Artificial Intelligence in the legal sector: Ethics on the spot

Mehmet B. Unver[*] and Lezel Roddeck[**]

Abstract

Artificial Intelligence (AI), while transforming software systems and products that are used by legal practitioners, raises several ethical issues concerning transparency, fairness, and accountability of legal tech tools powered by AI. Such issues concern whether professional conduct duties for lawyers respond to the overall ethical challenges and to what extent legal ethics need to interact with AI ethics. In response to these questions, this study upholds a broader viewpoint through which not only the professional conduct duties but also the regulatory landscape and principles on AI ethics is assessed. Furthermore, the AI life cycle including the stages of design, development, and deployment are discussed with a view to eliciting a holistic ethical viewpoint and strategy applicable for the legal sector. Overall, the interaction between the AI life cycle and the lawyers' use of AI is put into inquiry within the meaning of how to cope with the ethical challenges. It is concluded that the current professional conduct duties and ethical responsibilities need to be reviewed and revised to leverage ethical AI during its life cycle and beyond. What's more, all AI stakeholders, e.g., lawyers and legal tech companies, need to cross their boundaries, engage with ethical issues from a holistic viewpoint, and collaborate with each other. Ultimately, for a fruitful collaboration to realise these aims for the legal sector, regulatory bodies such as SRA should take the leading role.

Key words: Artificial Intelligence, AI ethics, legal ethics, professional conduct rules, legal sector, transparency, accountability, fairness.

1. Introduction

"For the rational study of the law, the black-letter man may be the man, but the man of the future is the man of statistics".[1] As foretold, the transformation of the legal industry through the application of quantitative legal methods to law, 'jurimetrics, has arrived.[2] Disruptive

[*] Lecturer in law, University of Hertfordshire, Hertfordshire Law School; *e-mail address: m.unver@herts.ac.uk.*

[**] Programme Director, Bucerius Law School, Germany; *email address*: *lezel.roddeck@law-school.de*

[1] Oliver Wendell Holmes, Jr., "The Path of the law", (1897) 10 Harvard Law Review 457, 460-461.

[2] Daniel Martin Katz, 'Quantitative Legal Prediction – or – How I Learned to Stop Worrying and Start Preparing for the Data Driven Future of the Legal Services Industry' (2013) 62 Emory Law Journal 909 <https://ssrn.com/abstract=2187752> accessed 23 March 2023.

technologies powered by Artificial Intelligence (AI) including machine learning (ML), deep learning (DL) and natural language processing (NLP) are transforming the nature of legal work.[3] As '[e]thics is the building blocks upon which our whole society is based, and business is a part of our society,'[4] a major concern is that jurimetrics may replace critical legal reasoning, potentially leading to algorithmic opacity and biases that perpetuate existing inequalities.

While existing principles for professional conduct (Conduct Rules) in the legal sector may address some of these challenges, ethical concerns arise when AI algorithms are designed and developed well before they are used. Automated processes used in legal practice, may lead to misinterpretation and/or indirect (cognitive) bias and discrimination, which typically fall outside lawyers' involvement. The limited role of lawyers in the AI life cycle raises questions about whether existing regulatory frameworks sufficiently address AI ethics in the legal sector.

Scholarly guidance on AI's use in the legal sector is scarce,[5] and regulators have yet to make specific modifications to Conduct Rules for AI. However, regulators such as the Solicitor Regulation Authority (SRA) in the United Kingdom (UK),[6] the American Bar Association (ABA) in the United States of America (US),[7] the Law Council in Australia,[8] or the Council of Bars and Law Societies of Europe (CCBE),[9] in the European Union (EU) have revised codes of conduct to tackle technology-related ethical concerns.

Ranging from non-binding self-regulatory codes of conduct to heavy, externally audited compliance requirements, AI is in the spotlight of law and policy makers. This is evident from

[3] Kevin D. Ashley, *Artificial Intelligence and Legal Analytics: New Tools for Law Practice in the Digital Age* (Cambridge University Press 2017) 11-31; Lam Chen Meng, 'Impact of Artificial Intelligence on the Professional Responsibilities of Lawyers' (2020) 37 Singapore Law Review 43, 49-52; Anthony E. Davis, 'The Future of Law Firms (and Lawyers) in the Age of Artificial Intelligence' (2020) 27(1) The Professional Lawyer <https://www.americanbar.org/groups/professional_responsibility/publications/professional_lawyer/> accessed 23 March 2023.

[4] Steve Sailah, 'Kerry Stokes: a profile of rags to riches' (The World Today Archive, 2 June 2000) <https://www.abc.net.au/worldtoday/stories/s134656.htm> accessed 23 March 2023.

[5] Drew Simshaw, Ethical issues in Robo-Lawyering: The need for guidance on developing and using Artificial Intelligence in the practice of law, Hastings Law Journal (2019) 70(173) 198.

[6] Solicitors Regulation Authority (SRA), SRA Code of Conduct for Solicitors, RELs and RFLs <https://www.sra.org.uk/solicitors/standards-regulations/code-conduct-solicitors/> accessed 23 March 2023.

[7] American Bar Association (ABA), Model Rules of Professional Conduct <https://www.americanbar.org/groups/professional_responsibility/publications/model_rules_of_professional_conduct/> accessed 23 March 2023.

[8] Law Council of Australia, Australian Solicitors' Conduct Rules <https://www.lawcouncil.asn.au/policy-agenda/regulation-of-the-profession-and-ethics/australian-solicitors-conduct-rules> accessed 13 March 2023.

[9] CCBE, 'Model Code of Conduct for European Lawyers' (CCBE.eu, 2021) <https://www.ccbe.eu/fileadmin/speciality_distribution/public/documents/DEONTOLOGY/DEON_CoC/EN_DEONTO_2021_Model_Code.pdf> accessed 23 March 2023.

recently adopted measures, either via hard law, such as the EU's Proposed AI Act,[10] or via soft law, such as the OECD recommendations,[11] which provide sector-independent guidelines and address wide-ranging issues during AI life cycle. Given the juxtaposition of rules and principles of 'AI ethics' and 'legal ethics', key questions revolving around how to develop and apply ethical AI for the legal sector. These questions include:

1. Whether the Conduct Rules and ethical standards in the legal sector (legal ethics) can cope with the challenges posed by AI? Would the recent modifications to Conduct Rules resolve all the concerns?
2. Whether current soft and hard law measures and governing mechanisms on AI governance adequately respond to the ethical issues in the legal sector?
3. Should AI ethics be taken as an issue to be handled by all the AI stakeholders including legal professionals; if so, on what grounds and to what extent?

To address these questions, this study will first establish a framework regarding the current state of AI usage in the legal sector. Then the paper will survey the globally recognized Conduct Rules within the context of legal ethics, focusing on the impact of AI on the legal profession. Next the paper considers AI ethics and its leading principles under both soft and hard law, which govern the relevant mechanisms and standards. Thereafter, the extent to which current rules and principles can successfully address the emerging challenges of AI will be evaluated from a holistic point of view, incorporating potential ways in which AI ethics and legal ethics can interact.

Finally, it is concluded that current professional conduct duties and ethical responsibilities must be revisited and revised from a *holistic* perspective. In this context, ethics is viewed as a subject-matter entailing not only 'use' but also 'design, development and deployment' of AI. Therefore, a new strategy is needed to effectively manage AI transparency, accountability and fairness during and after the AI life cycle in tandem with the legal ethics. Furthermore, stakeholders representing both 'legal ethics' and 'AI ethics' need to collaborate by crossing

[10] Council of the European Union, 'Proposal for a Regulation of the European Parliament and of the Council laying down Harmonised Rules on Artificial Intelligence (AI Act) and amending certain Union legislative acts', 25 November 2022 <https://data.consilium.europa.eu/doc/document/ST-14954-2022-INIT/en/pdf> ('EU's Proposed AI Act') accessed 23 March 2023.

[11] OECD, Recommendation of the Council on Artificial Intelligence, <https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0449> accessed 23 March 2023.

their boundaries, engaging with ethical issues from a holistic viewpoint, and interacting with each other from the beginning. Finally, regulatory bodies should take the leading role in promoting fruitful collaboration that leverages ethical AI.

2. AI in the legal sector

Artificial intelligence refers to "the science and engineering of making intelligent machines".[12] Today, the term is commonly used to describe the use of technological applications for automation to execute tasks typically requiring human intelligence.[13] AI presently used in the legal sector is known as Weak AI.[14] Learning algorithms are trained to perform specific tasks, like e-discovery, legal research, document analysis, due diligence, and prediction of case outcomes. Such AI tools are not yet able to imitate sophisticated cognitive processes, such as logical reasoning, comprehension or metacognition that are vital to legal thinking. Consequently, ethical issues arise in relation to design choices or data (quality) governance in the training, testing or validation processes. Furthermore, the way lawyers use AI tools can potentially affect the outputs derived from them. Such ethical issues will be further explored after surveying AI, specifically while examining the lawyers' ethical responsibilities for using AI.

2.1. Current landscape: types, models, and applications of AI

AI is used to harness intelligence found in datasets. Specifically, ML techniques, such as supervising learning, are employed to train algorithms used in the creation of legal tech tools. In supervised learning, the models are trained using a labelled dataset to classify data or predict outcomes.[15] When input data is added, the AI model is fine-tuned until it is fitted appropriately, which follows as part of the cross-validation process.[16] In other words, the model can estimate its accuracy and learn over time. Human intervention is required to structure data modelling during training, and there is a probability of error, which means algorithms may learn

[12] John McCarthy, 'What is Artificial Intelligence'(Stanford Education, 2011) <http://jmc.stanford.edu/artificial-intelligence/what-is-ai/index.html> accessed 23 March 2023.
[13] Ibid.
[14] "Weak AI" and "Strong AI" are terms coined by John Searle in the "Chinese room argument". See John Searle, 'Minds, brains, and programs,' (1980) Behavioral and Brain Sciences 3(3), 417.
[15] Julianna Delua, 'Supervised vs. Unsupervised Learning: What's the Difference?' (IBM Cloud, 12 March 2021) <https://www.ibm.com/cloud/blog/supervised-vs-unsupervised-learning> accessed 23 March 2023.
[16] Andrew Y. NG, 'Preventing "Overfitting" of Cross-Validation Data' (Proceedings of the14th International Conference on Machine Learning, Nashville, July 1997) <https://dl.acm.org/doi/abs/10.5555/645526.657119> accessed 23 March 2023.

incorrectly. This may lead to bias and discrimination in models, to which the legal sector is not immune.

Supervised learning models are trained on datasets that contain labelled data, to map input variables (often called features) onto desired outputs (also called target variables or labels).[17] Unsupervised learning aims to uncover patterns in a dataset that are difficult to discover or not explicitly known.[18] Unlike supervised learning, datasets are not trained but rather algorithms are applied to unlabelled datasets to find clusters (datasets that are similar to each other) or association, to draw links. For example, this method can be particularly useful in in e-discovery where litigators may be able to find 'the smoking gun' among a plethora of documents.

Reinforcement learning (RL), a form of semi-supervised learning, follows a trial-and-error approach. The AI agent learns to take actions based on its interaction with the environment with the aim to maximize rewards and map a series of inputs to outputs with dependencies (such as Markov Decision Processes).[19] Unlike supervised learning, which relies on labelling, the focus is on balancing the unknown and current knowledge.[20] RL may be useful in legal work as a method of training automatic summarization of legal texts.

Many ML techniques including DL, a subset of ML, use neural networks to mimic the learning process of the human brain. Specifically, DL requires large amounts of data and extensive training to be able to make non-linear and complex correlations that are more accurate, such as in NLP-based software and applications.[21] DL is one of the major driving forces behind the new wave of interest and applications in NLP, a sub-field of AI, originating in the fields of linguistic and computer science.[22] It assists machines in understanding the complexities of human languages by transforming unstructured textual data into numeric vectors that can be analysed using ML techniques.[23] This signifies a new trend of combining NLP with ML, as successfully used in development of large learning models (LLMs). They are pre-trained using

[17] David Leslie et al, 'Understanding artificial intelligence ethics and safety: A guide for the responsible design and implementation of AI systems in the public sector' (The Alan Turing Institute 2019) 8 <https://doi.org/10.5281/zenodo.3240529> accessed 23 March 2023.
[18] Ibid.
[19] Richard Sutton and Andrew G Barto, *Reinforcement Learning: An Introduction* (2nd edn, The MIT Press 2015) 2-25
[20] Leslie et al (n 17).
[21] Adam Coates et al, 'Deep Learning with COTS HPC systems' (Proceedings of the 30th International Conference on Machine Learning, Atlanta, 2013) <http://proceedings.mlr.press/v28/coates13.pdf> accessed 23 March 2023.
[22] Li Deng and Yang Liu, *Deep Learning in Natural Language Processing* (Springer Singapore, 2018) 7.
[23] Ibid.

large datasets from a variety of sources making it easier to understand how human language is used. For example, ChatGPT[24] is currently being used to summarise clause language, generate clauses in drafts, insert clauses from a playbook or produce suggested redlines.[25] As LLMs are continuously enhanced by learning patterns from legal texts, they have the potential to revolutionise the legal tech industry, together with other AI tools and techniques.[26]

2.2. Legal AI Tools: Commonly used applications in legal practice

There are tasks performed by lawyers that AI can either replace or augment.[27] Broadly speaking, many lawyers already employ AI to handle low value tasks such as contract drafting, proofreading and review of non-complex documents.[28] The value of AI in the legal profession is however found in the automation of non-routine or knowledge-based work. Below, commonly used AI tools are briefly surveyed under certain categories, namely research, automation tools, prediction tools and document review.

2.2.1. Research

Research is deemed the bread and butter of legal work. However, it can be laborious process leading to inaccuracies where important legal rules are overlooked. Therefore, it is not surprising that research tools are the most commonly used AI applications in the legal profession.[29] Platforms that use ML and/or NLP to search and retrieve information relevant to a legal question and then deliver the information in an accessible fashion may assist in expediting traditional legal research. It may also improve the accuracy of the research.[30] For

[24] OpenAI, 'Introducing ChatGPT Plus' (OpenAI, 1 February 2023) <https://openai.com/blog/chatgpt-plus> accessed 23 March 2023.
[25] Christel Stokel-Walker, 'Generative AI Is Coming For the Lawyers' (Wired, 21 February 2023) <https://www-wired-com.cdn.ampproject.org/c/s/www.wired.com/story/chatgpt-generative-ai-is-coming-for-the-lawyers/amp> accessed 23 March 2023.
[26] Jena Frankenreiter and Julian Nyarko, 'Natural Language Processing in Legal Tech' in David Engstrom (ed), *Legal Tech and the Future of Civil Justice* (Cambridge University Press, 2023).
[27] Richard Susskind, *Tomorrow's Lawyer: An Introduction to Your Future* (2nd edn, Oxford University Press, 2017) 32.
[28] Kathryn Betts and Kyle Jaep, 'The dawn of fully automated contract drafting: machine learning breathes new life into a decades-old promise' (2017) 15 Duke Law & Technology Review 216, 220.
[29] Judge Herbert B. Dixon, 'What Judges and Lawyers Should Understand About Artificial Intelligence Technology' (ABA, February 03, 2020). <https://www.americanbar.org/groups/judicial/publications/judges_journal/2020/winter/what-judges-and-lawyers-should-understand-about-artificial-intelligence-technology/> accessed 23 March 2023.
[30] Faraz Dadgostari and others, 'Modeling Law Search as Prediction'(2021) 29 Artificial Intelligence and Law, 3.

example, Westlaw Edge[31] leverages AI to expedite the research process and claims to deliver more accurate results.

2.2.2. Automation Tools

There is much repetition in law, with many processes and forms requiring the checking of boxes or completion of forms. In many firms, templates such as condition precedents in contracts have already been digitalised. However, modern digital tools employ supervised ML algorithms that automate tasks or systematize processes. Automation software can manage the completion of a simple worksheet in draft documents, while also providing personalised and user-tailored results based on the user's interaction. Companies such as Neota[32] use AI to automate the creation of non-disclosure agreements.

2.2.3. Prediction Tools

Lawyering inherently involves navigating the 'grey areas' of the law with care. Terms such as 'reasonable' or 'diligence' are elusive and require the ability to read the law to make legal predictions to remove legal uncertainty. This is how lawyers, especially experienced ones, are able to justify their value – the ability to predict the outcome of the case. However, human judgment is flawed, studies have shown that AI algorithms consistently outperform human beings in predicting outcomes.[33] Thanks to the advancement of NLP and ML, it is now possible to digest unstructured data and generate results that uncover patterns within court decisions,[34] predict legal outcomes in tax and employment law,[35] rate lawyers' success rate,[36] and predict case outcomes[37].

---

[31] Thomas Reuters, 'Case law, legislation, practice notes and guidance'(Thomas Reuters) <https://legalsolutions.thomsonreuters.co.uk/en/products-services/westlaw-edge-uk.html> accessed 23 March 2023.

[32] Neota, 'Document Automation' (Neotalogic.com) <https://www.neotalogic.com/platform/document-automation/> accessed 23 March 2023.

[33] R. M. Dawes, 'Clinical Versus Actuarial Prediction' in Neil Smelser and Paul Bates (eds), *International Encyclopedia of the Social & Behavioral Sciences* (Pergamon, 2001).

[34] Nikolaos Aletras and others, 'Predicting judicial decisions of the European Court of Human Rights: a Natural Language Processing perspective' (2016) 10 Peer J Computer Science <https://doi.org/10.7717/peerj-cs.93> accessed 23 March 2023.

[35] Benjamin Alarie, Athony Niblett and Albert Yoon, 'Using Machine Learning to Predict Outcomes in Tax Law' SSRN Electronic Journal <https://ssrn.com/abstract=2855977> accessed 23 March 2023

[36] Premonition, 'Premonition Methodology' (Premonition.ai) <https://www.premonition.ai/accuracy-faq/> accessed 23 March 2023.

[37] Theodore W. Ruger et al, 'The Supreme Court Forecasting Project: Legal and Political Science Approaches to Predicting Supreme Court Decision making' (2004) Columbia Law Review 104(4) 1150.

2.2.4. Document Review

In document review, much of the complexity relates to calculating risks and translating legal risks into business acumen, but it is painstaking work. The risk for error is high, as many similar and complex documents need to be reviewed and there is always the risk that seemingly similar paragraphs can alter the meaning when drafted differently. Keeping abreast of documents and relevant changes is an arduous and time-consuming task increasingly fulfilled by AI. For example, algorithms are trained to learn how to perform certain tasks to identify, extract and analyse the information contained in large volumes of contract data.[38] LawGeex[39] is a solution that automatically reviews contacts, identifies clauses and variations in a document. Some brief analyser tools perform comparable tasks.[40] For instance, tools like the Kira system[41] identify, extract, and analyse business information contained in large volumes of contract data used in due diligence.

3. Lawyers' professional conduct duties and ethical responsibilities for using AI

3.1. Duty of competence

The duty of competence is one of the key duties for legal professionals, as reflected by the International Bar Association (IBA) Principles which states that '[a] lawyer's work shall be carried out in a competent and timely manner. A lawyer shall not take on work that the lawyer does not reasonably believe can be carried out in that manner.'[42] This duty requires lawyers to have 'the legal knowledge, skill, thoroughness, and preparation reasonably necessary for the

[38] Premonition (n 36) (fn 27) 7-11.
[39] LawGeex, 'Contract Review Automation'(LawGeex.com) <https://www.lawgeex.com/cra/> accessed 23 March 2023.
[40] See Nicole Black, 'Lawyers have a bevy of advanced and AI-enhanced legal research tools at their fingertips' (ABA Journal, 22 November 2019) <https://www.abajournal.com/web/article/lawyers-have-a-bevy-of-advanced-and-ai-enhanced-legal-research-tools-at-their-fingertips> accessed 23 March 2023.
[41] David Curle, 'Legal Due Diligence: Evolving Roles, Expanding Benefits' (Kirasystems.com, 23 November 2021) <https://kirasystems.com/blog/legal-due-diligence/> accessed 23 March 2023.
[42] International Bar Association (IBA), 'IBA International Principles on Conduct for the Legal Profession' (28 May 2011) Principle 9 <https://www.icj.org/wp-content/uploads/2014/10/IBA_International_Principles_on_Conduct_for_the_legal_prof.pdf> accessed 23 March 2023.

representation' as indicated by the US ABA Model Rules[43] as well as by the Singapore Legal Profession Rules[44] and the Canadian Model Code of Professional Conduct.[45]

The 'instrumentalist' view of technology as a mere aid for lawyers may not suffice with the advent of AI. The duty of professional competence takes on a new meaning when lawyers use AI. Especially, considering 'black box AI' where AI is incomprehensible to the developers which design it.[46] This raises not only the issue of 'opacity' but also potential 'bias' and 'discrimination' particularly when AI has the capability to give automated legal advice. Such ethical concerns are particularly relevant when AI is used for case prediction. Further, when a lawyer relies on AI-enabled 'legal advice', it has ramifications for the 'duty of competence'. The ABA has reacted to such concerns by modifying its Model Rules in 2012. Accordingly, the duty of competence requires lawyers "to maintain the requisite knowledge and skill, …[and] keep abreast of changes in the law and its practice, *including the benefits and risks associated with relevant technology*."[47] This has been interpreted as requiring lawyers "to better understand any advances in technology that genuinely relate to competent performance of the lawyer's duties to a client"[48] or "to evaluate how and how much to include AI technologies in their practices".[49] The ABA has subsequently adopted a resolution in 2019, postulating as follows: Under the Rule 1.1, lawyers also must have a basic understanding of how AI tools operate. While lawyers cannot be expected to know all the technical intricacies of AI systems, they are required to understand how AI technology produces results. As one legal commentator notes, "[i]f a lawyer uses a tool that suggests answers to the legal questions,

[43] ABA (American Bar Association) Model Rules of Professional Conduct, Rule 1. 1 (2002). This represents the rule that precedes the modification that took place in 2012 with reference to technology.
[44] Singapore Legal Profession (Professional Conduct) Rules 2015, Rule 5(1)(b) <https://sso.agc.gov.sg/SL/LPA1966-S706-2015> accessed 23 March 2023.
[45] Canadian Model Code of Professional Conduct (as amended October 19, 2019) 3.1.1. < https://flsc-s3-storage-pub.s3.ca-central-1.amazonaws.com/Model-Code-October-2019.pdf > accessed 23 March 2023.
[46] Adrienne LaFrance, 'Not Even the People Who Write Algorithms Really Know How They Work' (The Atlantic, 18 September 2015) <https://www.theatlantic.com/technology/archive/2015/09/not-even-the-people-who-write-algorithms-really-know-how-they-work/406099/> accessed 23 March 2023.
[47] ABA Model Rules of Professional Conduct, Rule 1.1.
[48] John M. Barkett, Shook, Hardy & Bacon LLP, 'More on the Ethics of E-Discovery: Predictive Coding and Other Forms of Computer-Assisted Review' (2012) 30 <https://scholarship.law.duke.edu/bolch/8/> accessed 23 March 2023.
[49] Jason Tashea and Nicholas Economou, 'Be Competent in AI before adopting, integrating it into your practice', ABA Journal (April 23, 2019) <https://www.abajournal.com/lawscribbler/article/before-lawyers-can-ethically-adopt-and-integrate-ai-into-their-practices-they-must-first-be-competent> accessed 23 March 2023.

he must understand the capabilities and limitations of the tool, and the risks and benefits of those answers."[50]

Similarly, the CCBE[51] emphasised that the duty of competence should 'entail not only the need to use reliable providers, but also the ability to request and understand the information on the basic characteristics of the program' and that such information should include ways to verify its compliance with the five principles of the European Ethical Charter on the Use of AI in Judicial Systems.[52] It is also essential to verify and check the outputs of ML algorithms to ensure the quality of the information and check for potential bias in algorithms.[53] Finally, the CCBE recommends that lawyers receive professional training in this field and actively participate in the design of AI tools.[54]

From this perspective, the 'duty of competence' should encompass 'technology' competence, that is, reasonable knowledge and skill to comprehend the likely consequences of AI-driven decisions. Thus, in the era of AI, lawyers must be involved in AI design or, at minimum, understand its biases (including that of the design, designer, and data) and its limitations (including the limits of observational data and exclusion of information which has not been taken on board by AI).[55] Additionally, as the CCBE proposes, lawyers should prioritise active involvement in AI design with clear stakeholder roles, in contact and communication with legal tech companies.

3.2. Duty to act in the best interests of the client

Lawyers have a universally recognised duty to act in the best interests of the client. This principle is reflected in frameworks for legal professionals, such as the SRA's seven

[50] See also ABA Resolution 112 (2019) 5. <https://www.americanbar.org/content/dam/aba/directories/policy/annual-2019/112-annual-2019.pdf> accessed 23 March 2023; Canadian Model Code of Professional Conduct, 4A (commentary), 16-17.
[51] CCBE, 'CCBE Considerations on the Legal Aspects of Artificial Intelligence 2020' (CCBE.eu)32 <https://www.ccbe.eu/fileadmin/speciality_distribution/public/documents/IT_LAW/ITL_Guides_recommendations/EN_ITL_20200220_CCBE-considerations-on-the-Legal-Aspects-of-AI.pdf> accessed 23 March 2023.
[52] CEPEJ, 'European Ethical Charter on the use of artificial intelligence in judicial systems and their environment' (2018) 7-12 <https://rm.coe.int/ethical-charter-en-for-publication-4-december-2018/16808f699c> accessed 23 March 2023.
[53] CCBE (n 51).
[54] CCBE (n 51).
[55] Simshaw (n 5)

principles,[56] or the Australian Conduct Rules, which require a solicitor to 'act in the best interests of a client in any matter in which the solicitor represents the client`,[57]or the Canadian Codes of Conduct, which obliges lawyers to 'maintain loyalty to clients and avoid conflicts of interest'.[58]

Crucially, this duty is based on the 'trust' that is reposed to the lawyers for the relationship they have with their clients. Under English common law, at the core of this relationship lies the attitude of the fiduciary, that is, the lawyer, 'who has undertaken to act for or on behalf of another in a particular matter in circumstances which give rise to a relationship of trust and confidence.'[59] This fiduciary relationship requires lawyers to act solely in the interests of their clients, subordinating personal interests, and upholding the 'obligation of loyalty'.[60], avoiding conflicts of interest, and avoiding 'unauthorised or secret profits'.[61] Moreover, from a broader perspective, fiduciary duties typically encompass the 'duty of confidentiality' and the 'duty of care'.

The duty, in its broad sense, entails not charging clients excessively or unnecessarily, as the ABA Model Rules suggest.[62] Singapore's Legal Profession (Professional Conduct) Rules also state that a 'legal practitioner must not undertake work in a manner that unnecessarily or improperly increases the costs that are payable to the legal practitioner; and must, at all times, use the legal practitioner's best endeavours to complete any work for his or her client as soon as is reasonably possible'.[63] From a broader perspective, this duty may also imply a positive obligation to determine whether the use of AI is necessary for better serving the client.[64]

---

[56] Solicitors Regulation Authority (SRA), 'SRA Principles' <https://www.sra.org.uk/solicitors/standards-regulations/principles/> accessed 23 March 2023.
[57] Australian ACT Law Society, 'ACT Legal Profession (Solicitors) Conduct Rules 2015: Commentary and Guidelines', (2015) Rule 4.1.1.
[58] Federation of Law Societies of Canada, 'Law Society Codes of Conduct' <https://flsc.ca/what-we-do/model-code-of-professional-conduct/law-society-codes-of-conduct/> accessed 23 March 2023. See also IBA International Principles on Conduct for the Legal Profession, Principle 5; CCBE (n 9) 16.
[59] *Bristol & West Building Society v Mothew (t/a Stapley & Co)* [1998] Ch 1, [1996] EWCA Civ 533.
[60] Practical Law Corporate, 'Fiduciary duties' Practical Law UK Practice (2022) Note 8-107-4883.
[61] M. Atkins, 'What is the purpose of the ongoing use of fiduciary duties in English business law, with particular reference to breaches of duty in relation to bribery, secret profits, conflicts of interest and unconscionability?' Unpublished PhD dissertation (Lancaster University 2017) 32.
[62] According to the ABA Model Rules, "A lawyer shall not make an agreement for, charge, or collect an unreasonable fee or an unreasonable amount for expenses." (ABA Model Rules of Professional Conduct, Rule 1.5).
[63] Singapore Legal Profession (Professional Conduct) Rules, Rule 17(2).
[64] Meng (n 3) 55; Davis (n 3).

Such obligation may necessitate further evaluations such as: (i) assessing AI trustworthiness amidst challenges in training data diversity, outcome adaptability, and system resilience against security threats; (ii) determining AI's added value in terms of speed and accuracy for clients; (iii) evaluating the justifiability of any change in service fees with clients' best interests in mind; and (iv) deciding on the responsibility for retraining AI and its continuous monitoring to improve outcomes. Crucially, the assessment is closely linked with the notions of responsibility, accountability, and liability, all of which necessitate a broader ethical and legal perspective to fulfil the duty of acting in the best interest of the client, alongside other professional duties.

3.3. Duty of confidentiality

The duty of confidentiality is a well-established principle that lawyers must adhere to. According to the ABA Model Rules, '[a] lawyer shall not reveal information relating to the representation of a client unless the client gives informed consent'.[65] Similarly, the SRA's Code of Conduct[66] and Canadian Model Code[67] require lawyers to maintain the confidentiality of their clients' information and affairs, except when disclosure is mandated by law or authorised by client consent.[68] The IBA Principles also contain a similar provision.[69]

In English Law, the duty to preserve confidentiality is unqualified, meaning lawyers must keep client information confidential, not merely to take all the reasonable steps to do so.[70] It entails not disclosing the information to third parties or misusing it, that is, without the consent of a client or former client to make any use of it or to cause any use to be made of it by others otherwise than for the client's benefit.[71]

---

[65] ABA Model Rules of Professional Conduct, Rule 1.6.
[66] Code of Conduct for Solicitors, RELs and RFLs and of the Code of Conduct for Firms (referred to collectively as ("the Codes") Para. 6.3 <https://www.sra.org.uk/solicitors/guidance/confidentiality-client-information/> accessed 23 March 2023.
[67] Canadian Model Code of Professional Conduct, 3.3-1.
[68] The Australian (ACT) Conduct Rules underpin that a practitioner's obligation to maintain the confidentiality of a client's affairs is not limited to information which might be protected by legal professional privilege and is a duty inherent in the fiduciary relationship between the practitioner and client (See Australian (ACT) Legal Profession (Solicitors) Conduct Rules 9 and 4.1.1).
[69] See IBA International Principles on Conduct for the Legal Profession, Principle 4.
[70] SRA, 'Guidance: Confidentiality of client information' (30 June 2022) <https://www.sra.org.uk/solicitors/guidance/confidentiality-client-information/> accessed 23 March 2023.
[71] Ibid citing *Prince Jeffrey Bolkiah v KPMG* [1998] UKHL.

Notwithstanding, rather than being a negative obligation, this duty is referred to as a positive obligation '[t]o safeguard information relating to the representation of a client … against inadvertent or unauthorized disclosure' and to '[t]ake reasonable precautions to prevent the information from coming into the hands of unintended recipients' when transmitting a communication of such information.[72] As a matter of fact, the modified ABA Rule 1.6 requires lawyers to 'make reasonable efforts to prevent the inadvertent or unauthorized disclosure of, or unauthorized access to, information relating to the representation of a client.'[73] By the same token, the ACT Rules suggest 'as a matter of good practice, solicitors should de-identify this [confidential] information to the greatest extent possible.'[74]

Confidentiality requirements for lawyers cover both privacy and security measures, as protecting client information involves data protection and IT systems security. Given the AI-related risks 'duty of confidentiality' can be deemed an expansive obligation. Considering AI is usually exposed to vast datasets, AI tools can process the clients' confidential information, when sifting through the documents or drafting contracts. By the same token, the CCBE flagged the issue of reusing data, which is common practice on current online legal platforms, and recommend that lawyers require terms 'excluding profiling activity and the reuse of data even after supposed anonymisation of the data'.[75]

Overall, to protect confidentiality in the AI era, it is not enough to rely solely on security measures like encryption or pseudonymisation. Rather a broader understanding of how AI systems work, communicate with clients (and former clients) is needed to appreciate expectations and preferences.[76] Additionally, AI developers and providers should respect the Conduct Rules within their respective domains, such as considering ethics at the design level during the AI life cycle.

[72] Simshaw (n 5) 199.
[73] ABA Model Rules of Professional Conduct, Rule 1.6.
[74] ACT Legal Profession (Solicitors) Conduct Rules 2015: Commentary and Guidelines (2015) Rule 4.1.1.
[75] Guide on the use of Artificial Intelligence-based tools by lawyers and law firms in the EU 2022 published by the CCBE and the European Lawyer Foundation (ELF) published on 31 March 2022 under the project AI4Lawyers, 47 <https://www.ccbe.eu/fileadmin/speciality_distribution/public/documents/IT_LAW/ITL_Reports_studies/EN_ITL_20220331_Guide-AI4L.pdf > accessed 23 March 2023.
[76] Simshaw (n 5) 200.

3.4. Independence, Integrity, and Fairness

'Independence' is one of the universally accepted principles of the legal profession. ABA's Rule 2.1 mandates lawyers to 'exercise independent professional judgment',[77] while the IBA's principles require a lawyer to maintain independence to provide clients with unbiased advice and representation.[78] Under the Canadian Model Code, it is set out that "A lawyer who engages in another profession, business or occupation concurrently with the practice of law must not allow such outside interest to jeopardize the lawyer's professional integrity, independence or competence".[79]

Overall, the principle of 'independence' is regarded as both protective, meaning a lawyer should be protected from outside pressures that impair professional judgement, and as self-disciplinary, meaning a lawyer should not impair their professional judgement by pursuing personal interests or succumbing to outside pressures.[80] Mirroring the two facets of this principle, the IBA has identified a number of circumstances that would risk a lawyer's independence, on a non-exhaustive basis.[81]

While AI automation, research and prediction tools utilising ML and NLP techniques, offer sophisticated legal tech solutions, they use may also affect a lawyer's ethical standpoint, such as in the case of overfitting data being used for AI training. Given all such external and internal factors that can potentially compromise their judgment, a further question concerns whether lawyers who allow AI algorithms to make predictions are, in fact, exercising their own 'professional judgment'.[82] This must be evaluated along with the risk of 'automation bias', resulting from over-reliance on AI, which may harm the soundness and independence of a decision.[83]

---

[77] Regarding a similar provision see Australian ACT Legal Profession (Solicitors) Conduct Rules 2015, 17.1.
[78] IBA International Principles on Conduct for the Legal Profession, Principle 1.
[79] Canadian Model Code of Professional Conduct, 7.3-1.
[80] Bjorn Fasterling, 'The Managerial Law Firm and the Globalization of Legal Ethics' (2009) 88 Journal of Business Ethics 21, 26-27.
[81] IBA International Principles on Conduct for the Legal Profession, 12-13. Also, regarding examples and accompanying explanations as to the boundaries of 'lack of independence' see SRA, 'Case studies: Lack of independence (25 November 2019) <https://www.sra.org.uk/solicitors/guidance/lack-independence/> accessed 23 March 2023.
[82] Mike Hoeflich, 'Legal ethics and Artificial Intelligence' (2021) 2(3) Legal Ethics & Malpractice Reporter 2, 4.
[83] While over-reliance would need to be mitigated, lawyers should presumably avoid underutilising AI, which could cause them to serve their clients less efficiently (ABA Resolution 112, 7).

The potential harm caused by employing ML and NLP techniques can be significant because they use data inputs to achieve outputs that are not transparent to lawyers or even to AI programmers, due to black box algorithms. Therefore, making AI systems transparent and explainable is integral to maintaining the duty of 'integrity'. The broader notion of 'integrity' is linked to other ethical concerns including 'honesty' reflecting the UK approach that treat 'integrity' as a separate principle.[84] The latter links integrity to a lawyer's trustworthiness, public confidence in the administration of justice, and the legal profession.[85] The SRA emphasises that integrity encompasses more than merely acting dishonestly[86] and includes unethical behaviours related to a solicitor's integrity, independence, and honesty.[87]

Issues of integrity and independence would need a (re)solution in situations where lawyers face sophisticated AI results that may mislead users without verification and monitoring. Hence, the more transparent the AI algorithms are, the more ethical and independent the judgements based on AI will be. Thus, if ethical and legal responsibilities are transparently defined and monitored, AI-driven legal tech tools can be used more safely.

Another closely related issue: 'fairness', needs also to be taken into consideration for a holistic treatment of any client representation involving AI. Although fairness is referred to as one of the lawyers' duties towards the other party,[88] the likely controversies would require a broader conception incorporating the life cycle of AI and how it would eventually affect clients. Perceptions of technology as value neutral or incapable of error might result in dulled moral sensitivity and insufficient scrutiny being applied to AI systems.[89] In this regard, lawyers' professional duties should entail far-reaching accountability measures, also addressing bias and undue influence.[90]

---

[84] See SRA Principles; Canadian Model Code of Professional Conduct, 2019, 2.1-1 reading "A lawyer has a duty to carry on the practice of law and discharge all responsibilities to clients, tribunals, the public and other members of the profession honourably and with integrity".
[85] Ibid.
[86] SRA, 'Integrity and ethics' (23 November 2020) <https://www.sra.org.uk/sra/research-publications/risk-outlook-2020-21/integrity-and-ethics/> accessed 23 March 2023.
[87] Ibid. Under the English Law, the meaning of integrity has been considered several times by the courts in recent years including through *Wingate & Anor v SRA* [2018] EWCA Civ 366 Case No and Malins v SRA [2018 ECWA Civ 3666], considered together by the Court of Appeal in 2018 and also through *Ryan Beckwith v SRA* considered by the High Court in 2020 [2020 EWHC 3231 (Admin)] (Ibid).
[88] ABA Model Rules of Professional Conduct, Rule 3.4 (entitled 'Fairness to Opposing Party & Counsel').
[89] Justine Rogers and Felicity Bell, 'The ethical AI lawyer: What is required of lawyers when they use automated systems?' 2019 1(1) Law, Technology and Humans 80, 87.
[90] H. Felzmann, E, Fosch-Villaronga, and C. Lutz, 'Towards Transparency by Design for Artificial Intelligence' (2020) 26 Science and Engineering Ethics 3333, 3343.

Overall, 'transparency' and 'accountability' are key issues that require careful consideration in AI-based lawyering, as they are closely linked to independence, integrity, and fairness. Above all, a lawyer's legal and ethical responsibilities should extend beyond simply taking professional measures to address ethical concerns in the use of AI but a holistic approach that encompasses the entire AI life cycle.

3.5. Duty of supervision

The duty of supervision is recognised as another professional duty of lawyers,[91] which typically involves exercising reasonable supervision over their team and other staff engaged in the legal work. For instance, Canadian Model Code provides that '[a] lawyer has complete professional responsibility for all business entrusted to him or her and must directly supervise staff and assistants to whom the lawyer delegates particular tasks and functions'.[92] The implicit requirement of supervision of human employees, needs to be revisited given the possibility of delegating some legal work to AI. In fact, one could argue that both human and non-human assistants should be supervised since they contribute to the legal service and representation of the clients.

AI is transforming legal tech tools used by legal practitioners, and this situation requires consideration of supervision not only non-lawyer humans, but also non-human elements such as AI that are involved in the process of representing clients. This impact of AI seems to have influenced the ABA's approach to the lawyers' duty of supervision. The ABA model rules were changed in 2012 to include not only human but also non-human assistance under the lawyers' duty of supervision. By changing the Rule 5.3's title from 'Responsibilities Regarding Nonlawyer Assistants' to 'Responsibilities Regarding Nonlawyer Assistance', the ABA indicated that the rule is intended to have reach beyond human assistants, to other non-lawyers, human or not, involved in the representation of a client.[93]

[91] See ABA Model Rules of Professional Conduct, Rules 5.1 and 5.3(a)-(b); Canadian Model Code of Professional Conduct, 6.1.; Australian (ACT) Legal Profession (Solicitors) Conduct Rules 2015, Rule 37; Singapore Legal Profession (Professional Conduct) Rules 2015, Rule 36.
[92] Canadian Model Code of Professional Conduct, 6.1-1.
[93] See ABA Model Rules of Professional Conduct, Rule 5.3.

While it is ultimately up to lawyers to decide whether to employ AI-generated predictions for case outcomes or case analyses, recent trends, and advancement in AI's capabilities suggest AI-based outputs may be more relied upon when compared to that those of humans.[94] Notwithstanding, it is important to balance our expectations, with the need for proper tools and mechanisms to supervise AI, as it is not infallible.[95]

4. AI Ethics in general

4.1. General overview

Ethics, recognised as a branch of philosophy, encompasses a diverse range of aspects, such as moral values, norms, and practical application,[96] which extends to various fields, including the legal sector, as can be seen through the development of Conduct Rules. On the other hand, AI has the potential to cause revolutionary change, introducing a new ethos in various aspects of life, including the legal profession. AI permeating the legal sector would mean AI ethics meeting and interacting with legal ethics eventually. Therefore, investigating ethics in the legal sector involves reviewing underlying moral values and norms not only from the perspective of the legal professionals (legal ethics) but also from an AI-centric viewpoint, as echoed with 'AI ethics' in general.[97]

One may consider various ethical frameworks, including virtue ethics and top-down, bottom-up, and hybrid approaches, when applying moral philosophy to AI.[98] Without delving into details of ethics as a discipline, this section examines the leading concerns and principles that drive AI ethics. It is worth noting that rule makers have mainly adopted converging ethical rules or standards through soft law (high-level) principles, which also extend to hard law, as

[94] See also Chris Stokel, 'Generative AI Is Coming for the Lawyers' (Wired 21 February 2023) <https://www.wired.co.uk/article/generative-ai-is-coming-for-the-lawyers> accessed 23 March 2023.
[95] See also ABA Resolution 112, 7.
[96] Robin Attfield, *Applied Ethics: An Introduction* (Polity 2023) 6-10; Mark Dimmock and Andrew Fisher, 'Normative Ethics, Metaethics and Applied Ethics: Three Branches of Ethics' *Ethics for A-Level* (Open Book Publishers 2017) 93.
[97] See Leslie et al (n 17) 3-4; Central Digital & Data Office, Office for Artificial Intelligence, Guidance: Understanding artificial intelligence ethics and safety ('UK Guidance 2019') (10 June 2019) <https://www.gov.uk/guidance/understanding-artificial-intelligence-ethics-and-safety> accessed 23 March 2023).
[98] Wendell Wallach and Shannon Vallor, 'Moral Machines: From value alignment to embodied virtue' in S. M. Liao (eds.) 'Ethics of Artificial Intelligence' (OUP 2020) 388-391.

evident in the EU's Proposed AI Act. What follows is first an examination of AI ethics principles, including transparency, fairness, and accountability, followed by an investigation of regulatory tools in the field.

4.2. Ethical concerns in relation to AI

Algorithms present ethical challenges not only due to the scale of analysis and complexity of decision making but also because of the uncertainty and opacity of their operations.[99] These concerns, along with bias and discrimination, drive AI ethics to have a consolidated nature based on the principles of 'transparency', 'fairness' and 'accountability'.[100] Although other concerns such as privacy and security can be added to the list of relevant issues and challenges, these are not examined within the domain of AI ethics, mainly for their very nature being not directly related to ethics.[101]

4.2.1. Transparency

Transparency is a key component of AI ethics, as it addresses the 'opacity' or 'black box' problem.[102] The value of transparency lies in the explainability and interpretability of the algorithmic system, including the model and data used.[103] While full openness represents the far end of the spectrum of transparency, proprietary systems that are closed and protected by intellectual property rights (IPRs), such as trademarks against third party access, lies at the other end. Not only the presence of IPRs but also special skills required to create and understand the operation of a complex AI/ML system pose an impediment for transparency.

[99] Mittelstadt, B. D. et al., 'The ethics of algorithms: Mapping the debate' (July-December 2016) Big Data & Society 1, 3.
[100] See Leslie et al (n 17) 112; Independent High-Level Expert Group (IHLEG) on Artificial Intelligence (set up by the European Commission) Ethics Guidelines for Trustworthy Artificial Intelligence (2019) <https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai> accessed 23 March 2023; UK Cabinet Office, Central Digital & Data Office, Office for Artificial Intelligence, 'Guidance: Ethics, Transparency and Accountability Framework for Automated Decision-Making' (13 May 2021); UK Guidance 2019 (n 97); Sebastian Hallensleben and Carla Hustedt, 'From Principles to Practice: An interdisciplinary framework to operationalise AI ethics' (Artificial Intelligence Impact (AIEI) Group 2020).
[101] See ibid.
[102] AIEI Group (n 100) 19.
[103] AIEI Group (n 100) 19.

It is a commonly held belief that there is a transparency paradox in that revealing how an algorithm works, even if it were possible to predict consistently, would mean 'revealing information handling practices in ways that are relevant and meaningful to the choices individuals must make'.[104] Even if one did so, describing 'every flow, condition, qualification and exception' would neither be read nor understood by the user.[105] Given this, in the context of any transparency regulation, potential benefits would need to be balanced against the adverse implications, including any contradiction to the duty of confidentiality.

Transparency obligations for AI systems or products can be based on the information requirements specified in Articles 12 to 22 of the GDPR.[106] In addition to these horizontal obligations, sector-specific transparency regulations can be developed to address the unique societal needs related to distinct sectors. These regulations should consider factors such as the criticality of the sector, the importance of the system operator for the sector, the impact on individual and collective rights, the type of data involved (such as special categories under article 9 GDPR), the integration of the technology into decision-making processes, and the need to protect the legitimate interests of system operators. To that end, meaningful and effective transparency obligations tailored to each sector's unique needs and challenges should be adopted.[107]

4.2.2. Fairness

Fairness is a major issue regarding AI systems or algorithms, as they have the potential to introduce bias and discrimination. Bias relates to the decision making process itself, whereas discrimination describes the negative and disproportionate impact of algorithmic decision making.[108] Unfair outcomes might originate from the quality of the training data, for example in relation to unrepresentative data being used, as well as result from the algorithmic design or programming.[109] Unfairness or its root causes are not visible every time, for example, when an

[104] Thomas Wischmeyer, 'Artificial Intelligence and Transparency: Opening the Black Box' in Thomas Wischmeyer and Timo Rademacher (eds.) *Regulating Artificial Intelligence* (Springer 2020) 75, 94.
[105] Ibid.
[106] Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC ('General Data Protection Regulation' or 'GDPR') [2016] OJ L 119/1, art. 12-22.
[107] Ibid.
[108] Mittelstadt (n 99) 8.
[109] P. W. Grimm, M. R. Grossman, and G. V. Cormack, 'Artificial Intelligence as Evidence' (2021) 19(1) Northwestern Journal of Technology and Intellectual Property 9, 44; N. Criado and J. M. Such, 'Digital

AI solution actually works better for one gender or race than another, but the beneficial effect is masked by an overall, that is, a combined accuracy rate that is low, or because the protected data is either not collected or not considered by the algorithm.[110] Crucially, using data from the same stereotypes in an AI model may perpetuate unfair outcomes,[111] meaning decisions based on similar data can entrench biased or discriminatory practices from the past.[112]

Against this background, a holistic approach is advisable, taking in account all stages of the AI life cycle, such as modelling, training and usage, where algorithmic design and choices, data quality, and usage would have an impact.[113] It is argued that four overlapping strategies would enable discrimination prevention in analytics: (i) controlled distortion of training data, (ii) integration of anti-discrimination criteria into classifier into the classifier algorithm, (iii) post-processing of classification models and (iv) modification of predictions and decisions.[114] Such an approach is persuasive for embodying all the stages of AI life cycle and offering a vision to review, monitor and update the underlying model and variables towards elimination of unfair outcomes.

Ensuring fairness or preventing bias and discrimination in AI systems may not be measured on an even basis across different jurisdictions due to differences in laws. However, the factors for investigating fairness should remain consistent across different sectors, given the fundamental rights and non-discrimination laws that apply universally,[115] including 'prohibition of

Discrimination' in K. Yeung and M. Lodge (eds.) *Algorithmic Regulation* (OUP 2019) 82, 85-86; K. Yeung, 'Why worry about decision-making by machine?' in K. Yeung and M. Lodge (eds.) *Algorithmic Regulation* (OUP 2019) 21, 32-33.

[110] Grimm, Grossman and Cormack (n 109) 42.

[111] Some ML algorithms operate through a continuous process of feedback and learning from new data, mostly in an unsupervised manner, resulting in perennial unfinished predictive models. In other words, the algorithm updates itself as it learns bias in the data; it does not follow a standard static program (See also Fernando Ávila, Kelly Hannah-Moffat and Paula Maurutto, 'The seductiveness of fairness: Is machine learning the answer? - Algorithmic fairness in criminal justice systems' in M. Schuilenburg and R. Peeters (eds.) *The Algorithmic Society: Technology, Power and Knowledge* (Routledge 2022) 87, 92.

[112] Mark Coeckelbergh, *AI Ethics* (The MIT Press 2020) 134.

[113] Criado and Such, 85-87. See also K. Swapnil, 'Types of Biases in Data' (Towards Data Science, 26 August 2021) <https://towardsdatascience.com/types-of-biases-in-data-cafc4f2634fb> accessed 23 March 2023.

[114] Mittelstandt et al (n 109) 8.

[115] See Lance Eliot, 'Emerging Rise of "AI Localism" is getting bigger, bolder, and badder says AI Ethics and AI Law' (Forbes, 12 December 2022) <https://www.forbes.com/sites/lanceeliot/2022/12/12/emerging-rise-of-ai-localism-is-getting-bigger-bolder-and-badder-says-ai-ethics-and-ai-law/?sh=44d056647e08> accessed 23 March 2023.

discrimination' under Article 14 of the European Convention on Human Rights, as well as non-discrimination statutes, like the Equality Act in the UK.[116]

4.2.3. Accountability

Accountability refers to the requirement for the system to be able to explain and justify its decisions to users and other relevant actors.[117] Accountability of AI systems requires both the function of guiding action (by forming beliefs and making decisions) and the function of explanation (by placing decisions in a broader context and by classifying them along moral values).[118] However, the design or development of AI would result in an accountability gap across the values or principles established earlier, for instance in the legal sector, by means of statutory and/or professional rules. The rationale and default outcome based on AI algorithms can diverge from the well-established values and principles that used to drive decisions or assessments made by humans. This misalignment might result from 'technical', 'social' or 'emerging' bias and would be augmented with the 'many hands problem' or the 'traceability of harm'.[119]

To ensure accountability, decisions should be (i) derivable from, and explained by, the decision-making mechanisms used and (ii) be in harmony with the moral values and societal norms that inform the purpose of the AI system as well as its operation.[120] From this point of view, identification of not only the causes for decision making process but also the responsibility of the actors for any traceable harm and/or breach of rule that might be caused by AI is key to accountability. In practice, it would entail identifying and mitigating risks in a transparent way that can be explained to and audited by third parties in an AI-driven socio-technical system.[121]

---

[116] Generally, laws related to data protection, human rights and non-discrimination may be used to prevent and address unfair outcomes resulting from bias and discrimination. See Frederik J. Zuiderveen Borgesius Strengthening legal protection against discrimination by algorithms and artificial intelligence, The International Journal of Human Rights (2020) 24(10) 1572-1593.

[117] Virginia Dignum, 'Responsibility and Artificial Intelligence' in M. D. Dubber, F. Pasquale and S. Das (eds) 'The Oxford Handbook of Ethics of AI' (OUP 2020) 218.

[118] Virginia Dignum, 'The ART of AI - Accountability, Responsibility, Transparency' (Medium, 4 March 2018) <https://medium.com/@virginiadignum> accessed 23 March 2023.

[119] World Health Organization (WHO), 'Ethics and governance of artificial intelligence for health' WHO guidance (2021) 43 <https://www.who.int/publications/i/item/9789240029200> accessed 23 March 2023.

[120] Dignum (n 118).

[121] Rowena Rodrigues, 'Legal and human rights issues of AI: Gaps, challenges and vulnerabilities' (2020) 4 Journal of Responsible Technology <https://doi.org/10.1016/j.jrt.2020.100005> accessed 23 March 2023.

When managing related risks and potential harms, a distinction can be made between accountability for 'content' - such as the algorithmic model and design choices - 'operation', and 'usage'. For instance, if harm is caused by a developer's design choices, the user (the legal professional) should not be presumed responsible, as they should only be held responsible for any improper usage of AI. Notwithstanding, there is a debate on how to apply human-centric ethical values in AI-driven socio-technical systems, including the extent to which developers should be held responsible for the harm caused by AI.[122] Overall, it is necessary to ensure accountability from a holistic perspective through the interaction of Conduct Rules (legal ethics) with AI ethics, with a view to eliminate the accountability gaps during and after AI life cycle.

4.3. Main threads to regulate AI ethics

4.3.1. Soft law and hard law measures

AI has increasingly come under the spotlight of the law and policy makers due to ethical concerns related to human rights, data protection, security, and individual and societal well-being. Over 90 diverse organisations have attempted to define ethical AI principles[123] Around 173 ethical guidelines that outline principles for developing and implementing automated decision-making ethically have been published thus far.[124] These guidelines include share common principles such as transparency, equality/non-discrimination, accountability and safety, while some demand that AI should be socially beneficial and protect human rights.[125]

This initial response suggests a preference for a principles-based regulatory approach to AI ethics. Such an approach involves analysing the key precepts in each field regulated by international instruments vertically, followed by a second phase that considers the

[122] Mittelstadt, et al (n 109) 11-12.
[123] Maria Luciana Axente and Ilana Golbin, 'Ten principles for ethical AI' (PwC Australia, 13 April 2022) <https://www.pwc.com.au/digitalpulse/ten-principles-ethical-ai.html> accessed 23 March 2023.
[124] AlgorithmWatch, AI Ethics Guidelines Global Inventory <https://inventory.algorithmwatch.org/about> accessed 23 March 2023.
[125] Ibid.

commonalities across all fields.[126] Contextualising guiding principles and legal values is a key element in ensuring effective implementation of binding instruments. Additionally, utilising non-binding instruments can provide granular applications of these principles to specific contexts.[127] Non-state actors taking these steps might pre-empt stricter risk-based regulation by state actors, while signalling AI users and implementors regarding their solutions and services. Soft law instruments in the field of AI ethics include the OECD's 2019 Recommendation,[128] the Asilomar AI Principles,[129] the Montreal Declaration for Responsible AI,[130] the IEEE Principles,[131] the EU's Ethics Guidelines for Trustworthy Artificial Intelligence,[132] and reports issued by 'Working Group on Responsible AI' under the auspices of the Global Partnership on AI.[133] Notwithstanding, the use of soft law is characterised by substantive expectations that lack enforcement or oversight mechanisms.[134] For instance, even the IEEE's ethical guidelines have proven ineffective, as large technology companies like Facebook, Google and Twitter do not implement them, despite having IEEE members among their engineers and developers.[135]

Consequently, the EU has taken a firmer stance by initiating a legislative process with the Commission's proposal of the AI Act (Regulation) in 2021, followed by a modified version by the Council of the EU in November 2022.[136] A finalised version is expected after a trialogue in near future. The proposed Regulation is a hard law measure that aims to create a uniform, horizontal legal framework for AI governance, based on a risk-based approach.[137] It requires AI to be legally, ethically, and technically robust, while upholding democratic values, human

[126] Braun Tomasz, *Regulating Artificial Intelligence: Binary Ethics and the Law (Routledge Research in the Law of Emerging Technologies)* (Routledge 2021)144).
[127] Ibid.
[128] OECD (n 11).
[129] Future of Life Institute, AI Principles (2017) <https://futureoflife.org/open-letter/ai-principles/> accessed 23 March 2023.
[130] Université de Montréal, 'The Montréal Declaration for a Responsible Development of Artificial Intelligence' (2018) <https://www.montrealdeclaration-responsibleai.com/> accessed 23 March 2023.
[131] IEEE Standards Association (IEEE SA), 'Version 2 of Ethically Aligned Design: A Vision for Prioritizing Human Well-being with Autonomous and Intelligent Systems' (2017) <https://standards.ieee.org/wp-content/uploads/import/documents/other/ead_v2.pdf> accessed 23 March 2023.
[132] See IHLEG (n 100).
[133] Global Partnership on Artificial Intelligence <https://gpai.ai/projects/responsible-ai/> accessed 23 March 2023.
[134] Ida Varošanec, 'On the path to the future: mapping the notion of transparency in the EU regulatory framework for AI' (2022) 36(2) International Review of Law, Computers & Technology 95, 107.
[135] Algorithm Watch, 'In the realm of paper tigers – exploring the failings of AI ethics guidelines' (28 April 2020)<https://algorithmwatch.org/en/ai-ethics-guidelines-inventory-upgrade-2020/> accessed 23 March 2023.
[136] See Tambiama Madiega, 'Briefing on EU Legislation in Progress: Artificial intelligence Act' (European Parliamentary Research Service, January 2022) 1-2 <https://www.europarl.europa.eu/RegData/etudes/BRIE/2021/698792/EPRS_BRI(2021)698792_EN.pdf> accessed 23 March 2023.
[137] Ibid.

rights and the rule of law,[138] based on a four-category 'product safety framework', including (i) unaccepted risk, (ii) high-risk, (iii) limited risk and (iv) low or minimal risk.[139] High-risk and limited-risk AI systems are subject to specific obligations, while low or minimal risk systems have none. AI systems under the unaccepted risk category are prohibited from entering the EU market.[140] Furthermore, it aims to regulate AI systems that pose a high risk to safety or fundamental rights in pre-selected areas.[141] While 'administration of justice' is among the pre-selected areas, it does not yet cover AI systems used in legal practice. The proposed regulation, includes measures related to addressing ethical concerns and, along with the institutional oversight[142] and enforcement powers,[143] will ultimately shape the business environment for AI design, development, deployment and use.

4.3.2. Standards and specifications

Industry stakeholders involved in the AI life cycle are establishing governing mechanisms to develop ethical and legal AI. There are diverse initiatives signifying a self-regulatory environment focused on company ecosystems. Companies are aware that investing in their own frameworks and ethics policies is a way to presumably pre-empt interventions and avoid more stringent regulations.[144] For instance, Facebook created a Responsible AI team to build and test approaches to ensure that their ML systems are designed and used responsibly.[145] Likewise, in

[138] Mauritz Kop, 'EU Artificial Intelligence Act: The European Approach to AI' Stanford - Vienna Transatlantic Technology Law Forum, Transatlantic Antitrust and IPR Developments (Stanford University, Issue No. 2/2021). <https://law.stanford.edu/publications/eu-artificial-intelligence-act-the-european-approach-to-ai/> accessed 23 March 2023.
[139] Madiega (n 136) 5.
[140] See EU's Proposed AI Act, art. 5.
[141] See EU's Proposed AI Act, Annex III.
[142] The EU's proposed regulation aims to create a European AI Board, composed of representatives of member states and the Commission. Being chaired by the European Commission, the Board can issue opinions, recommendations, and written contributions on matters related to the implementation of the proposed regulation (Benjamin Mueller, 'The Artificial Intelligence Act: A Quick Explainer' (Center for Data Innovation, 4 May 2021) <https://datainnovation.org/2021/05/the-artificial-intelligence-act-a-quick-explainer/> accessed 23 March 2023).
[143] As with GDPR, these rules apply to providers and users outside of the EU if the output of the system is used in the EU, along with administrative fines to be applied under certain circumstances (Ibid).
[144] See Lance Eliot, 'AI ethics and legal AI are flustered by deceptive pretences known as AI ethics - washing which are false claims of adhering to ethical AI, including for autonomous self-driving cars' (Forbes, June 2022) <https://www.forbes.com/sites/lanceeliot/2022/06/09/ai-ethics-and-legal-ai-are-flustered-by-deceptive-pretenses-known-as-ai-ethics-washing-which-are-false-claims-of-adhering-to-ethical-ai-including-for-autonomous-self-driving-cars/?sh=5c895a402b65> accessed 23 March 2023.
[145] Meta AI, 'Facebook's five pillars of Responsible AI' (22 June 2021) <https://ai.facebook.com/blog/facebooks-five-pillars-of-responsible-ai/> accessed 23 March 2023.

June 2018, Google published its 'Objectives' to assess their AI applications.[146] As opposed to the contrary views, 'ethics washing' policies may find it difficult to prevail given the increasing efforts to under soft and hard law through horizontal or vertical approaches. On the other hand, the convergence of principles may not necessarily lead to concrete harmonized standards and specifications. Despite this uncertainty, standard setting organisations (SSOs) could foster standards and specifications that certify AI systems based on ethical norms. Such certification can potentially encourage private or public organizations to adopt ethical practices. Many SSOs develop standards and manage internal processes to ensure their certified AI systems comply with certain ethical principles, such as transparency, accountability, and fairness. For instance, ISO process standards and certification, such as ISO/IC JTC 1/SC7, address the processes and procedures organisations should follow in systems design.[147] Similarly, the IEEE P7000 series of standards, particularly IEEE P7001 Transparency of Autonomous Systems, provide good options within the area of algorithmic transparency and accountability.[148] Such standards provide a process model that embodies various stages, such as system initiation, analysis, and design, within which ethical issues can be addressed by system and software developers.

The responsiveness of SSO standards and processes to the ethical needs in the legal sector is up for debate. Should SSO standards embed ethical norms by implementing or drawing on legal ethics (Conduct Rules) at the design level? In other words, is 'ethics by design' necessary or derivable as a principle under legal ethics?[149] What safeguards are necessary in earlier phases to create and implement ethical AI when legal tech companies do not follow the SSO standards or design processes? Moreover, should there be a platform for collaboration on AI ethics among various stakeholders, including SSOs, big tech companies, governmental bodies, and NGOs? Accordingly, the next section narrows its focus to the AI life cycle and explores how to address ethical challenges from a holistic perspective.

[146] Sundar Pichai, 'AI at Google: Our principles' (7 June 2018) <https://blog.google/technology/ai/ai-principles/> accessed 23 March 2023.
[147] Ansgar Koene et al, 'A governance framework for algorithmic accountability and transparency' (European Parliamentary Research Service, April 2019) 33 <https://www.europarl.europa.eu/RegData/etudes/STUD/2019/624262/EPRS_STU(2019)624262_EN.pdf> accessed 23 March 2023.
[148] Ibid.
[149] 'Ethics by design' means the technical/algorithm integration of ethical reasoning capabilities as part of the behaviour of artificial autonomous system (Dignum (n 117) 216). Ethics by design could include the requirement that traceability is ensured at all stages, thus contributing to the accountability of AI (Coeckelbergh (n 112) 162).

5. Boundaries and strategies for the legal sector: How to respond to the ethical challenges?

5.1. Revisiting the AI life cycle: How to understand ethics?

Legal technologies have been turning to a new chapter with the advent of AI, signifying a transformation from understanding of legal tech as a 'tool' to employing it as an 'agent'. The advantages of AI tools, such as speed, accuracy and cost-effectiveness, make them compelling options; but there is a need for a long-term vision on how to govern their use in the legal sector. The questions that still need to be addressed regarding 'legal ethics' go beyond revisiting lawyers' obligations under Conduct Rules and require a broader perspective that considers the ethical implications of leveraging AI. Should lawyers' Conduct Rules be limited to simply understanding the risks and benefits of using AI, as suggested by the model rules? They would then be required to upskill their AI competences concerning its limits, risks, and benefits under the guidance issued by regulatory bodies. This perspective, however, advocates minor modifications to the already established ethical discourse, where lawyers are required by the duty of competence to address the implications of AI use, including necessary safeguards such as training and monitoring. From this narrow perspective, lawyers are not obligated to interact with other AI stakeholders during the design, data curation, or creation of AI models. Although the former would bear the professional ethical implications entirely on their own, their lacking a role during AI life cycle would mean disparate and disconnected ethical discourses.

Therefore, a holistic approach to ethics is recommended to address wide-ranging ethical considerations that arise in AI design and development, including issues of transparency, fairness, and accountability, as well as corresponding soft and hard law measures. The need for a holistic approach becomes more urgent with the growth of generative AI and its encroachment on the legal sector.[150]

Presently, lawyers have little or no involvement in the various stages of the AI lifecycle, yet they are tasked with using AI competently and independently, safeguarding integrity, honesty, confidentiality while following the best interests of their clients, along with ensuring standards through supervision. Overall, lawyers' limited involvement in the AI life cycle is in stark

[150] See Stokel-Walker (n 25).

contrast to their significant role and responsibility in providing legal services that increasingly rely on AI. This means a quandary across the roles and responsibilities on the part of the AI stakeholders, as demonstrated in Figure 1.

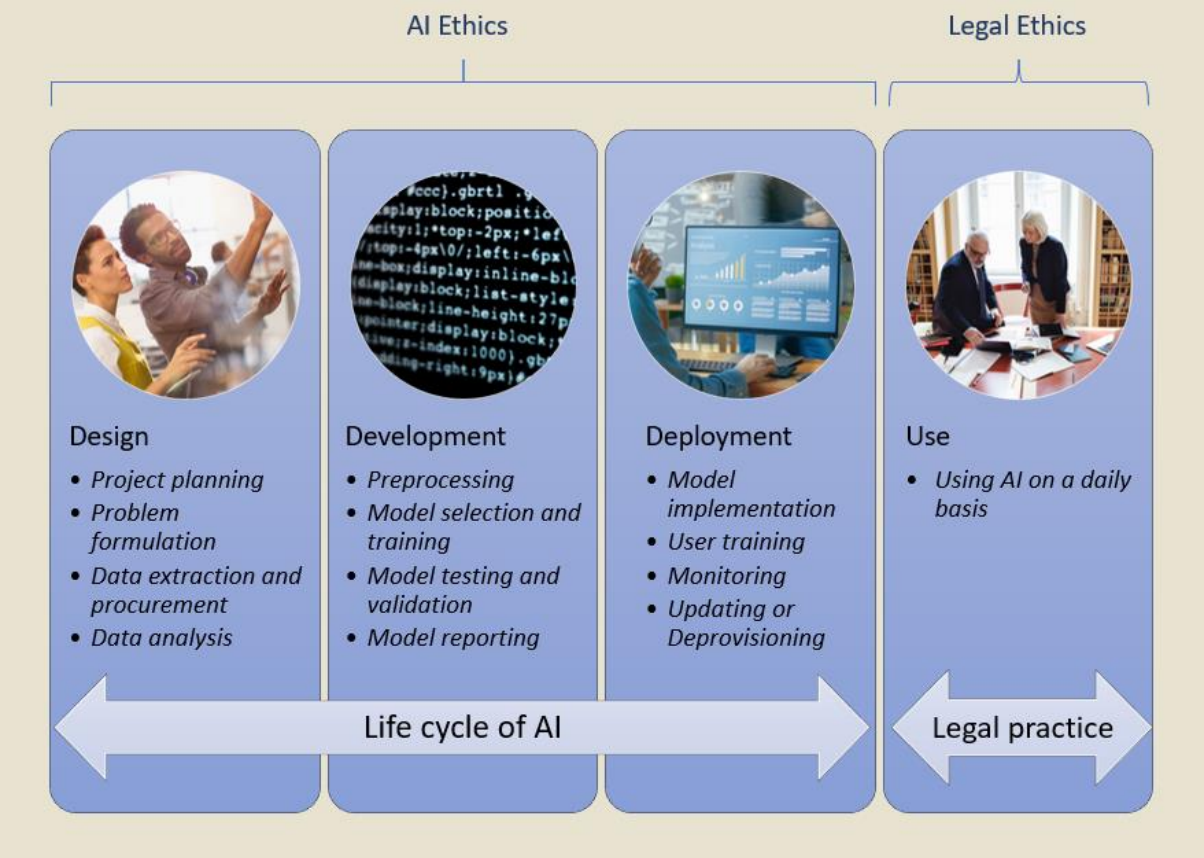

Figure 1: Stages within the life cycle of AI[151]

The AI life cycle encompasses several stages that system and software developers typically undertake with minimal involvement of AI users. It begins with the 'design' stage, where a project is planned, formulated and carried out with a view to create an AI model. This model goes through further stages of training, testing and validation. Next comes the 'development' stage, which encompasses model implementation, user training, monitoring, updating or deprovisioning.[152] Then, the AI system or software is ready for 'deployment', which incorporates user training, monitoring, updating or deprovisioning. While it can be accepted that the lack of significant input from users is due to practical and technical reasons, it is

[151] Inspired by Leslie et al (n 17) 9-11.
[152] Leslie et al (n 17).

important to recognise that AI ethics cannot be isolated from AI use, when evaluating ethical norms and obligations.

In summary, although lawyers have professional obligations that reach out to the key stages of the AI life cycle, their influence on the creation of AI tools is limited. Nonetheless, they bear the responsibility of ensuring their clients are not adversely affected by the use of AI in legal practice. Currently, there is limited interaction between lawyers and industry stakeholders to promote ethical AI based on a mutual understanding, which is a key criticism.

5.2. Moving towards a holistic view of 'AI ethics' and 'legal ethics': How to harmonise?

The absence of interaction between 'AI ethics' and 'legal ethics' is a critical concern that needs revisiting. Both ethical strands have their own formation and implementation discourses, but they do not factor each other in. Consequently, the way and the extent to which stakeholders representing each discourse interact with each other is crucial to harmonising AI ethics and legal ethics. Harmonisation between these ethical strands cuts across a two-pronged analysis entailing review of each discourse.

Regarding AI ethics, governance mechanisms, standards and processes developed by industry stakeholders should be assessed in several ways. The mainstream approach is for industry stakeholders to focus on the goals of company ecosystems by pursuing ethics-oriented approaches in their design, development and deployment of AI.[153] However, AI developers and providers would find that despite the availability of detailed frameworks, a plethora of practical challenges still remain before standards can be implemented in their data science processes and procedures.[154] Given the lack of standards and underlying practical challenges, one can conclude that AI ethics is a field denoting fragmentation despite being based on converging principles, namely transparency, fairness and accountability. Further, AI ethics is an issue that needs to be considered at the design level by industry stakeholders, which can ideally be achieved by means of standards and specifications. Irrespective of whether legal tech

[153] Koene et al (n 147) 24.
[154] Krijger et al, 'The AI ethics maturity model: a holistic approach to advancing ethical data science in organizations' AI and Ethics <https://doi.org/10.1007/s43681-022-00228-7> accessed 23 March 2023.

companies develop their AI products following standards of the SSOs or create their respective ethics-oriented AI solutions, one can question whether such products are aligned with lawyers' understanding of ethics, that is, the discourse underlying legal ethics.

To re-emphasize, the current Conduct Rules predominantly view AI as a mere technological tool for lawyers to use and lawyers are usually not concerned, nor familiar with, the life cycle of these AI products. Although it is not their responsibility to scrutinise whether the abovementioned principles are met by the AI providers, this situation can be criticised not only for the disparate ethical discourses but also for the lack of interaction. Such criticism resulting from the disparate ethical discourses and ensuing lack of interaction can also be made for the AI ethics and providers. Overall, one can derive two suggestions from these key findings. First, there is a pressing need to re-evaluate Conduct Rules to ensure alignment with the professional ethical duties underlying the use of AI in the legal sector. Second, the AI stakeholders, e.g. SSOs, AI developers, and providers, ought to interact to seek the ways how to revitalise and apply a harmonised understanding of ethics for the legal sector.

5.3. Interactions between 'AI ethics' and 'legal ethics': How to revitalise?

Although there is already some interaction between industry stakeholders and lawyers during the 'deployment' stage, it is often controlled by industry stakeholders, such as legal tech companies. These stakeholders have the option to involve lawyers or not, which means that the involvement of lawyers in the process is not always guaranteed. They would likely prefer to acquire data and feedback from lawyers as passive users rather than involving them as active participants. Fruitful engagement would permit AI developers/providers to tailor their tools and services for lawyers, resulting in a more effective and personalised experience. Nonetheless, lawyers find themselves on the outskirts of the AI lifecycle, confined within the bounds of the 'use' stage, and only able to provide limited feedback to the developers/providers.

To combat this constraint, an active role can be rendered for lawyers by expanding and enriching feedback channels. Lawyers would be able to not only to share passive data but also provide feedback to the industry stakeholders regarding their proposed design choices and/or data quality after (re)using AI. Such active contribution would offer new insights for

stakeholders to re-design or re-develop their AI model, including altering the model's variables. In this context, performance metrics could also prompt new insights based on the user feedback. Lawyers, as AI users, can be deemed 'implementors,' as their feedback can be used not only to monitor AI usage, but also to review the variables and overall performance of the the AI model, and if necessary, to re-design it.

Passive or active contribution do not necessarily imply direct and active involvement in the design or development processes. To increase the effectiveness of both types of contribution, lawyers should be included as peers or team members in the AI life cycle to discuss key steps. This would allow for a third category of contribution -'collaboration-' where AI users and providers work together to unlock ethical AI that considers user needs and anticipates any potential ethical issues at the design level. Collaboration would effectively enable the holistic viewpoint to cope with the AI-centric ethical challenges for the legal sector, since it provides a fully-fledged medium of debate and deliberation.

Thus far, there are various examples of collaboration between the AI providers and lawyers, however, they do not fully represent the needs of lawyers in a holistic manner.[155] To ensure lawyers are effectively represented and heard by the industry stakeholders, a peer-to-peer platform is needed for collaboration, ideally led by the regulatory bodies, such as SRA in the UK. A wide range of stakeholders, including SSOs, could take part in the collaboration to discuss how to revitalise the interactions between AI ethics and legal ethics. A more inclusive approach would enable lawyers to be represented more broadly, and regulations could be clarified to better meet their needs. This would allow their voices to be effectively heard at the highest level, also setting the stage for an appropriate response from industry stakeholders. By this means, all participants would gain exposure to a diverse range of ethical issues and perspectives. Moreover, collaboration under the leadership of the regulators, would enable wider goals to be achieved, including the dissemination of best practices. This would stimulate and enable everyone to work together towards the common goal of leveraging ethical AI in the legal industry.

[155] For instance, Slaughter and May Collaborate help entrepreneurs develop legal tech products in collaboration with their in-house lawyers. Slaughter and May, 'Collaborate' (2022) <https://www.slaughterandmay.com/our-firm/innovation/collaborate/> accessed 23 March 2023.

6. Conclusion

AI is rapidly transforming the legal industry, providing lawyers with powerful tools that can outperform them in many areas. Notwithstanding, the use of AI by lawyers also raise ethical issues that go beyond daily usage. Not at least because design choices or data quality problems could cause biased, discriminatory, or misleading outputs, but also the proprietary nature of training databases and issues of algorithmic transparency can surface as other potential problems that need acceptable solutions.

Potential solutions to these ethical issues typically rest in the hands of AI stakeholders which design, develop and deploy AI, often with little involvement from lawyers. As a result, ethical norms are typically imposed directly or indirectly by industry stakeholders such as developers, SSOs, and legal tech companies. Lawyers frequently encounter AI tools in plug-and-play mode and often rely on default system choices configured by system and software developers. However, while this may be practical, addressing ethical concerns such as transparency, fairness, and accountability, require more than fine-tuning the *status quo*. As AI tools become more mainstream in legal practice, ethical AI cuts across the current discourses, requiring a robust interplay between 'legal ethics' and 'AI ethics' to address ethical challenges from a holistic viewpoint. Currently, the former tackles ethics as a subset of Conduct Rules, while the latter struggles with a lack of industry-wide standards and diverse ecosystems. This results in disparate ethical discourses under which the ethical principles, such as transparency, accountability, and fairness, are not fully harmonised and Conduct Rules have a shortcoming vision and reach.

Against this backdrop, this study suggests upholding a holistic viewpoint for the broader ethical issues related to the creation and implementation of AI in the legal sector. This requires considering lawyers not only as 'users' but also 'implementors' of AI. If legal practitioners are viewed merely as users, tasks would be limited to revisiting the Conduct Rules and addressing risks and benefits of AI-driven legal tech tools. While this can be extended to necessary training, supervision and cybersecurity measures, lawyers would still not be involved in the AI life cycle. We argue that this involvement is inevitable and necessary to develop ethical AI not only for the legal sector but also for the whole society incorporating all the stakeholders, including users, and clients. Such a partnership, can also pave the way for well-filtered codes

of conduct on the part of lawyers, reflecting on the AI trajectory, standards and governing mechanisms. Last but not least, such a collaboration can leverage ethical AI, both enabling industry stakeholders to better tailor their products and involving lawyers in AI ethics.