Singapore Journal of Legal Studies
[September 2026 Online] 1–25

ARTIFICIAL INTELLIGENCE AND TRADE MARK REGISTRATION: WHAT'S NEXT FOR THE HUMAN DECISION-MAKER?

Michael **Handler**[*]

Artificial intelligence is increasingly being used by trade mark offices around the world. Offices are developing and making available AI-based "pre-clearance" tools that are designed to provide automated "assessments" of whether applications for trade mark registration are likely to face obstacles in examination. This article considers some fundamental questions about such AI-based tools, working on the assumption that offices will continue to invest in them and consider placing greater reliance on their output. It explores some general, rule-of-law-based concerns which such tools raise, as well as more specific problems with the operation of these tools in the context of the trade mark registration system. These problems include the fact that such tools are unable to replicate what we demand from human assessors applying complex, multi-factorial and subjective legal tests, and that the "data" on which such tools are likely to be trained may include decisions that are flawed, erroneous, or based on reasoning that has been overturned.

I. Introduction

When considering the impact of artificial intelligence ("AI") on intellectual property, it is fair to say that less attention has been paid to how such technologies impact on trade mark law, as compared with copyright and patent law. To the extent that the interface between trade marks and AI has been explored in litigation, or considered by commentators, this has tended to be in the context of trade mark *infringement*. Most notably, one of the first major decided cases in the world relating to the activities and outputs of an AI model involved a significant trade mark infringement component. In *Getty Images (US) Inc v Stability AI Ltd*,[1] the High Court of England and Wales was asked to consider whether Stability, the operator of an AI model that had been trained on images including photographs bearing "Getty Images" and "iStock" watermarks, was liable for infringement of various Getty-owned UK registered trade marks when the model, following user prompts, generated synthetic images containing identical or similar watermarks. Joanna Smith J held that Stability was

---

[*] Professor and Deputy Dean (Education), School of Private and Commercial Law, Faculty of Law & Justice, UNSW Sydney. ORCID iD: <https://orcid.org/0000-0001-8980-5099>. My thanks go to David Tan for the invitation to present the paper on which this article was based at *Intellectual Property and Technology in the 21st Century: Challenges for the Next Decade*, National University of Singapore, 4–5 August 2025, and to Graeme Austin, Lyria Bennett Moses, Robert Burrell, Dev Gangjee, Justin Hughes, Megan Jones, Kimberlee Weatherall and Ronald Yu.

[1] *Getty Images (US) Inc v Stability AI Ltd* [2025] EWHC 2863 (Ch) [*Getty Images*].

liable in such circumstances, given that each watermark would be seen by the average consumer as a "commercial communication by Stability" that suggested a link in the course of trade between the synthetic image and Getty,[2] and was being used in relation to goods or services covered by Getty's registrations.[3] Although the case required the court to come to grips with the workings of a novel and complex form of technology, the decision involves a relatively straightforward application of settled law on matters such as the identity/similarity of marks, the identity/similarity of goods and services, and the likelihood of confusion.[4] Further, the decision is likely to have a relatively limited impact on the design and operation of AI models more generally. This is because Stability was found liable only in relation to a very early version of its model, before it started a "filtering" process designed to prevent the emergence of watermarks on its generated synthetic images.

Beyond the specific infringement scenario addressed in *Getty Images*, there is a growing body of scholarship that looks to a future where many purchasing decisions will be effectively outsourced to AI-powered tools that will make recommendations and purchases based on assumptions as to past consumer preferences, and asks what this might mean for our tests for trade mark infringement. Amongst other things, it is argued that the increasing role played by such AI tools calls into question the role of the "average" or "reasonable" consumer in determining how a likelihood of confusion is to be assessed, and might increase the risks of certain types of confusion amongst consumers.[5] This is a valuable strand of scholarship – in particular, it helps to shed light on a fundamental and ongoing question in trade mark laws around the world about the construct of the "average" or "reasonable" consumer, and the

---

2 *Ibid* at [398]–[399].

3 Her Honour was not, however, prepared to find that Stability's activities constituted infringement by dilution or passing off.

4 Arguably, the only issue involving a contentious application of the law related to the finding that Stability, as the operator of the AI model, was in fact "using" marks when synthetic images were generated. This is in light of *Google France SARL v Louis Vuitton Malletier SA, Viaticum Lucetiel SA and Centre National de Recherche en Relations Humaines (CNRRH) SARL* Joined Cases C-236/08 to C-238/08, [2010] ECR I-2417 (CJEU), where it was held that Google, as the party operating a keyword advertising service allowing advertisers to bid on and use keywords containing trade marks, was not using the marks in the course of trade merely by setting up the "technical conditions necessary for the use of the sign" by third parties. This issue was addressed in detail in *Getty Images*, *supra* note 1 at [335]–[351].

5 For representative examples, see Michael Grynberg, "AI and the 'Death of Trademark'" (2020) 108(2) Ky LJ 199; Lee Curtis, "Will Artificial Intelligence Become the Gatekeeper of Trade Mark Law?" (2021) 43(1) Eur IP Rev 1; Daryl Lim, "Computational Trademark Infringement and Adjudication" in Ryan Abbott, ed. *Research Handbook on Intellectual Property and Artificial Intelligence* (Cheltenham: Edward Elgar, 2022) 259 [Lim, "Computational Trademark Infringement"]; Anke Moerland & Christie Kafrouni, "Online Shopping with Artificial Intelligence: What Role for Trade Marks?" in Ryan Abbott, ed. *Research Handbook on Intellectual Property and Artificial Intelligence* (Cheltenham: Edward Elgar, 2022) 290; Rob Batty, "Trade Mark Infringement and Artificial Intelligence" (2022) 26(3) NZBLQ 137; Christine Haight Farley, "Trademarks in an Algorithmic World" (2023) 98(4) Wash L Rev 1123; Alpana Roy & Althaf Marsoof, "Removing the Human from Trademark Law" (2024) 55(5) Intl Rev IP & Competition L 727; Alexandre Piron, "Do We Still Need Brands? Revisiting the Essential Function of Trade Marks and the Consumer Model in AI-mediated Commerce" (2026) 21(2) J Intell Prop L & Prac 83. See also Dev S Gangjee, "Panoptic Brand Protection? Algorithmic Ascendancy in Online Marketplaces" (2024) 46(7) Eur IP Rev 448 (on the increasing use by online marketplaces of AI-enabled filtering tools to detect and remove counterfeit products).

extent to which the law should seek to ensure that this construct reflects the likely behaviours of actual consumers (or automatons) or also serve other normative goals, such as ensuring freedom of competition and expression.[6] These issues are complex but likely to be addressed over time as infringement cases come before tribunals around the world for determination.

In this article, I wish to consider a different set of issues about the impact of AI on trade mark law, which has received more limited attention in the literature to date. This relates to the role that AI is playing, and might continue to play, in the trade mark *registration* process.[7] My focus is primarily on the increasing use of AI by government agencies, in particular the AI-based "pre-clearance" tools that some of these agencies are developing and making available for public release. As will be explained, these tools provide automated "assessments" of whether an application for registration of a trade mark is likely to face obstacles in examination. These obstacles include a finding that the mark in question lacks distinctiveness, and/or that it might conflict with an earlier mark – these being two of the most contentious and difficult issues that tend to arise pre-registration. Although such tools are being developed cautiously, and without any immediate suggestion that they are intended to supplant the role of a human decision-maker in the registration process, it is possible to imagine a future where such tools become closely integrated into decision-making workflows and practices within government agencies. In that sense, such tools have the potential to revolutionise the registration process. Whether that will lead to a revolution in the quality of decision-making within that process is, however, a much more open question.[8] The issues to which such automated tools, and their use by trade mark offices, potentially give rise are complex, and worthy of close and critical investigation.

I start in Part II with a brief overview of how trade mark offices around the world have invested in novel technologies and sought to embed those technologies within

---

6 See generally Graeme B Dinwoodie & Dev S Gangjee, "The Image of the Consumer in EU Trade Mark Law" in Dorota Leczykiewicz & Stephen Weatherill, eds. *The Images of the Consumer in EU Law: Legislation, Free Movement and Competition Law* (Oxford: Hart Publishing, 2015) 339; Kimberlee Weatherall, "The Consumer as the Empirical Measure of Trade Mark Law" (2017) 80(1) Mod L Rev 57 [Weatherall, "The Consumer as the Empirical Measure"]; Robert Burrell & Kimberlee Weatherall, "Towards a New Relationship between Trade Mark Law and Psychology" (2018) 71(1) Current Leg Probs 87; Graeme B Dinwoodie, "Trade Mark Law as a Normative Project" [2023] Sing JLS 305. See further Part IV below.

7 Notable foundational scholarship on this topic includes Dev S Gangjee, "Eye, Robot: Artificial Intelligence and Trade Mark Registers" in Niklas Bruun *et al*, eds. *Transition and Coherence in Intellectual Property Law: Essays in Honour of Annette Kur* (Cambridge: Cambridge University Press, 2021) 174 [Gangjee, "Eye, Robot"]; Anke Moerland & Conrado Freitas, "Artificial Intelligence and Trademark Assessment" in Jyh-An Lee, Reto Hilty & Kung-Chung Liu, eds. *Artificial Intelligence & Intellectual Property* (Oxford: Oxford University Press, 2021) 266 [Moerland & Freitas, "Artificial Intelligence and Trademark Assessment"]; Dev S Gangjee, "A Quotidian Revolution: Artificial Intelligence and Trade Mark Law" in Ryan Abbott, ed. *Research Handbook on Intellectual Property and Artificial Intelligence* (Cheltenham: Edward Elgar, 2022) 325 [Gangjee, "Quotidian"]; Ángela Martínez López, "An Artificial Intelligence Toolkit for Future-proof Trade Mark Offices across the Globe" (2024) 46(6) Eur IP Rev 382; and Enrico Bonadio & Akshita Rohatgi, "Trademarks and Artificial Intelligence: Some Preliminary Considerations" (2025) 47(11) Eur IP Rev 657 [Bonadio & Rohatgi, "Trademarks and Artificial Intelligence"].

8 *Cf* Bonadio & Rohatgi, "Trademarks and Artificial Intelligence", *supra* note 7 at 659.

their work practices over the past decade. It will be seen that there has been something of a shift in this regard: while offices initially sought to invest in and deploy new technologies to assist them with standard operations forming part of the administration of the registration system, such as searching the register and classifying goods and services, they have in more recent years sought to develop AI-based tools that could be used to *automate* aspects of decision-making processes within that system. This Part explores some of the automated tools that have been developed in key jurisdictions, such as Australia's "TM Checker" service, explaining how these tools operate and how they have been used to date. I also highlight some of the early criticisms that have been levelled at these tools, especially as to the accuracy and utility of the outputs they are producing, and the signals offices are sending to would-be applicants for registration by the way these tools are being publicly presented.

I then turn to consider some of the more fundamental questions we need to ask about such AI-based tools, working on the assumption that offices will continue to invest in their development and consider placing greater reliance on their outputs in office decision-making. In Part III, I draw on the wider literature on the challenges posed by automated decision-making within government agencies, to explore some of the general, rule-of-law-based concerns to which tools designed to automate assessments of the registrability of trade marks give rise. These go to matters such as the transparency of the design of such tools and their methods of producing outputs, the accountability of agencies relying on such tools, and the fundamental importance of requiring comprehensible reasons for decisions.

In Parts IV and V, I focus more specifically on the functioning and limitations of AI-based tools that can be used to automate assessments of trade mark registrability. First, in Part IV, I address the differences between what such AI-based tools are capable of doing and what we demand of human assessors, both in terms of the legal reasoning skills we ask of them and the obligations we place on them to explain their decisions. This calls into question the extent to which it is possible to design systems that can, in essence, replicate the results that might be reached by a human decision-maker on complex, multi-factorial and subjective matters – such as the extent to which a mark is sufficiently inherently distinctive, or whether two marks are "confusingly similar" – and forces us to confront what it would mean to rely on such outputs in the absence of a set of reasons that led to them. Then, in Part V, I consider a related difficulty with the operation of such AI-based tools, which concerns the datasets on which they are likely to be trained (that is, past decisions of examiners and tribunals on the legal issue in question). The assumptions underpinning projects to develop such automated tools are that these datasets are likely to be relatively stable, and that they provide an accurate reflection of the state of the law. I seek to challenge those assumptions about stability and accuracy by considering situations where AI tools have been trained on datasets involving decisions based on faulty reasoning, or which involve significant internal inconsistency. In addition, because of the way the doctrine of precedent operates, it will be suggested that particular problems arise where the law subsequently changes as a result of new court decisions, in ways that must "falsify" some of the data on which an AI model has been trained. In doing so, I seek to show that the very prospect of automated decision-making in the trade mark registration process makes it more important than ever that we critically scrutinise the state of our trade mark laws (for example, on when marks are confusingly similar, or on inherent distinctiveness) and focus our attention on whether those laws are operating in an optimal manner.

I conclude with some brief thoughts in Part VI about some of the challenges involved for decision-makers and their managing organisations in working out whether and how to integrate AI-generated assessments of similarity and distinctiveness in examination practice.

II. AI in Trade Mark Offices: What's Been Happening? What's Next?

Over the past decade, both the private sector and government agencies have invested in the development of new technologies to assist with various aspects of the trade mark registration process.[9] Initially, the focus of much of this investment was on finding ways of harnessing new technologies, especially machine learning and natural language processing, to improve tasks routinely undertaken as a part of that process.

One such task is that of searching registers for potentially conflicting marks – something commonly undertaken by private parties for clearance and monitoring purposes, and by trade mark offices primarily in the course of examining applications for registration. Searching registers has long presented challenges.[10] This is not simply because of the sheer volume of marks contained on registers. It is also because the accuracy of any search has historically depended on both prior administrative decisions (for example, how precisely an image-based mark has been "indexed"[11]) and an appreciation of the complexity of the legal tests as to when marks are in conflict (which generally turn on the interplay between the visual, aural and conceptual similarities between marks). It is therefore understandable that both the private sector and government agencies focused their energies on using new technologies to improve the quality of searches of the register to identify potentially "similar" marks. This has led to considerable advances in search practices. For example, Dev Gangjee observed in 2021 that the development and adoption of new, AI-assisted image recognition systems had enabled offices around the world to move beyond conducting searches for pictorial marks using more limited, text-based methods, such as those contained in the International Classification of the

---

9 For a rolling list of various AI-related initiatives being undertaken in agencies around the world, see World Intellectual Property Organization, "Index of AI Initiatives in IP Offices" <https://www.wipo.int/en/web/ai-tools-services/ipos-initiatives> [WIPO, "Index of AI Initiatives"]. For an earlier list, see International Trademark Association, Artificial Intelligence Subcommittee of the Emerging Issues Committee, *Adoption of Artificial Intelligence by IP Registries* (2023) [INTA, *Adoption of Artificial Intelligence by IP Registries*]. On the use of privately-developed AI tools designed to assist businesses and their advisors, see Sonia K Katyal & Aniket Kesari, "Trademark Search, Artificial Intelligence and the Role of the Private Sector" (2021) 35(2) BTLJ 501. On the ethical issues raised by the use of AI by professionals conducting trade mark work, see Rob Batty, "AI and Trade Mark Practice", *University of Auckland Faculty of Law Research Paper Series* (9 March 2026) [Batty, "AI and Trade Mark Practice"].

10 For an historical perspective, see Jose Bellido & Hyo Yoon Kang, "In Search of a Trade Mark: Search Practices and Bureaucratic Poetics" (2016) 25(2) Griffith L Rev 147 at 157–166.

11 That is, in the sense of being described in text-based form by offices during the registration process.

Figurative Elements of Marks, thus improving the overall quality and accuracy of image-based searches.[12]

Another task in the registration process that offices have sought to streamline through the use of new technologies is the drafting and classification of specifications of goods and services. For such purposes, most countries use the International Classification of Goods and Services established under the Nice Agreement.[13] The Nice Classification not only divides all goods into 34 classes and all services into 11 classes, but also contains a detailed list of specific product descriptions falling within each class, with over 11,000 items listed across the 45 classes.[14] Over and above this, many offices have developed their own databases, consisting in many cases of tens of thousands of additional product descriptions, grouped by Nice class.[15] To assist applicants for registration, offices have developed tools using natural language processing that are designed to provide guidance to potential applicants for registration by providing recommendations as to suitable language that might be used to describe their goods or services and as to the appropriate Nice class for those specified goods or services.[16]

The key point to note about the above technologies is that they were adopted and are being utilised to streamline and improve aspects of the registration process, but not to supplant any of the decision-making involved in that process. That is, while new tools have been used to enhance the quality of searches to identify potentially conflicting marks, and new technologies have been deployed to guide applicants in their choice of language for their specifications, the actual decisions as to whether identified marks are in fact in conflict, or whether goods and services have been clearly specified and correctly classified, have remained with human decision-makers.

More recently, however, offices and private actors have moved beyond developing what might be described as "assistive" technologies, and have sought to invest in new AI-based technologies, based primarily on machine learning and natural language processing, that are designed to make a range of automated "assessments" on

---

12 Gangjee, "Eye, Robot", *supra* note 7 at 180–183 (referring to image-search technologies adopted in the USA and Australia and by the EUIPO). See also Moerland & Freitas, "Artificial Intelligence and Trademark Assessment", *supra* note 7 at 269–278 (describing and testing the performance of search tools used by offices in the EU, Australia and Singapore and by the World Intellectual Property Organization ("WIPO")); Thomas Vandamme, Julien Cabay & Olivier Debeir, "A Quantitative Evaluation of Trademark Search Engines' Performances through Large-Scale Statistical Analysis" (*Conference Paper, ICAIL '23: Proceedings of the Nineteenth International Conference on Artificial Intelligence and Law*, 19–23 June 2023) at 2 <https://doi.org/10.1145/3594536.3595149> [Vandamme, Cabay & Debeir, "A Quantitative Evaluation"] (referring to the image recognition technologies being used in the EU and Benelux offices); Batty, "AI and Trade Mark Practice", *supra* note 9 at 7 (referring to search tools produced by private parties).

13 Nice Agreement Concerning the International Classification of Goods and Services for the Purposes of the Registration of Marks (15 June 1957, as amended on 28 September 1979).

14 See World Intellectual Property Organization, "Nice Classification Publication" <https://nclpub.wipo.int/enfr/>; World Intellectual Property Organization, "Frequently Asked Questions: Nice Classification" <www.wipo.int/classifications/nice/en/faq.html>.

15 See Robert Burrell & Michael Handler, "Who Reads the Trade Marks Register?" (2025) 45(2) Oxford J Leg Stud 272 at 275–276 [Burrell & Handler, "Who Reads?"].

16 Gangjee, "Eye, Robot", *supra* note 7 at 178–179 (describing developments in offices in the EU, Singapore, China, Germany and Japan).

matters going to the registrability of marks.[17] While these matters differ according to the tool in question, the two matters that appear to be most commonly assessed are whether a chosen mark is inherently distinctive, and whether it is "confusingly similar" to an earlier registered mark, which involves considering similarities between both the marks themselves and the goods or services in question. This is a much more far-reaching development, in that it raises the spectre of such technologies being used by government agencies to produce preliminary evaluations on the registrability of marks that can be taken into account by human decision-makers within such agencies, or potentially to produce "decisions" on registrability.

By way of illustration of what such technologies are and how they operate, a fascinating example is the model developed by Shivam Adarsh, Elliott Ash, Stefan Bechtold, Barton Beebe and Jeanne Fromer. The intention of this model is to see if the assessment of whether a word mark is inherently distinctive under US law could be automated.[18] This model was trained on a dataset of over one million US applications for registration made between 2012–2017, and used natural language processing to generate assessments of whether several hundred thousand word marks were inherently distinctive under US law, with these results being compared with actual examination decisions relating to such marks made during 2018–2019.[19] The authors reported that their model was 86% accurate overall, with confidence in the model's predictions ranging from around 93% for clear-cut cases (*ie*, those where the model considered that there was a 90% probability that the mark was inherently distinctive), down to 66% for borderline cases (*ie*, those where the model considered that there was only a 50% probability of the mark being inherently distinctive).[20]

Among trade mark offices, the European Union Intellectual Property Office ("EUIPO") has played a leading role in developing automated "pre-assessment" tools and services. In June 2025 it announced that it had updated its EUTM EasyFiling service to include an "AI-powered pre-assessment service".[21] This allows a user to enter details of a proposed application for registration (in the form of a mark and draft specification), and is designed to identify earlier EU trade marks with an "identical verbal denomination and at least one matching class of goods and services by checking it against entries in the TMView database".[22] It is also designed

---

17 Whether offices in the future might consider whether Large Language Models, trained on datasets consisting of legal decisions, and supplemented by Retrieval-Augmented Generation, can be used to generate "reasoned" decisions on these matters, is beyond the scope of this article. As to the problems with such projects, see, *eg*, Matthew Dahl *et al*, "Large Legal Fictions: Profiling Legal Hallucinations in Large Language Models" (2024) 16(1) J Legal Analysis 64.

18 Shivam Adarsh *et al*, "Automating *Abercrombie*: Machine-learning Trademark Distinctiveness" (2024) 21(4) J Empirical Leg Stud 826 [Adarsh *et al*, "Automating *Abercrombie*"].

19 *Ibid* at 837.

20 *Ibid* at 841–845. For an example of a model developed in Canada to assess "likelihood of confusion" under EU law, see Samuel Dahan *et al*, "Analytics and European Union Courts: The Case of Trademark Disputes" in Tamara Capeta, Iris Goldner Lang & Tamara Perišin, eds. *The Changing European Union: A Critical View on the Role of Law and the Courts* (Oxford: Hart Publishing, 2022) 77.

21 EU Intellectual Property Office, "Real-time Alerts and AI Plug-ins in EUTM EasyFiling Form Help Reduce Uncertainty, Improve Predictability and Enhance the User Experience" <https://www.euipo.europa.eu/en/news/real-time-alerts-and-ai-plug-ins-in-eutm-easyfiling-form> (30 June 2025) [EUIPO, "Real-time Alerts"].

22 *Ibid*.

to indicate whether the mark raises offensiveness, morality or public policy grounds for rejection, or whether it is deceptive as to the nature, quality or geographical origin of the specified goods or services.[23] In the same announcement, it was said that the tool "demonstrates the EUIPO's approach to the gradual adoption of AI into its systems, rooted in ethical and responsible practices" and also that "[t]he results of the pre-assessment services … never replace the final decision taken by a human EUIPO examiner".[24] The United Kingdom Intellectual Property Office ("UKIPO") offers a similar AI-based tool, originally called "Trade Mark Pre-apply".[25] This tool:

> [i]s aimed at assisting and educating novice users with their application; it helps identify goods and services to protect any aspects of a trademark that might not be appropriate (such as offensive words or protected symbols …), and if there are any similar trademarks that could cause conflict. It does not provide legal advice, and does not prevent a user submitting an application, as a trademark examiner always examines the submitted trademark.[26]

Perhaps the most well-developed AI-assisted service is IP Australia's "TM Checker".[27] IP Australia has, for some time, been interested in exploring how new technologies such as machine learning and natural language processing can be used in the development of public-facing trade mark-related tools and products. In 2017, it launched its "TM Assist" service, which was designed to guide potential applicants through the initial stages of the registration process[28] by generating search results for potentially conflicting marks and helping applicants with their choice of language for their specifications.[29] Following the results of a 2021 study showing that small businesses lacked both an awareness of the importance of trade mark registration and confidence as to how to secure rights in their marks,[30] IP Australia invested in a new AI-assisted product that was formally launched in March 2024 as TM Checker. This publicly available product is prominently advertised on the home page for Australian Trade Mark Search ("ATMS"), the main search tool used to access information on the Australian Register,[31] and is presented as an alternative

---

23 *Ibid*.

24 *Ibid*.

25 See UK Intellectual Property Office, "How Can Artificial Intelligence Improve Customer Experience at the IPO?" <https://ipo.blog.gov.uk/2020/04/17/how-can-artificial-intelligence-improve-customer-experience-at-the-ipo/> (17 April 2020); UK Intellectual Property Office, "Introducing the Trade Mark Pre-apply Service" <https://ipo.blog.gov.uk/2020/10/29/introducing-the-trade-mark-pre-apply-service/> (29 October 2020). The tool can be accessed at <https://trademarks.ipo.gov.uk/ipo-improve>.

26 WIPO, "Index of AI Initiatives", *supra* note 9.

27 IP Australia, "TM Checker" <https://www.ipaustralia.gov.au/trade-marks/search-existing-trade-marks/tm-checker> [IP Australia, "TM Checker"].

28 Australian Government, Department of Industry, Innovation and Science, *Annual Report 2017–18* (2018) at 212.

29 INTA, *Adoption of Artificial Intelligence by IP Registries*, *supra* note 9 at 5; Moerland & Freitas, "Artificial Intelligence and Trademark Assessment", *supra* note 7 at 272, 277, 290.

30 Australian Government, Australian Public Service Commission, *State of the Service Report 2024–25* (2025) at 92 [APSC, *State of the Service Report 2024–25*].

31 IP Australia, "Australian Trade Mark Search" <https://search.ipaustralia.gov.au/trademarks/search/quick>.

to ATMS elsewhere on IP Australia's website.[32] TM Checker allows a user to input their desired sign (either a word or figurative mark) and specification of goods and services (limited to a "pick list" of terms). Following this, the user is provided with: an automated distinctiveness "warning" if the tool considers that the mark might lack inherent distinctiveness; and/or an automated similarity "warning" if potentially conflicting marks on the Australian Register have been identified, with up to 20 conflicts being listed. Users who wish to pursue an application for registration at this point are taken to a page that explains the types of application that can be made and the fees involved,[33] and contains a further link to the login page for applications for registration.[34]

It is important to emphasise that, at the moment, these services offered by the EUIPO, UKIPO and IP Australia provide only high-level "pre-clearance" information to potential applicants for registration. That is, they are presented as "guides" that are intended to "simplify the user experience and provide efficiencies for users",[35] but without guarantees as to legal accuracy or how an examiner might decide on an application for registration of the mark in question.

Even taking this into account, what is striking about many of these tools is how inaccurate they are. For example, Chris Sgourakis undertook numerous searches on TM Checker in January 2025 and identified problems with the quality of the results generated on distinctiveness, similarity of marks and similarity of goods and services, concluding that it is a "basic search tool with limitations"[36] and "not a substitute for a full search of the Register"[37] on ATMS. Indeed, spending any time on TM Checker demonstrates that it consistently generates large numbers of false positives (especially in the number of unwarranted "warnings" it provides about potentially conflicting marks) as well as false negatives (especially in relation to distinctiveness).[38] This is consistent with the findings of Thomas Vandamme, Julien Cabay and Olivier Debeir in their studies of the functionality of the "image search"

---

32 IP Australia, "How to Search Existing Trade Marks" <https://www.ipaustralia.gov.au/trade-marks/search-existing-trade-marks>.

33 IP Australia, "TM Checker: Options for Applying" <https://www.ipaustralia.gov.au/trade-marks/search-existing-trade-marks/tm-checker/options> [IP Australia, "TM Checker: Options for Applying"].

34 See IP Australia, "TM Checker: Options for Applying: Apply with Headstart" <https://www.ipaustralia.gov.au/trade-marks/search-existing-trade-marks/tm-checker/headstart> [IP Australia, "TM Checker: Options for Applying: Apply with Headstart"] and IP Australia, "TM Checker: Options for Applying: Apply with Standard" <https://www.ipaustralia.gov.au/trade-marks/search-existing-trade-marks/tm-checker/standard> [IP Australia, "TM Checker: Options for Applying: Apply with Standard"]. Currently, on neither page is there any reference to ATMS.

35 See IP Australia, "TM Checker Privacy" <https://www.ipaustralia.gov.au/about-us/our-agency/privacy/tm-checker-privacy>.

36 Chris Sgourakis, *Australian Trade Mark Practice: Selecting, Registering and Managing Trade Marks* (Sydney: LexisNexis, 2025) at 408 [Sgourakis, *Australian Trade Mark Practice*].

37 *Ibid* at 409. See also Gilbert + Tobin, "IP Australia's TM Checker: How to Balance Automation and Strategy in Trade Mark Registration" <https://www.gtlaw.com.au/insights/ip-australias-tm-checker-how-to-balance-automation-and-strategy-in-trade-mark-registration> (23 July 2025) ("it is important to recognise its limitations, particularly in relation to more descriptive marks or where there is crowding around use of marks").

38 For example, searches of TM Checker undertaken on 24 February 2026 and 24 March 2026 for colloquial Australian descriptive marks for or relating to "beer", being "MIDDY", "TINNY", "COLD ONE", "BREWSKI", "TURPS", "SLAB" and "CASE", did not generate any distinctiveness "warnings".

tools used by the Benelux Office of Intellectual Property and the EUIPO. In their first study, which compared the operation of the Benelux office's tool with the outcomes of 8,126 EUIPO opposition decisions from 2016–2022, they found that the Benelux tool identified potentially similar and conflicting image marks only 7.36% to 11.72% of the time.[39] In their follow-up study comparing the EUIPO's tool with 553 Benelux opposition decisions from 2011–2023, they found that the EUIPO tool identified potentially similar and conflicting image marks only 15.17% to 26.9% of the time.[40] Returning to the Adarsh *et al* model designed to make assessments of inherent distinctiveness under US law, the authors note that their 14% error rate is "relatively high" and that their model is precise only in relation to the most straightforward cases.[41]

Relatedly, it might be asked how the (often inaccurate) outputs produced by these AI-based tools will be received and internalised by potential self-filers and SMEs. For example, notwithstanding the presence within TM Checker of disclaimers and recommendations for users to seek professional assistance, it is arguable that self-filers and SMEs might misunderstand the nature of the outputs generated and place undue reliance on them. They might thus choose not to proceed with an unproblematic application for registration (where the outputs of the tool consist of false positives) or proceed with an application that will face unexpected grounds of refusal (in the case of false negatives).[42] Given that these new AI systems are purportedly designed to benefit ordinary traders, there would be something of an irony if this were to entrench a two-tiered trade mark system, where sufficiently well-resourced traders would seek advice about the outputs of an agency's pre-clearance tools and choose to proceed or not to proceed with their applications accordingly, with the remainder simply accepting the bureaucracy's algorithmically-generated and potentially inaccurate results.

Despite these problems and concerns, offices are clearly continuing to invest in and look to improve the accuracy and operation of these AI-assisted pre-clearance tools. It is hardly speculative to say that this is being done with a view to seeing whether these tools might be integrated into offices' formal decision-making processes. This would be broadly consistent with what some offices are already doing in this space. For example, IP Australia currently uses a "Trade Mark Precedent Identification" tool in examination, which relies on a combination of automated business rules and natural language processing techniques to "retrieve, rank, and display in order of relevance, *substantially identical* text trade marks from the Australian trade marks register" [emphasis added].[43] Similarly, the EUIPO has indicated that

---

39 Vandamme, Cabay & Debeir, "A Quantitative Evaluation", *supra* note 12 at 6.

40 Julien Cabay, Thomas Vandamme & Olivier Debeir, "Looking Through the Crack in the Black Box: A Comparative Case Law Benchmark for Auditing AI-Powered Trade Mark Search Engines" (2025) 59 Computer L & Sec Rev 106167:1 at 10 [Cabay, Vandamme & Debeir, "Looking Through"].

41 Adarsh *et al*, "Automating *Abercrombie*", *supra* note 18 at 828.

42 The risks of users proceeding with applications in such circumstances are increased by the design of TM Checker: see IP Australia, "TM Checker: Options for Applying", *supra* note 33, IP Australia, "TM Checker: Options for Applying: Apply with Headstart", *supra* note 34 and IP Australia, "TM Checker: Options for Applying: Apply with Standard", *supra* note 34.

43 WIPO, "Index of AI Initiatives", *supra* note 9. On the operation of the doctrine of "substantial identity" and its relationship with "deceptive similarity", see Robert Burrell & Michael Handler, *Australian*

it is making internal use of an AI-assisted "G&S Comparer" tool that draws on past decisions on the similarity of goods and services.[44] Further, it would not be surprising if offices were to look to employ new technologies to make productivity gains in administering the registration system. This is in light of the massive increase in the number of applications for registration filed worldwide between 2014–2024,[45] and recognising that even trade mark offices (many of which operate in a self-sustaining manner based on the revenue they generate) might not be immune from broader policies to reduce the size and increase the efficiency of government agencies.[46]

If offices do go down the path of considering whether AI-assisted products might be more closely integrated into their decision-making processes, the more likely scenario is that they will look to encourage examiners to use and consider the outputs of such products as part of the "arsenal of tools" at their disposal in undertaking their decision-making. However, especially if the accuracy of such tools improves over time, it cannot be discounted that offices might seek to use the outputs of the tools as *substitutes* for the decisions that are required to be made.[47] This raises a number of questions about these AI-assisted tools, their operation and their limitations, which form the basis of the remaining Parts of this article.

III. Fundamental Limitations: Transparency, Bias, Accountability, and Reasons for Decisions

The most fundamental concerns that might be raised about the use of AI-assisted tools by government agencies within the trade mark registration process are not unique to trade mark law. Rather, they are best seen as illustrative of broader, rule-of-law-based concerns about the reliance by public bodies on automated decision-making more generally. In essence, these concerns relate to transparency, or the lack thereof, as to how such automated decision-making tools in fact operate, including as to the data on which they are trained, what legal tests and standards have been built into their design, how they are generating their outcomes, the level of accuracy of these assessments,[48] and about the accountability of those relying on

Trade Mark Law, 3d ed (Sydney: LexisNexis, 2024) at [6.3]–[6.7], [9.5] [Burrell & Handler, *Australian Trade Mark Law*]. See further IP Australia, *IP Australia Artificial Intelligence and Automated Decision-making Transparency Statement*, 2d version (2026) [IP Australia, *AI and ADM Transparency Statement*] (noting that the agency currently uses AI to support "decision making or the taking of administrative action by guiding, assessing or making a recommendation to a human decision maker").

44 EU Intellectual Property Office, *2023 Consolidated Annual Activity Report* (2024).

45 World Intellectual Property Organization, "Trademarks Highlights" <https://www.wipo.int/web-publications/world-intellectual-property-indicators-2025-highlights/en/trademarks-highlights.html> (showing a 123% increase in filings over this period).

46 See, *eg*, Josh Gerben, "DOGE Threatens the Stability of the US Trademark System" <https://www.gerbenlaw.com/blog/doge-threatens-the-stability-of-the-u-s-trademark-system/> (20 February 2025).

47 Notably, Australia already recognises that a decision required to be made by the Registrar of Trade Marks can be made entirely by computer programs under the control of the Registrar: Trade Marks Act 1995 (Cth), s 222A [TMA (Cth)].

48 On the sort of information that might be expected to be disclosed to ensure transparency, both before and after release of an AI system, see Christopher S Yoo, "Beyond Algorithmic Disclosure for AI" (2024) 25(2) Colum Sci & Tech L Rev 314 at 318–321, 324–330.

such tools. Closely related to these is the broader concern that the transparency and predictability of the law, and its very authority in being able to guide human actions, requires decision-makers to be able to explain the reasons for their decisions, as distinct from announcing mere outcomes.[49]

The concern about transparency is particularly acute in relation to the use of AI tools by trade mark offices. Perhaps because these tools have, to date, been presented to self-filers and SMEs for limited "advisory" purposes, offices have not sought to provide much, if any, public information about the actual workings of such tools. For context, what public authorities in any particular jurisdiction are required to disclose about their use of AI will depend on local laws and regulatory approaches. In a number of countries, in the absence of prescriptive regulation, governments have adopted frameworks and principles on the use of AI by government agencies.[50] In Australia, by way of example, the federal government has established a policy on the responsible use of AI by non-corporate government bodies, which includes a requirement for "transparency statements" as to those bodies' adoption and use of AI.[51] IP Australia has published two versions of such a statement,[52] but although it has recently been rated as having one of the better ones in place,[53] its current statement reveals little about TM Checker, other than that it is being used.[54] The most expansive statement about TM Checker, which is contained in an Australian Public Service Commission report, discloses that it:

> [p]erforms AI-integrated steps, including:
> (a) image pre-processing (neural networks and computer vision models to perform text detection, upscaling, inpainting and text removal);
> (b) object detection through deep learning model;
> (c) a re-trained vision transformer for image evaluation against IP Australia's unique criteria; and
> (d) natural language processing for text-based evaluation.[55]

Yet this falls well short of the sort of information that would enable the public to have an understanding of the operation, accuracy and reliability of the tool. At the

---

49 For an overview of the full range of rule-of-law-based concerns raised by automated decision making, see Michael Guihot & Lyria Bennett Moses, *Artificial Intelligence, Robots and the Law*, 2d ed (Sydney: LexisNexis, 2025). See also Monica Zalnieriute, Lyria Bennett Moses & George Williams, "The Rule of Law and Automation of Government Decision-Making" (2019) 82(3) Mod L Rev 425.

50 See, *eg*, Singapore Government, Info-communications Media Development Authority and Personal Data Protection Commission, *Model Artificial Intelligence Governance Framework*, 2d ed (2020); UK Government, *Artificial Intelligence Playbook for the UK Government* (2025).

51 Australian Government, Digital Transformation Agency, *Policy for the Responsible Use of AI in Government (Version 2.0)* (2025).

52 IP Australia, *AI and ADM Transparency Statement*, *supra* note 43.

53 Kimberlee Weatherall, Jose-Miguel Bello y Villarino & Alexandra Sinclair, *AI Transparency in Practice: An Evaluation of Commonwealth Entities' Compliance with their Obligations Regarding AI Transparency Statements* (Melbourne: ARC Centre of Excellence for Automated Decision-Making and Society, 2026) at 21.

54 IP Australia, *AI and ADM Transparency Statement*, *supra* note 43 at 3.

55 APSC, *State of the Service Report 2024–25*, *supra* note 30 at 93.

most basic level, it is difficult to get even a meaningful sense of what *data* TM Checker is being trained on. The TM Checker home page refers only to "internally trained data and examinations conducted by IP Australia".[56] What is covered by this statement – that is, as to which decisions have been taken into account, within which timeframes, and whether the dataset extends to material such as court decisions or training manuals – is entirely unclear. Similar criticisms can be made about the lack of detail provided by the UKIPO in relation to its Trade Mark Pre-apply tool.

Even in jurisdictions with much stronger regulation of the use of AI, the situation appears to be much the same. For example, the EU's "AI Act" of 2024[57] sets out detailed requirements on the deployment and use of "AI systems"[58] throughout the private and public sector in the EU. Particular obligations are placed on developers and operators of "high risk" AI systems,[59] including as to transparency of the design of such systems and their outputs, with users being given the right to obtain explanations of the role of the system in any decision-making procedure.[60] However, the sort of AI tools being used or considered by the EUIPO do not currently appear to meet the threshold of being "high risk".[61] The absence of mandatory requirements of transparency gives important context to the EUIPO's limited provision of information on its website or elsewhere about the functioning of its EUTM EasyFiling service and its internal "G&S Comparer" tool.

What these examples reveal is a significant lack of transparency about how the AI tools being used by offices in fact operate. The absence of clear information about the data on which these tools have been trained and the way they produce their outputs raises difficult questions about whether biases may have inadvertently been embedded in their design. Even if it is assumed that the tools in question have been trained on exclusively "internal" decisions, a number of questions might be asked. What is the size, currency and representativeness of those datasets? Was the data "weighted" in any way, and if so how? To what extent were legal rules and standards incorporated into the designs of these tools (for example, were the models trained on the basis that particular elements of marks might be given particular weight, or be de-emphasised, in any assessment of similarity, in accordance with long-standing legal principles[62])? Who made the decisions about the design of the tools? How were these tools tested, what errors were identified, and what changes or improvements were adopted?[63] Is the government authority

---

56 See IP Australia, "TM Checker", *supra* note 27.

57 Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 Laying Down Harmonised Rules on Artificial Intelligence [2024] OJ L 2024/1689.

58 *Ibid*, Art 3.

59 *Ibid*, Art 6.

60 *Ibid*, Art 86.

61 *Cf ibid*, Annex III.

62 These might include, for example, rules that in any assessment of similarity more weight tends to be given to the beginning of words (see, *eg*, *London Lubricants (1920) Ltd's Application* (1925) 42 RPC 264 (CA, Eng)), or that descriptive or common components of words are likely to be given less weight (see, *eg*, *Cooper Engineering Co Pty Ltd v Sigmund Pumps Ltd* (1952) 86 CLR 536 (HCA)).

63 See generally Sarah Crossman & Rachel Dixon, "Government Procurement and Project Management for Automated Decision-Making Systems" in Janina Boughey & Katie Miller, eds. *The Automated*

even free to disclose any of this information if it has worked on the tool with a private sector partner?[64]

These questions about transparency matter, especially if offices look to place greater reliance on the outputs of such tools. The extent to which these questions matter in circumstances where such tools might only be used within offices as "reference points" to support the work of a human decision-maker is something to which I will return in Part VI – for now it is enough to note that even treating the outputs in such a manner carries significant risks. But if the quality of the outputs of the tools improves to a point where it is considered that they can largely replicate the outcomes that might be reached by human decision-makers, not only will these questions carry even greater importance, but we will also need to confront a different issue – what would it mean to rely on AI-produced outputs in the absence of *reasons*?

Even if an office is more transparent about how its AI tools have been trained and how they operate, this is likely to fall well short of telling an affected party *why* a particular decision was reached.[65] At a fundamental level, the accountability and authority of governmental decision-making requires the provision of sufficient, comprehensible information as to how and why a decision was reached.[66] Doing so promotes predictability in the law and is designed to ensure that decision-making is not arbitrary or biased. In the trade mark registration context, this takes the form of decision-makers providing a set of reasons for their decisions, in particular as to why a particular ground of refusal was raised. Knowing those reasons will enable an affected party to shape its conduct accordingly.[67] It might choose to contest the decision with the examiner or a more senior official by developing arguments in response (for example, as to why its mark does not in fact conflict with a cited mark, or to show that its mark does not have the meaning the examiner has ascribed to it and should be considered to be inherently distinctive). Equally, it might choose to accept the decision and adopt a different position, such as applying to register a different mark or rebranding. The provision of a mere "outcome", without reasons, casts doubt on whether the decision was in fact made in accordance with the law. It makes it harder and costlier for an impacted party to contest an adverse decision if that party does not know the grounds for the decision.[68] In short, it has the real potential to undermine the core rule-of-law principles of transparency, accountability and predictability that should underpin the law. Since the prospect of "explainable" AI that can provide plausible reasons for its decisions remains some way off,

---

State: Implications, Challenges and Opportunities for Public Law (Sydney: Federation Press, 2021) 154 at 169–170.

64 See generally Rita Matulionyte, "Government Automation, Transparency and Trade Secrets" (2024) 47(3) Melbourne UL Rev 679.

65 See, *eg*, Jenna Burrell, "How the Machine 'Thinks': Understanding Opacity in Machine Learning Algorithms" (2016) 3(1) Big Data & Society 1.

66 See, *eg*, Tatiana Dancy, "Guiding Action in the Age of Artificial Intelligence" (2025) 48(4) UNSWLJ 1236 at 1245–1246.

67 See generally Joseph Raz, *The Authority of Law: Essays on Law and Morality*, 2d ed (Oxford: Oxford University Press, 2009) at 214 (on the authority of the law requiring that "*it must be capable of guiding the behaviour of its subjects*" [emphasis in original]).

68 Equally, the affected party might incur unnecessary costs in seeking to register a different mark or rebrand by choosing to accept an adverse decision that should not have been made.

the adoption of AI tools by offices as substitutes for reasoned human decisions is something that needs to be considered with extreme caution.

IV. Design Limitations: The Gap between What AI Tools Can Produce and What Trade Mark Law Requires

The fundamental concerns about the importance of providing reasons for decisions outlined in the previous Part tie into a more specific set of concerns about the outputs of AI tools used to make "assessments" of particular legal issues within the trade mark registration process. In essence, these more specific concerns are that the outputs of such tools are not the product of the sort of *legal* reasoning that we expect of human decision-makers applying legal tests. This is because there is an inevitable gap that exists between, on the one hand, what an AI tool (even one trained on high-quality data and whose workings are transparent) can produce when asked to make an "assessment" of distinctiveness or whether two marks are in conflict, and, on the other hand, what the law demands of decision-makers, in the sense of a reasoned application of the relevant law on these issues.[69]

Here, it is important to appreciate that the AI tools under consideration are generally trained to look for and develop patterns based on and consistent with pre-existing data. For example, the expectation is that with sufficient training, a model will be able to produce an output indicating that two images are sufficiently "similar", or that a particular mark has reached a threshold of "descriptiveness" in relation to a specified set of goods or services, and that these outcomes might replicate those that would be reached by a human decision-maker after a much more laborious process. However, there is a difference between pattern recognition to generate a broadly internally consistent set of outcomes, and what we demand of the legal reasoning process in a system, based on a doctrine of precedent, containing complex, multivalent and open-textured legal rules.[70]

This gap might, at first glance, seem to be less significant for assessments of inherent distinctiveness. In jurisdictions such as the EU, UK, Singapore and Australia, this assessment looks to what appears to be a limited question of the semantic relationship between a sign and a specified set of goods or services.[71] It might therefore appear that an AI-assisted tool, trained through natural language processing on a sufficiently large dataset, could be well placed to provide a fairly

[69] This issue has been well explored in some of the foundational literature on the topic: see, *eg*, Gangjee, "Eye, Robot", *supra* note 7 at 185–188; Moerland & Freitas, "Artificial Intelligence and Trademark Assessment", *supra* note 7 at 284–287; Bonadio & Rohatgi, "Trademarks and Artificial Intelligence", *supra* note 7 at 660.

[70] See, *eg*, Mireille Hildebrandt, "Law as Computation in the Era of Artificial Legal Intelligence: Speaking Law to the Power of Statistics" (2018) 68 supp 1 UTLJ 12 at 21–23. See especially at 23 ("Whereas machines may become very good in … simulation, judgment itself is predicated on the contestability of any specific interpretation of legal certainty in the light of the integrity of the legal system, which goes way beyond a quasi-mathematical consistency").

[71] See Regulation (EU) 2017/1001 of the European Parliament and of the Council of 14 June 2017 on the European Union trade mark (codification) [2017] OJ L154/1, art 7(1)(*b*)–(*c*) [EUTMR]; Trade Marks Act 1994 (UK), s 3(1)(*b*)–(*c*) [TMA (UK)]; Trade Marks Act 1998 (2020 Rev Ed) (Singapore), s 7(1)(*b*)–(*c*) [TMA (S'pore)]; TMA (Cth), s 41.

accurate assessment of whether any given mark is or is not inherently distinctive in relation to any given set of goods or services, given the patterns it might be able to detect from a sufficiently large and useful dataset of prior decisions. The Adarsh *et al* model, discussed in Part II, which appeared to perform well in clear-cut cases, might appear to lend support to such a conclusion.

However, on closer inspection, the situation is much more complex. In cases that clearly involve non-descriptive marks, an AI-assisted tool will be of no benefit. A human decision-maker will be able to determine with negligible intellectual effort that such a mark is inherently distinctive – a decision-maker will know almost instantly that "QYSTROX" for pharmaceuticals or "GORILLA" for fruit juice will be inherently distinctive, and an AI-based "assessment" to the same effect will add little. An AI-based tool could only ever usefully supplement human decision-making on inherent distinctiveness in borderline cases. In such cases, the law asks decision-makers to consider how notional consumers would understand the sign in question and whether they would recognise it to serve a trade mark function[72] or whether the sign is one that other traders are likely to wish to use for the sake of what it ordinarily signifies.[73] These tests, especially the latter, are underpinned by normative considerations as much as empirical ones, and they have proven to be highly complex to apply in fact-specific determinations of whether particular marks are or are not inherently distinctive. This is especially the case for words that might be understood to have a geographical meaning, or are neologisms, or are in a foreign language, or are slang terms, or are slogans.[74] The extent to which an AI model might be able to be trained in a way to produce meaningful outcomes in such fact-dependent borderline cases – in circumstances where a human decision-maker might need to engage in research as to how the sign in question would be understood and exercise a degree of judgment in considering whether it would be perceived as a mark or whether it would be a term that other traders might wish to use in relation to the specified goods or services – is highly doubtful. It is notable that Adarsh *et al* recognise that their model "performs poorly" in borderline cases, mimicking the uncertainty that human decision-makers experience in assessing such marks.[75]

Additional difficulties are likely to arise in developing AI tools that seek to make assessments of the distinctiveness of marks other than word marks (which were outside the scope of Adarsh *et al*'s project). For example, in the case of figurative marks that combine descriptive and distinctive elements, it is often difficult for a human

---

[72] In the EU, on the "distinctive character" ground, see *Linde AG, Winward Industries Inc and Rado Uhren AG* Joined Cases C-53/01 to C-55/01, [2003] ECR I-3161 (ECJ) at [41], [47]. In Singapore, see *Societe Des Produits Nestlé SA v Petra Foods Ltd* [2017] 1 SLR 35 (CA) at [33].

[73] In the EU, on the "descriptiveness" ground, see *Windsurfing Chiemsee Produktions- und Vertriebs GmbH v Boots- und Segelzubehör Walter Huber and Attenberger* Joined Cases C-108/97 and C-109/97, [1999] ECR I-2779 (ECJ) at [25]–[26], [30]. In Australia, see *Cantarella Bros Pty Ltd v Modena Trading Pty Ltd* (2014) 254 CLR 337 (HCA) at [59], [71].

[74] On the challenges involved in assessing for inherent distinctiveness in Australia, the UK/EU, and Singapore, see Burrell & Handler, *Australian Trade Mark Law*, *supra* note 43 at [4.1]–[4.19]; Lionel Bently *et al*, *Intellectual Property Law*, 6th ed (Oxford: Oxford University Press, 2022) at 992–1016 [Bently *et al*, *Intellectual Property Law*]; Ng-Loy Wee Loon, *Law of Intellectual Property in Singapore*, 3d ed (London: Sweet & Maxwell, 2022) at [21.3.22]–[21.3.38] [Ng-Loy, *Law of Intellectual Property in Singapore*].

[75] Adarsh *et al*, "Automating *Abercrombie*", *supra* note 18 at 857.

decision-maker to determine how the mark *as a whole* would be perceived, and the weight that should be given to the various elements.[76] For non-traditional marks such as shapes and colours, where relevant tests often look to the extent to which such shapes and colours depart from the customs or norms of the sector,[77] an AI tool will be of minimal assistance. Further, in some jurisdictions such as Australia, decision-makers who have found that a mark lacks inherent distinctiveness are still required to identify the *extent* of the mark's descriptiveness in the course of conducting the factual distinctiveness inquiry.[78] Quite different consequences attach to whether a non-distinctive mark is classified as being to no extent adapted to distinguish the specified goods (such as "MONTE CARLO" for gambling services) as distinct from being adapted to a slight extent, but not enough to make the mark inherently distinctive (such as "MONTE CARLO" for ice cream). This is a complex task with which Australian decision-makers routinely struggle, and it is entirely unclear how an AI-generated output could be expected to provide a meaningful or helpful assessment of the extent of a non-distinctive mark's adaptation to distinguish, as required by Australian law.

The gap between what the law demands of human decision-makers and what an AI tool can produce is arguably even greater when considering conflicts between marks. Such conflicts are addressed differently throughout the world, but a common feature tends to be that decision-makers are required to look for a "likelihood of confusion",[79] the assessment of which involves a wide range of factors that need to be considered holistically. The focus of these tests is on the degree of aural, visual and/or conceptual similarity between the marks and the proximity of the parties' specified goods and/or services, judged from the perspective of the "average" or "notional" consumer. But the tests require much more than an assessment of mere "similarity", guided by prior comparisons and findings. Whether potential similarities are such that they would give rise to a likelihood of confusion involves decision-makers taking into account a variety of further factors. These include: the nature and knowledge base of the potential consumers of the goods or services; the levels of attention they would be expected to pay in making purchasing decisions; the fact that they are likely to have an imperfect recollection of the earlier mark (and thus misremember certain details, or give particular weight to other details); general marketplace factors such as the circumstances in which the goods are commonly sold; whether the marks contain descriptive elements or elements that are "common" to the trade; and, in some jurisdictions, the reputation of the marks or elements within them.[80] The relevance of and weight to be given to each of these factors depend entirely on the marks and specifications that are before the decision-maker.

---

76 In the EU, see European Trade Mark and Design Network, *Common Communication on the Common Practice of Distinctiveness – Figurative Marks Containing Descriptive/Non-distinctive Words* (2015).

77 In the EU, see *Henkel KGaA* Case C-218/01, [2004] ECR I-1725 (ECJ) at [49].

78 This is the position in Australia: see TMA (Cth), s 41(3)–(4).

79 EUTMR, Art 8(1)(*b*); TMA (UK), s 5(2); TMA (S'pore), s 8(2). See also TMA (Cth), ss 10, 44(1)–(2) (on "deceptively similar" marks).

80 On the relevant tests in the EU/UK, Australia and Singapore, see Ilanah Fhima & Dev S Gangjee, *The Confusion Test in European Trade Mark Law* (Oxford: Oxford University Press, 2019); Bently *et al*, *Intellectual Property Law*, *supra* note 74 at 1045–1066; Burrell & Handler, *Australian Trade Mark Law*, *supra* note 43 at [6.8]–[6.34]; Ng-Loy, *Law of Intellectual Property in Singapore*, *supra* note 74 at

In addition, it needs to be appreciated that the "average consumer" has both an empirical and a normative dimension.[81] Many of the factors identified in the previous paragraph are not designed exclusively to reflect how actual consumers are likely to respond to two given marks in a marketplace. Rather, they are also designed to achieve certain broader policy goals, such as to ensure free and fair competition within markets. For example, the idea that the "average consumer" would give limited weight to descriptive components within marks, such that marks that contain common descriptive material are routinely allowed to co-exist,[82] would seem to be intended to ensure that a first-mover within a market is not able to secure a *de facto* monopoly over descriptive terms (irrespective of how actual consumers might in fact respond to such marks). This is reinforced by the sceptical approach taken in some jurisdictions, most notably the UK and Australia, towards both survey evidence and expert opinion evidence that might be used to show confusion among actual consumers.[83]

Some commentators have expressed faith in the idea that the various factors that decision-makers are required to take into account in assessing for likelihood of confusion, which can be gleaned from relevant case law, can be codified into clear rules, which can then be used by AI tools to produce consistent outcomes on confusing similarity.[84] However, such a view downplays the complexity of the legal tests involved, in particular how the factors outlined above interoperate within individual assessments depending on specific circumstances, as well as the policy-based nature of the tests. It might well be possible to ensure that an AI tool can be trained to identify "similarities" between marks and between goods and services. But beyond this, rules would not easily be able to be formulated on matters such as: the relationship between the degrees of visual, aural and conceptual similarities between the marks in any given case; how lower levels of similarity between goods/services can potentially be overcome with a higher degree of similarity between the marks, and *vice versa*; when a particular marketplace factor (for example, noise levels in the locations where the specified goods might ordinarily be sold) might take on greater prominence; what weight to give common or descriptive material within marks; or whether the reputation of a mark counts towards or against confusion.[85] These are matters that appropriately trained decision-makers ought to be well-placed to take into account and weigh up, based on the specific marks and specifications before them, and guided by legal standards. Such factors are, however, likely to be either absent from or play an uncertain role in any AI-generated assessment of

[23.1.14]–[23.1.17], [23.5.32]–[23.5.35], and see further David Tan & Benjamin Foo, "The Extraneous Factors Rule in Trademark Law: Avoiding Confusion or Simply Confusing?" [2016] Sing JLS 118.

81 See the literature cited at *supra* note 6.

82 See, *eg*, Burrell & Handler, *Australian Trade Mark Law*, *supra* note 43 at [6.21]–[6.22].

83 See generally Weatherall, "The Consumer as the Empirical Measure", *supra* note 6 at 61–66, 82–83.

84 Lim, "Computational Trademark Infringement", *supra* note 5 at 267–272.

85 See the literature cited at *supra* note 80. This will be a problem in other jurisdictions, such as the USA which applies a multi-factor assessment of "likelihood of confusion" and obliges decision-makers to engage in a complex exercise weighing these factors: see, *eg*, *Polaroid Corp v Polarad Elecs Corp* 287 F 2d 492 at 495 (2d Cir, 1961); *AMF Inc v Sleekcraft Boats* 599 F 2d 341 at 348–349 (9th Cir, 1979). For critique, see Barton Beebe, "An Empirical Study of the Multifactor Tests for Trademark Infringement" (2006) 94(6) Cal L Rev 1581.

"likelihood of confusion" – for example, it is especially difficult to see how an AI tool could be expected to deal with effective parodies, where despite what are likely to be substantial similarities between the marks and the goods, the parodic intention behind the later mark means that there would be no likelihood of confusion.[86]

A further complicating issue – and one that is common to assessments of distinctiveness and conflicts between marks – relates to how a human decision-maker is likely to engage with prior decisions in their reasoning processes and in reaching their conclusions. In any system with a doctrine of precedent, some decisions will carry more authority than others, and not all prior decisions will be considered equally by a subsequent decision-maker. Most obviously, decisions of a court on a particular issue are likely to be seen as more authoritative than decisions of an examiner or hearing officer – such cases might be influential not only because they contain specific statements of legal principle, but also because subsequent decision-makers might wish to ensure that they reach outcomes broadly consistent with the outcome reached in that earlier, authoritative case. But this is not invariably so: offices might well recognise that particular outcomes reached in court cases are hard to reconcile with long-standing practice, and will choose to persist with the latter in their decision-making.[87] In addition, certain decisions might be seen as more authoritative or valuable simply because of the frequency with which they are cited, or because they were singled out in another decision (for example, as a useful illustration of a particular point), or even on the basis of the perceived expertise of the decision-maker. Human decision-makers will therefore place weight on or de-emphasise prior decisions for a range of complex and subtle reasons in reaching their own decisions on a particular outcome.[88] In contrast, the "weight" of a decision relative to another, understood in this way, is not something that can easily be taken into account in any AI model designed to provide "assessments" based on past decisions on that legal issue.

What all of the above suggests, and building on a point first made by Gangjee, is that we need to be much more open about the limits of what AI tools can produce in the trade mark registration process. An AI tool might well be able to generate an output that is superficially consistent with others in the dataset on which it was trained, and which might well be the same as a decision reached by a human decision-maker after considerable research and mental effort – for example, it may well be able to produce a plausible finding that two marks are confusingly similar. Yet such a finding will inevitably involve a much more "attenuated" or "shrunken" assessment of the legal issue in question than what a human decision-maker with

---

86 See, *eg*, *Louis Vuitton Malletier SA v Haute Diggity Dog LLC* 507 F 3d 252 (4th Cir, 2007); *VIP Products LLC v Jack Daniel's Properties Inc* 2025 WL 275909 at 19–22 (DC Ariz, 2025). My thanks go to David Tan for this point.

87 On this phenomenon, see Burrell & Handler, *Australian Trade Mark Law*, *supra* note 43 at [4.13] (on the distinctiveness of foreign language marks), [4.18] (on the distinctiveness of surnames), [6.7], [7.27], [9.5] (on the interpretation of "substantial identity" and related doctrines).

88 See generally Neil Duxbury, *The Nature and Authority of Precedent* (Cambridge: Cambridge University Press, 2008), especially at ch 3 (on the various ways precedents can be understood to capture "reasons", including as to the weight decision-makers may choose to give to earlier decisions) and ch 4 (on the various ways in which decision-makers might seek to distinguish or overrule earlier decisions).

adequate legal training applying legal reasoning is likely to be able to produce.[89] The AI tool might have been able to take into account *some* of the factors that go towards the legal test in question – for example, the abstract visual, aural and conceptual similarities between the marks, or the relationship between the mark and the specified goods and services. But it will not have been able to take into account the full range of factors and policy considerations that the law requires be considered, and that we would expect a human decision-maker to consider. Equally, it will not have been able to determine the weight that should be given to prior decisions in reaching a conclusion. This seriously calls into question the value of the outputs that such tools generate – that is, there is a significant risk that AI-produced outcomes will meaningfully differ from those that would be reached through an application of the law.

V. Training Data Limitations: Problems with "Bad Law"

In addition to the problems identified in the previous two Parts, a further broad issue of concern relates to the quality and limitations of the data used to train AI models that are designed to produce "assessments" as to the registrability of marks. As indicated in Part III above, it is difficult to find publicly available information from trade mark offices about the data they have been or are using to train their AI models. It is, however, a fair assumption that the most likely sources of such data will be decisions of examiners, as well as other parties such as hearing officers within opposition divisions.[90] This data may well be supplemented by other sources of information such as internal training materials, public-facing manuals and court decisions.

It might be tempting to think that the potential issues relating to training data are largely practical ones. That is, it might be thought that the key challenges involve producing a sufficient quantity of data that "reflects the legal concepts, tests, and case law of the relevant legal system", which would involve identifying relevant legal material, and structuring that material by reference to relevant points of law.[91] These practical concerns about assembling and organising sufficiently large and representative datasets are clearly important. But it should not be assumed that if such concerns are addressed, the resulting datasets are likely to be relatively stable and provide an accurate reflection of the state of the law, and thus a sound basis for training an AI model (subject to the concerns outlined in the previous two Parts). I wish to suggest that the limitations involved in using what might be termed "legal decisions" as training data for AI models go much deeper than this.

To start, while datasets consisting of prior legal decisions might well provide useful factual information about outcomes (that is, that particular marks were or were not held to be inherently distinctive, or that two marks were found to be or not

---

89 Gangjee, "Eye, Robot", *supra* note 7 at 187.

90 See Cabay, Vandamme & Debeir, "Looking Through", *supra* note 40 at 10 (drawing this inference from their study of the EUIPO's search tool).

91 Moerland & Freitas, "Artificial Intelligence and Trademark Assessment", *supra* note 7 at 281. See also Gangjee, "Quotidian", *supra* note 7 at 338.

to be in conflict), it cannot be assumed that all of the decisions contained in the datasets are in fact *correct*. Some of the decisions within training datasets will involve misapplications of the law. Datasets, especially those consisting of examiners' decisions that have not been the subject of internal or external review, will inevitably include marks that should not have been found to be inherently distinctive, or marks that were blocked from registration where a conflict should not have been found. In addition, and of even greater importance, a set of legal decisions must be seen as *fluid*. Both the law and administrative practice can change over time, which in turn might impact on the "correctness" of former decisions, in ways that may not be obvious. The use of any body of legal decisions as data to train AI tools therefore contains significant limitations. As Robert Burrell and I have argued elsewhere, where there are problems with the bodies of law that are used as AI training data, this leads to a risk of trade mark law developing its own variation of the "garbage-in, garbage-out" phenomenon that will limit the utility of the resulting AI tool and what it can generate.[92] These problems can be demonstrated by reference to a range of scenarios, which I have chosen to explore by reference to Australian trade mark law, although the problems identified are likely to arise in any jurisdiction.

The most obvious problem would appear to be where the dataset involves a group of decisions where the law has consistently been misapplied. For example, over the last 25 years Australian hearing officers have routinely made overbroad findings as to when two marks are "substantially identical" with each other in particular contexts, contrary to long-standing authority.[93] These contexts have involved comparisons where both marks contain the same essential feature, but where one contains additional descriptive or non-distinctive material resulting in striking visual dissimilarities between the marks. Another, more subtle problem is where the dataset involves decision-makers adopting a normatively undesirable position in applying an established legal test. For example, it may be the case that examiners, in assessing whether two marks are sufficiently similar so as to generate a likelihood of confusion, are routinely giving excessive weight to whether one mark is contained within another.[94] The concern in both scenarios is that "biases in the training data will be unintentionally embedded in the training model[s] developed from that data"[95] – the outputs might well be consistent with the results reached within the training data, but they do little more than reproduce previous errors. The suggestion that such biases can be washed out by "flooding the system with voluminous data"[96] seems to rest on the faulty premise that additional data (even if it exists) will somehow be "unbiased", rather than containing and thus entrenching existing biases, or even adding new ones.

More challenging problems arise when there is significant inconsistency within the dataset. Given the subjective and multi-factorial nature of both the distinctiveness inquiry and the inquiry into whether two marks are in conflict, some degree

---

92 Burrell & Handler, "Who Reads?", *supra* note 15 at 274.

93 See Michael Handler, "Trade Mark Law's Identity Crisis (Pt 1)" (2021) 44(1) UNSWLJ 394 at 415–419.

94 This is an especially complex issue, as recently considered in Australia in *Paige LLC v Sage and Paige Collective Pty Ltd* (2025) 187 IPR 430 (FCA).

95 Gangjee, "Quotidian", *supra* note 7 at 338.

96 Lim, "Computational Trademark Infringement", *supra* note 5 at 281.

of inconsistency between decisions of examiners and hearing officers is clearly to be expected. However, there may be situations where the legal test in question is so uncertain that it has generated a largely incoherent body of decisions. An example from Australia is the set of nearly irreconcilable opposition-level decisions that has developed over the last three decades as to when the retailing of specified goods will be held to be a service that is "closely related" to those goods,[97] in the absence of clear court authority as to how this test is to be applied.[98]

The subtlest, but also perhaps most significant, problem with relying on datasets consisting of legal decisions is when the law subsequently changes. The issue here is that legal decisions, in addition to having a particular "weight" (as discussed in Part IV), also have a "valency", in that they have the capacity to bind to earlier decisions and render them inert.[99] Australia provides two useful case studies of this phenomenon.

First, before March 2023 a decision-maker in determining whether two marks were "deceptively similar" for the purposes of assessing conflicts between marks was able to take into account whether the earlier mark, or an element of that mark, had a reputation. Taking this factor into account meant that the notional consumer would be considered to be less likely to have an imperfect recollection of that earlier mark, thus lessening the risk of confusion and of the marks in question being deceptively similar.[100] However, in its March 2023 decision in *Self Care IP Holdings Pty Ltd v Allergan Australia Pty Ltd*, the High Court of Australia held that in assessing "deceptive similarity", albeit in the context of an infringement action, it was *impermissible* to take into account the reputation of the earlier, registered mark.[101] It was quickly recognised that such an approach would apply equally in assessing for deceptive similarity in the registration context.[102]

Second, before March 2023 Australian examiners, in interpreting what certain words used in specifications of goods and services meant, could refer to the Nice class in which those goods or services had been classified for interpretive guidance. For example, it is likely that they would have read a claim to "apparatus (hand tools)" in Class 8 as not extending to electrically operated hand tools, since the latter were classified in Class 7.[103] This interpretative approach had the effect of restricting the scope of specifications. This situation changed following the decisions of the Full Court of the Federal Court in *Energy Beverages LLC v Cantarella Bros Pty Ltd*[104] in March 2023 and *Killer Queen LLC v Taylor*[105] in November 2024. The effect of those decisions is that specifications are now to be interpreted

---

97 See TMA (Cth), s 44(1)–(2).

98 See Sgourakis, *Australian Trade Mark Practice*, *supra* note 36 at 752–765; Burrell & Handler, *Australian Trade Mark Law*, *supra* note 43 at [6.34].

99 My thanks go to Robert Burrell for this concept.

100 Burrell & Handler, *Australian Trade Mark Law*, *supra* note 43 at [6.27].

101 *Self Care IP Holdings Pty Ltd v Allergan Australia Pty Ltd* (2023) 277 CLR 186 (HCA) at [36]–[51] [*Self Care*].

102 *Volvo Trademark Holding AG v Vovo Corporate Co Ltd* [2023] ATMO 48 at [28].

103 *Tornado Trade Mark* [1979] RPC 155 (UK Reg) at 159.

104 *Energy Beverages LLC v Cantarella Bros Pty Ltd* (2023) 407 ALR 473 (FCAFC) [*Energy Beverages*].

105 *Killer Queen LLC v Taylor* (2024) 306 FCR 199 (FCAFC) [*Killer Queen*]. The High Court did not consider this aspect of the Full Court's decision on appeal in *Taylor v Killer Queen LLC* (2026) 100 ALJR 335.

from the perspective of an average trader in the goods in question, with words being given their ordinary dictionary definitions, and not from the perspective of an expert with knowledge of the classification system.[106] This means, for example, that "non-alcoholic beverages" classified in Class 32 will include coffee, even though "coffee" is properly classified in Class 30.[107]

Self Care and *Energy Beverages*/*Killer Queen* significantly changed Australian law and practice. The effect of these cases is that there must now be a large subset of Australian decisions in which decision-makers found that two marks were not deceptively similar, or that two sets of goods or services were dissimilar to each other, on the basis of principles that have now been overturned. That is, there will be a body of earlier decisions that must now be regarded as either wrong or at least of dubious value. However, if these decisions are contained in training datasets, there would be no easy way of identifying, let alone extricating, them. Merely adding new decisions to the dataset, and even assigning these new decisions a higher "weighting", would not ensure that an AI tool would make "deceptive similarity" or "similarity of goods or services" assessments consistent with the new approach mandated by *Self Care* and *Energy Beverages*/*Killer Queen*.[108]

The above problems not only show some of the limitations of the AI tools currently being developed for use within offices, but also call into question the idea that there are technological fixes for the problems identified. They also show that it is more important than ever that we need to look over our shoulders and scrutinise our existing bodies of law, and assess whether these contain errors, or are internally inconsistent, or reflect normatively undesirable principles. Since these bodies of law are likely to continue to form the basis of training datasets, the automated decision-making tools on which governments increasingly wish to rely will serve only to enshrine unsatisfactory practices if the problems with the law contained within these datasets are not identified and addressed.[109]

VI. Concluding Thoughts: Avoiding Automation Bias, and Ensuring High-Quality Human Decision-Making

Projects to develop AI tools to make automated assessments of matters going to the registrability of trade marks, such as distinctiveness and conflicts between marks, are not being undertaken with the explicit goal of removing human decision-making from the registration process. Rather, they are premised on the idea that

---

106 *Energy Beverages*, *supra* note 104 at [132]; *Killer Queen*, *supra* note 105 at [57]. See Burrell & Handler, "Who Reads?", *supra* note 15 at 278–280.

107 *Energy Beverages*, *supra* note 104 at [132].

108 See Burrell & Handler, "Who Reads?", *supra* note 15 at 299. Further, even though it might be possible to identify particular "overturned" decisions (for example, it is clear that *Self Care* overturned *Registrar of Trade Marks v Woolworths Ltd* (1999) 93 FCR 365 (FCAFC) [*Woolworths*] at [61] on the "reputation being relevant to deceptive similarity" issue), there would be problems removing such decisions from training datasets. Those decisions might still stand for other valid propositions, and they might well have been decided in the same way in the absence of the "overturned" issue (*Woolworths* being a good example on both counts).

109 See Burrell & Handler, "Who Reads?", *supra* note 15 at 299.

these tools can increase efficiency in the examination process, which will manifest itself when those assessments are (broadly) equivalent to those that would be made by the (slower) human decision-maker. For the reasons outlined in this article, this idea should be treated with a high degree of caution. Not only is this "equivalence" likely to be some way off, but the AI tools being developed have a range of significant limitations. Some of these limitations – such as a lack of transparency as to, and problems with, the data on which such tools are being trained, and the inevitable gap between what the tools can produce and what trade mark laws in fact demand – need to be taken particularly seriously. But by way of conclusion, it is worth briefly addressing an objection to the above analysis. That is, how much might the limitations identified in this article really matter, if all that trade mark offices are ever going to do is provide these AI tools to examiners to produce non-binding "pre-assessments" on limited matters of registrability, while still requiring those examiners to make the ultimate decisions?[110] Will these tools, despite their known limits, not still be helpful resources that generate some efficiency gains, and should be seen as a net positive as a result?

Even these seemingly benign suggestions should not be accepted so readily. It might be the case that some existing AI tools can be trusted to provide accurate information, such as an indication that there is a high probability that a mark is inherently distinctive, or that it conflicts with a near-identical earlier mark for near-identical goods or services. But such an assessment will lead to negligible gains in efficiency: a human decision-maker could very likely reach the same decision almost instantly, or with minimal searching, without the input of AI. The more complex issue relates to what a human decision-maker is meant to do when provided with less clear-cut AI-generated outputs. Here, there is a danger in seeing this information as simply one more "reference point" for examiners to consider in their decision-making. This is not simply because such information may well contain limitations or inaccuracies, as outlined in Parts III to V above. It is also because, in borderline cases in particular, there is a very real risk that any pre-assessment might result in "automation bias" in the human decision-maker. This is a well-recognised phenomenon, where a decision-maker uses "automated cues as a heuristic replacement for vigilant information seeking and processing".[111] That is, when presented with an automatically-generated output, the decision-maker is likely to engage in a degree of "cognitive offloading", placing an undue level of trust in that output and not exercising the level of critical engagement with the task at hand.[112] Automation bias is especially likely where the human decision-maker considers the issue in question to be complex, is time-poor in having to deal with large case loads

---

[110] The suggestion that AI tools will be used only to assist human decision-makers, who bear the ultimate responsibility for their decisions, is reflected in offices' current public statements: see, *eg*, IP Australia, *AI and ADM Transparency Statement*, *supra* note 43 at 2 ("decisions that are likely to significantly impact our customers, including decisions to grant IP rights, are always made by our skilled people"); EUIPO, "Real-time Alerts", *supra* note 21.

[111] Kathleen L Mosier & Linda J Skitka, "Human Decision Makers and Automated Decision Aids: Made for Each Other?" in Raja Parasuraman & Mustapha Mouloua, eds. *Automation and Human Performance: Theory and Applications* (Boca Raton: CRC Press, 1996) 201 at 205.

[112] For recent consideration see Michael Gerlich, "AI Tools in Society: Impacts on Cognitive Offloading and the Future of Critical Thinking" (2025) 15(1) Societies 6: 1–28.

and performance targets, and implicitly trusts the automated system.[113] This would seem to reflect the situation of many trade mark examiners, and casts real doubt on the notion that an AI-generated pre-assessment of registrability would simply be a benign "reference point" for examiners.

If examiners within offices are to be asked to consider AI-generated pre-assessments of registrability – even if these improve substantially in quality and accuracy – it is incumbent on offices to provide additional resources and training to their staff to help them understand exactly what these pre-assessments mean and how they should be interpreted. That is, staff will need to be properly trained as to how the AI tools in question operate, how the pre-assessments were undertaken, what the information contained in them (including probabilities and confidence scores) signifies, and the risks that the information will be underinclusive, overinclusive or inaccurate. It may well be that asking examiners to consider AI-generated pre-assessments, and providing them with sufficient support and training to help them overcome any automation biases, will in fact add new *inefficiencies* into the examination system, suggesting that any such projects should be undertaken cautiously.

As a related and final point, offices will need to weigh up how their resources are best allocated in attempting to ensure the highest quality examination processes – something that goes beyond mere efficiency. While investing in new technologies to assist in the registration process is clearly important, this should not come at the expense of ensuring that examiners, hearing officers and their managers are adequately trained *in the law* on an ongoing basis. It is more important than ever that offices allocate appropriate resources to ensure that these people have the knowledge, skills and capabilities to understand and be able to apply complex bodies of trade mark law, so that they can make the best possible decisions in difficult cases – whether this is with or without the assistance of AI.

[113] Tatiana Kazim & Joe Tomlinson, "Automation Bias and the Principles of Judicial Review" (2023) 28(1) Judicial Rev 9 at 11–13. See generally Daniel J Solove and Hideyuki Matsumi, "AI, Algorithms and Awful Humans" (2024) 92(5) Fordham L Rev 1923 at 1934–1938.