-based reasoning, cooperation in the same case between a human judge and an AI is the first step on a slippery slope of relinquishing decision-making power – independent of the degree of transparency provided.

However, a clear separation of human-generated decisions on the one hand and machine-generated decisions on the other hand allows us to channel AI adjudication. The metaphor of the *AI Official with Limited Judicial Powers* corresponds with the idea of such institutional separation. It would allow the AI, functioning as black box, to decide on cases on its own. In order to balance the lack of explanation, a subsequent decision by a human judge would invalidate the AI's decision. Regarding plain cases this is a mere safety measure in case of an AI error. However, when it comes to hard cases, there is more than one correct solution. When turning to (possibly) equally correct results, the reasoning matters most. Regarding hard cases, a human-generated decision that provides justification is worth more than a machine-generated decision that lacks explanation, even if it is more likely to be correct.

---

[57] See: n 23.

In conclusion, institutional separation allows us to delegate plain cases to an AI without relinquishing control. Delegating hard cases to an AI requires both effort to facilitate explanations of an AI's decision and caution to prevent slipping on the slope of relinquishing control due to evidence-based reasoning.

Generally, introducing AI to the judiciary is a highly delicate matter. It requires precaution and patience rather than hurried efficiency optimisation at all cost. Most importantly, it requires transparency:[58] Divergence between appearance and reality regarding the extent of AI's decision-making power in the judiciary makes a public dialogue impossible.[59] We cannot discuss a problem we are unaware of. But we need to have this discussion. And we need to have it soon.[60]

[58] Wachter et al., 'Transparent AI' (n 16) 2.

[59] Eisenberger, *Innovation* (n 18) 320. Also see: Tina Ehrke-Rabel et al, 'Bitcoin-Miner als Prosumer: Eine Frage staatlicher Regulierung? Dargestellt am Beispiel des Glücksspielrechts' [2017] 3 ALJ 188, 223; Iris Eisenberger, 'Digitalisierung und Selbstbestimmung' [2017] 2 ALJ 140, 149.

[60] ibid.