

Risks (and Safeguards) in the Use of AI by IP Professionals¹

John Gray

John Gray IP Limited, UK | www.johngrayip.com

Abstract

Generative AI tools are transforming the working environment for intellectual property professionals, offering real gains in quality, speed, and cost. But they also introduce risks that are new in character, even if the professional rules that govern them are not. This article surveys the principal risks of AI use in IP practice — covering competence, confidentiality, communication, and cost — and examines the guidance issued by bodies including the epi, the ABA, IPREG, CIPA, and CITMA. It discusses the structural limitations of large language models, the confidentiality dangers of free and enterprise AI tools, and implications for attorney-client privilege in some jurisdictions. It also identifies less obvious and/or longer-term risks, including the "truth tax" of verification, the hazards of "feature creep" in AI tools, and the challenge of maintaining professional skills in an AI-assisted environment. The article concludes with practical guidance for practitioners beginning to adopt AI tools safely and responsibly, while keeping the human professional firmly in the loop.

1. Introduction: Putting Generative AI in Perspective

The emergence of large language models (LLMs) as practical tools did not happen overnight. The underlying technology – artificial neural networks (ANNs) – has been developing for decades, with key theoretical foundations laid as far back as the early 1980s.² What changed was the ready availability of powerful GPU chips, a new ANN architecture (the ‘transformer’³) and of course the accumulation of enormous training datasets. The result, after many decades of false promises and niche applications, is that LLM-based generative AI tools are now available from major technology providers and are accessible to any practitioner with an internet connection.

For IP professionals, this creates both opportunity and responsibility. Used well, generative AI can assist with technical and legal research, prior art searches, portfolio management, drafting support, prosecution strategy, and office administration. The potential benefits are real: improved quality, greater speed, and lower cost. However, these same tools introduce risks that, if not carefully managed, could lead to professional liability, loss of client confidentiality, or harm to the very intellectual property rights practitioners are engaged to protect.

For those not yet immersed in the use of modern AI tools, the author hopes to illuminate some of the hazards, so that a safe path may be found, into this new world.

As IP professionals, we engage with the law on behalf of the diverse creative individuals and companies who are our clients. Modern developments in AI raise

¹ This paper is based on a presentation given at the XLI Jornadas de Estudio Grupo Español AIPPI, 27 February 2026, Madrid.

² J. J. Hopfield, 'Neural networks and physical systems with emergent collective computational abilities', Proceedings of the National Academy of Sciences, 79(8), pp. 2554–2558 (1982).

³ Vaswani et al, 'Attention Is All You Need', <https://arxiv.org/abs/1706.03762>

fundamental ethical and practical concerns for our clients in their own work, and for society as a whole. While those wider concerns are beyond the scope of this paper, engaging with AI tools in our narrow field of work can lead us to an improved appreciation of the wider issues as well.

2. The Promise: Generative AI with the Human in the Loop

The most productive model for AI use in IP practice places the skilled professional firmly at the centre of the workflow. So-called ‘Generative AI’ tools such as LLMs have a wide range of applications in IP practice, including technical and legal research, prior art searching, landscaping and data analysis, prosecution strategy, portfolio management, minute-taking, drafting assistance, and quality checking. Specialised tools are now available for many of these tasks, alongside general-purpose platforms. However, it is vital that a qualified human reviews, verifies, and takes responsibility for the output. This “human in the loop” approach is directly reflected in the “human-centred” approach to AI adopted by the European Patent Office.⁴

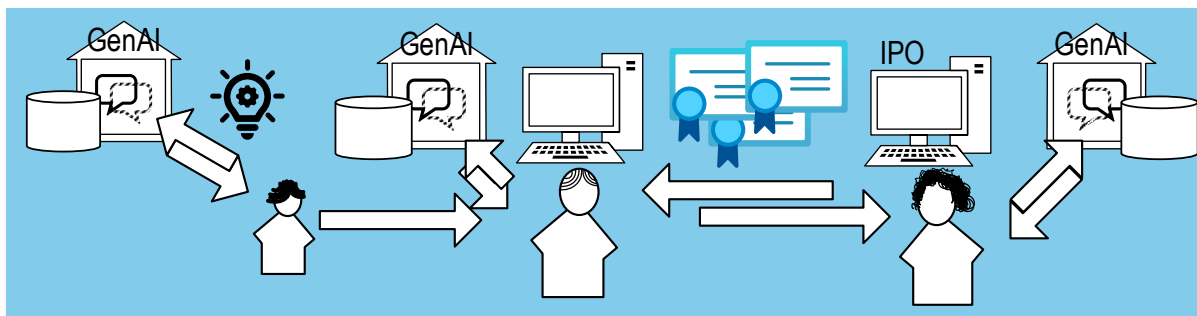


Figure 1: Generative AI + “human in the loop”, illustrating the multiple AI-assisted stages in a patent workflow from inventor to attorney to IPO, with the humans attentive and in control.

When this model is followed, the potential gains are significant. Quality can improve because AI can rapidly process large amounts of information and suggest options the practitioner might not have considered. Speed increases because routine drafting and research tasks take less time. Costs may fall as a result. The danger arises when the human steps out of the loop, whether through overconfidence in the AI, time pressure, or simple unfamiliarity with the tool’s limitations.

3. Limitations of Large Language Models and Generative AI

Before examining the specific professional risks, it is worth understanding when and why those risks arise in the first place. While risks can arise for all forms of machine learning and artificial intelligence, the LLM-based tools are proving uniquely especially seductive and apparently applicable to the everyday tasks of the ‘generative’ IP professional. LLMs are of course trained on vast databases of diverse knowledge, but that does not make them reliable for every purpose. Several limitations are particularly relevant to IP practice.

First, there is the question of knowledge coverage and currency. The training data may not adequately cover a practitioner’s technical specialty, a particular national legal system. It may not include recent developments in case law or office practice. The

⁴EPO, 'Human-centred approach to AI in the patent system' (2023). Available at: <https://www.epo.org>

model reflects the state of knowledge at the time of its training, and the sources that were available and selected by its ‘teachers’. Thus the model may be biased towards certain branches of technical or legal knowledge, certain languages, or certain legal traditions.

Second, there is the question of how much thought the user puts into the prompt and what comes out as a result. If the practitioner does not provide sufficient context, the output will reflect that gap. And the output may, in some circumstances, incorporate material derived from copyrighted or confidential sources in ways that are difficult to detect.

Third, and more fundamentally, the process by which an LLM generates output is stochastic rather than deterministic. The model predicts the most statistically likely response to a given prompt: it does not reason from first principles, and it does not “know” whether its output is correct. This can have critical consequences for professional use. Most models will not simply say “I don’t know”. They will produce an answer, in the most confident and convincing style, even when that answer is fabricated. This phenomenon of “hallucination” is not an occasional glitch, but a structural feature of how these models work. References might be invented, genuine references might be misquoted, and technical facts might be stated with complete apparent authority, even when they are entirely wrong.

For IP professionals, who work with technically complex subject matter, rely on precise legal language, and operate under strict obligations of accuracy, these limitations are not minor inconveniences. They define the conditions under which AI tools can and cannot be trusted for use in our work.

4. Professional Conduct Obligations: The Rules Are Not New

Fortunately, various professional bodies and authorities around the world are issuing guidance on AI use, and the common starting point for all is that existing professional obligations apply fully to AI-assisted work.⁵⁶⁷⁸⁹ That is to say, the rules of professional conduct are not new. Only the temptations and risks are new. Purely for the sake of convenience, we will discuss the most immediate considerations with reference to four ‘C’s: Competence, Confidentiality, Communication and Cost. Certain ‘hidden’ or long-term risks are discussed in Section 5.

4.1 Competence

Professionals are obliged to act competently. The *epi* Guidelines on the Use of Generative AI make clear that “members remain at all times responsible for their

⁵*epi* Guidelines: Use of Generative AI in the Work of Patent Attorneys (*epi*, 2024). Available at: <https://www.epi.org/resources/generative-ai-guidelines/>

⁶American Bar Association, Formal Opinion 512: Generative Artificial Intelligence Tools (2024). Available at: <https://www.americanbar.org/>

⁷IPREG, Interim Guidance on the Use of AI by Patent and Trade Mark Attorneys (2024). Available at: <https://www.ipreg.org.uk/guidance>

⁸CIPA, AI Tools and Patents - Risks and Safeguards for Inventors and SMEs” (2025). Available as PDF download at: <https://www.cipa.org.uk/news/ai-guidance-for-smes-and-creators/>

⁹CITMA, Choosing Your AI Tools & Vendors: Top 5 Tips for Trust and Partnership Assessment (2024). Available at: <https://www.citma.org.uk>

professional work and cannot cite the use of generative AI as any excuse for errors or omissions”.¹⁰ Members “must check any work product produced using generative AI for errors and omissions”. The American Bar Association’s Formal Opinion 512 reaches the same conclusion, placing competence at the head of its list of professional obligations engaged by AI use.¹¹

“Don’t trust: verify.”

In practice, this means that the practitioner must verify AI output rather than simply trust it. When an AI cites a document, for example, the practitioner should confirm that the document exists and actually says what the AI claims.

This raises a question that practitioners will increasingly confront: if everything the AI produces must be checked, where is the efficiency gain? The economics inevitably depend on the task, the tool, and the practitioner’s familiarity with both. As one write puts it, a “truth tax” must be factored into any honest assessment of AI’s value, representing the hidden cost of checking AI output¹².

Hopefully it will be found that verification is generally faster than original drafting or research. Lessons can be learned to improve prompting in future, and the AI output, even when imperfect, often provides a useful starting point. The AI itself can also help with the verification. For example, once can ask the AI to identify the specific passage in a prior art document where a particular feature is described, or one can ask the same question to a second AI system as a cross-check.

4.2 Confidentiality

Confidentiality is perhaps the area of greatest practical risk for IP professionals, and it requires careful attention to the type of AI tool being used.

Free or consumer-grade AI tools typically operate on a business model in which the user’s inputs may be used to train or improve the model. This creates a direct conflict with the practitioner’s duty to protect a client’s confidential information. The risks are multiple and serious. Confidential technical information fed into a free AI tool may cease to be confidential in any meaningful sense. For patent matters, any non-confidential disclosure of an

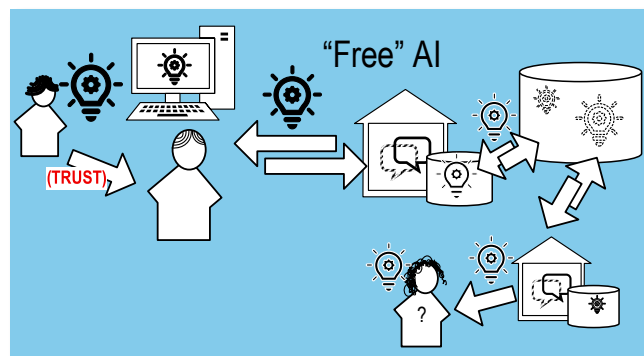


Figure 2 - Beware – ‘FREE’ AI may train on your data (or worse, your client’s data).

invention before the patent filing theoretically destroys novelty, the same as if you told it to a stranger in a bar. For trade secrets, leakage may be irrecoverable. And, as confirmed by recent US litigation, using a consumer-grade AI platform whose terms of

¹⁰*epi* Guidelines, above n 5, at §4.

¹¹ABA Formal Opinion 512, above n 6.

¹² Jeffrey Fazio, ‘The “Truth Tax”: 4 Hard Truths About AI in High-Stakes Fields’, Jan 2026, <https://substack.com/home/post/p-183461819>

service disclaim any expectation of privacy may result in the loss of attorney-client privilege over communications containing that information.¹³

The *epi* Guidelines advise that “if there is doubt that confidentiality will be maintained to a level appropriate to the prevailing context, the AI model in question should not be used”.¹⁴ IPREG puts it more directly: “practitioners have a fundamental duty to safeguard their client’s sensitive information, and must understand whether and how information provided when using an AI product will be stored and secured”.¹⁵

Different clients and different matters may warrant different tools and different settings. Some types of task may not involve confidential information at all. After all, patent applications and their associated documentation are highly confidential up to a point, but then become published and accessible for all. Publicly listed prior art and examiner objections are not sensitive (although your advice to the client about those public things may still be).

Enterprise-grade cloud solutions offer a higher level of protection. Reputable providers typically commit contractually not to train their models on customer data, and confidentiality is maintained to everyday commercial standards. This is sufficient for many purposes, and patent novelty and privilege can generally be preserved. However, some clients will have particular requirements for data security, and the practitioner may still be advised to investigate where data is stored geographically, who has access to it, how long it is retained, and what technical security measures are in place.

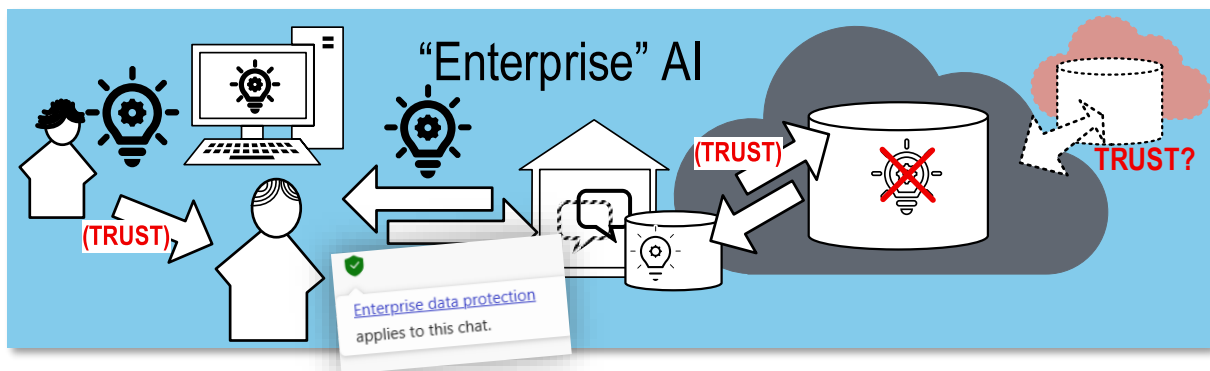


Figure 3 - Benefits and residual risks of “enterprise” cloud-based AI deployments

In case the cloud solution cannot satisfy these requirements, a “sovereign” or locally hosted AI deployment will be required, in which the model runs entirely on the practitioner’s own infrastructure (or the client’s) and data never leaves the premises. This provides the highest assurance of confidentiality and allows the practitioner to custom-train models on their own data. The practical limitation is cost: running a

¹³United States v. Heppner, No. 25 Cr. 503 (JSR) (S.D.N.Y., ruling 10 February 2026, written opinion 17 February 2026, Rakoff J). For commentary see: Harvard Law Review Blog, 'United States v. Heppner' (March 2026), <https://harvardlawreview.org/blog/2026/03/united-states-v-heppner/>; Venable LLP, 'AI, Privilege, and the Heppner Ruling' (February 2026), <https://www.venable.com/insights/publications/2026/02/ai-privilege-and-the-heppner-ruling-what-the-court/>; Inside Privacy, 'AI and Legal Privilege: Key Takeaways from US v. Heppner' (February 2026), <https://www.insideprivacy.com/artificial-intelligence/ai-and-legal-privilege-key-takeaways-from-us-v-heppner/>

¹⁴*epi* Guidelines, above n 5, at §5.

¹⁵IPREG Interim Guidance, above n 7.

capable LLM locally requires substantial computing resources and expertise that may be beyond the means of most small and medium-sized practices.

The Privilege Question: United States v. Heppner

The confidentiality risk is not merely theoretical, as the recent case of *United States v. Heppner*¹⁶ illustrates. Prior to instructing lawyers, the defendant had used a free AI chat bot to prepare defence strategy documents and later shared them with his lawyers. The court held that the communications were not confidential, because the platform's terms of service did not support any reasonable expectation of privacy; that the AI tool was not an attorney; and that the documents were not prepared at counsel's direction. Judge Rakoff of the Southern District of New York ruled that, for all these reasons, the documents generated were not protected by attorney-client privilege. The outcome might have been different had the defendant used a tool with appropriate confidentiality protections, and had done so at counsel's direction. Now, this decision is subject to some criticism in the *Harvard Law Review* article, cited in the footnote, but it nevertheless highlights the importance of selecting the correct tool and of keeping the qualified human professional in the loop: you can hardly trust the AI to tell you whether it's OK to trust the AI with your confidential data.

4.4 Communication and Transparency

The question of whether to disclose one's use of AI to clients and to intellectual property offices is addressed in the *epi* Guidelines. Members are required to establish their clients' wishes regarding the use of generative AI in advance, i.e. before the tool is deployed on a particular matter.¹⁷ Client preferences may differ, and the practitioner may need to operate different tools or settings depending on the instructions received from each client.

On disclosure to the EPO and the Unified Patent Court, the *epi* Guidelines take a measured position: members are not currently required to state that generative AI has been used in the production of work, unless the relevant procedural rules require this. This position may evolve as offices and courts develop their own policies, and practitioners should monitor developments. An individual professional or firm may have to comply in different scenarios with the obligations from multiple professional bodies and authorities. It is to be hoped that harmonisation is generally achieved, and that different regulatory requirements do not come into actual conflict.

4.5 Cost (Fees and Billing)

Both the ABA and IPREG guidance addresses the impact of AI use on billing.¹⁸ There is no general obligation to pass on cost savings achieved through AI use to the client, but practitioners must maintain transparency about billing and act in the client's best interests. The ABA's guidance draws a practical distinction: a lawyer may charge for the time spent inputting information into an AI tool and reviewing the output, but

¹⁷*epi* Guidelines, above n 1, at §7.

¹⁸IPREG Interim Guidance, above n 7; ABA Formal Opinion 512, above n 6.

ordinarily cannot charge a client for the time spent learning to use the tool in the first place¹⁹.

For firms used to applying time-based billing models, there is an immediate conundrum: if an AI tool allows a fee earner to complete a given task in less time, but it costs additional money to set up and use the AI tool (requiring IT infrastructure and/or subscription to external AI tool providers), does the billing model need adapting to give a fair benefit to both the client and the firm? Is the best solution just to adapt the hourly rate?

5. Hidden and Long-term Risks

Beyond the immediate professional conduct issues, there are longer-term risks that are less visible but potentially more consequential.

5.1 Economic and Commercial Risks

One hidden cost is the “truth tax” already mentioned: the time spent verifying AI output. Whether AI use is genuinely cost-effective depends on the task, the tool, and the quality of the practitioner’s prompting. For some tasks, AI assistance delivers clear net gains today. For others, the checking burden may currently outweigh the drafting benefit, though this balance is likely to shift as tools improve. As already mentioned part of the solution is to refine one’s practice to improve results, and to use the AI itself to perform some of the checking.

A subtler commercial risk concerns clients’ expectations. As AI tools become more widely known, some clients may begin to question the value added by professional expertise. CIPA’s AI guidance for SMEs and creators²⁰ offers a salutary warning:

“The patent system fundamentally relies on human invention and absolute confidentiality before filing, and AI tools, whilst offering efficiency gains, threaten both foundations.”

“A single confidentiality breach or over-reliance on AI-generated content could render years of research and development worthless, and the true cost of AI errors may only emerge during litigation or licensing negotiations years later.”

Evidently, educating ourselves as professionals is only a first part of the task: we must also educate our clients and the wider innovation community. And as professionals we need to be clear in our own minds where we are adding value, and how we can demonstrate this to our paying clients and prospects. Key to adding value, as always, is a solid human relationship and knowledge of the client’s background and commercial interests.

5.2 Legal and Procedural Risks

An AI tool, unless carefully instructed, may generate pages of superfluous/erroneous analysis as to why some right or another is or is not valid or enforceable. Where an AI

¹⁹ ABA Formal Opinion above n 6, provides detailed discussion at p.14 “Lawyers must remember that they may not charge clients for time necessitated by their own inexperience”.

²⁰CIPA, above n 8.

tool has been used in the preparation of a case, the AI's analysis or "opinions" may even in some circumstances become discoverable in litigation. The *Heppner* decision, discussed above, is a concrete illustration of this risk. Clients, particularly from the US, do not want to see such things on file. Similarly, an AI tool used in patent drafting may spontaneously include speculative embodiments, or lists of possible 'advantages'. These may be useful, or they may turn out to be pointless and damaging. The skilled attorney may know when "less is more", but the machine does not.

5.3 Skills, Recruitment and Professional Development

AI adoption raises important questions about how professional skills are acquired, maintained and transmitted. If a newly qualified attorney has relied heavily on AI throughout their training, have they developed the underlying understanding needed to recognise when the AI is wrong? If an experienced attorney is using an AI assistant in place of a human trainee, where will the next generation of experienced and skilled attorneys come from?

None of what has been said can be an argument against AI adoption, but it is an argument for conscious management of how AI is integrated into training and supervision. Perhaps a more positive question is: how can we exploit AI in the training of the next generation of professionals?

5.4 "Feature Creep" – AI arrives uninvited and unnoticed

One of the most underappreciated risks is the progressive integration of AI functionality into conventional tools that practitioners already use, often without any active decision to adopt AI. Microsoft Copilot, Google Gemini, and Meta AI are being embedded into widely used productivity applications. Meeting transcription tools such as Zoom AI Companion are becoming standard features of video conferencing platforms. Each of these integrations may involve the transmission of data – including potentially confidential client data – to external AI systems, on terms that the practitioner has not specifically reviewed.

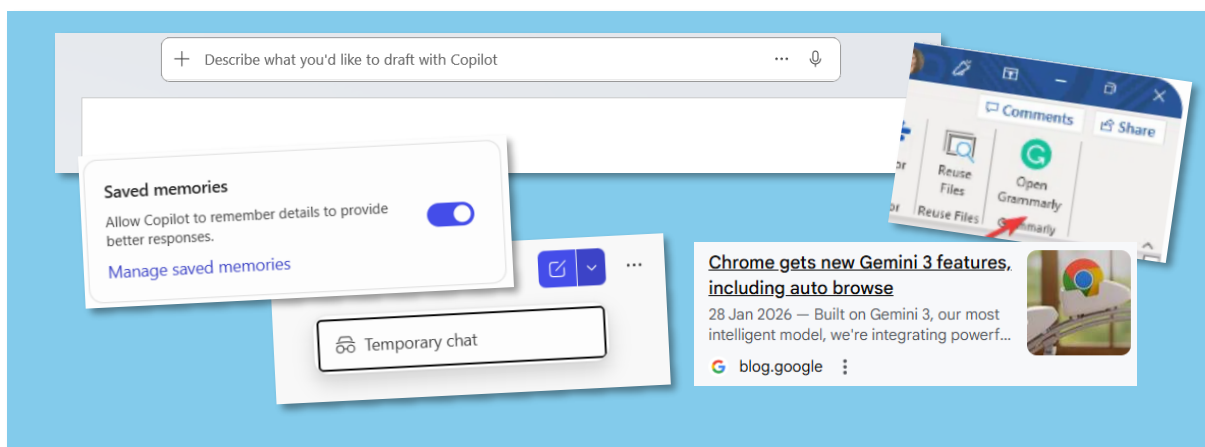


Figure 4 - 'Feature creep' – AI being embedded in familiar tools such as MS Copilot, Gemini, and Zoom AI Companion, potentially processing client data without explicit consent.

Practices should audit the tools they use, identify where AI functionality has been introduced, and make deliberate decisions about whether those features are appropriate for professional use. In some cases, AI features may need to be disabled, or separate accounts maintained for work involving sensitive client information.

One feature appearing recently in AI tools is “memory”: the tool will remember what you have asked it in the past, and may use that as context for today’s query. This may be good for some circumstances, but not for others. Suppose you fed it confidential information of Client A last week, and it thinks that might be something useful to add when it is helping you write something for Client B...?

6. Practical Guidance: How to Begin

6.1 Start Safely, but do experiment and learn

The first step is to develop familiarity with AI tools through low-risk use. Non-confidential tasks, such as general research, drafting on hypothetical scenarios and summarising publicly available documents, provide an opportunity to understand how different tools behave, where they are helpful, and where they are unreliable. A first draft of the present paper was generated by AI, based on the slides shown at the meeting in February. The first draft needed work, but it was an excellent starting point and opportunity for the author to try out a new tool.

Practices should consider how and when different members of the team will develop AI skills, and whether there is a risk that some team members are left behind.

6.2 Choose the Right Tool for the Right Task

Not all AI tools are equal, and not all tasks are equally suited to AI assistance. General-purpose tools are capable across a wide range, while specialist tools – purpose-built for patent drafting, prior art searching, or legal research – may offer greater accuracy in their target domain. Special purpose tools, albeit at extra cost, will have been trained with domain-specific expertise and/or ‘guardrails’, easing the ‘learning curve’ and reducing the risk of users falling into error.

It is also worth remembering that not every automation task requires generative AI: for deterministic, rule-based processes, conventional automation (macros and the like) may be more reliable than a stochastic LLM. Generative AI can itself be a useful tool for designing and implementing such automation, even where it would not be appropriate as the front-line tool for the substantive task.

6.3 Supervise and Challenge the AI (and let it challenge you)

AI tools should be treated as capable but fallible assistants, not as authoritative sources. Practitioners should adopt a habit of challenging AI output: asking the tool to justify its conclusions, to identify the sources it relies upon, and to consider counterarguments. Using a second AI tool to review the output of a first is a practical technique for surfacing errors. Where an AI cites a specific passage in a document, that passage should be verified directly against the source. And even if you prefer to do certain tasks yourself (like a patent attorney keeping tight control of their Claim 1 wording), you can always benefit from asking the AI to review and criticise your work.

Referring again to the writing of this paper as an example, revision of the AI-generated draft was onerous but, by comparing the AI-drafted and human-edited versions, the AI was able to derive some style preferences which can be saved for use in future tasks.

6.4 Assess and Select Vendors Carefully

Practitioners should not rely on general marketing claims but should seek contractual commitments and, where appropriate, independent verification. CITMA's guidance²¹ proposes a systematic framework for the evaluation and introduction of AI tools in an IP firm. They define a "trust equation" built on four elements:

- credibility,
- reliability,
- sector familiarity, and
- a user-focused approach.

To these should be added specific due diligence on data privacy, confidentiality protections, data residency, retention policies, and governance frameworks.

With a current boom in AI offering directed at IP and legal professionals, two things seem certain: not all of these providers and tools are equally suited to your particular needs, and not all of these providers and tools will be in business in a few years' time.

7. Conclusion

Artificial Intelligence, and Generative AI in particular is not a passing trend, and IP professionals cannot afford to ignore it from a commercial or, ultimately, professional perspective. The efficiency gains are real, the tools are improving rapidly, and clients will increasingly expect practitioners to use them. The risks are equally real, but they can be assessed and mitigated within the framework of existing professional obligations.

The practitioners who will navigate this transition most successfully are those who approach AI with the same rigour they apply to their substantive work: learning the tools thoroughly, understanding their limitations, choosing them carefully, and remaining firmly in control of the process and its outcomes.

The core message of the guidance issued by the *epi*, IPREG, CIPA, CITMA, and the ABA is consistent: the professional is always responsible. AI can assist, accelerate, and augment the work of an IP practitioner, but it cannot substitute for professional judgment, and it cannot transfer professional liability. Every output must be checked. Every client's confidentiality preferences must be respected. Every AI tool, in every iteration, must be assessed at some level before use.

The human must remain in the loop not as a 'rubber stamp', but as the professional whose judgment, expertise, and accountability the client has engaged and the system depends upon.

²¹See above, n 9

Further Reading and Resources

The following resources are just a few potential starting points for practitioners wishing to develop their understanding of AI tools and their professional implications:²²

- EPO podcast: Exploring AI in the Patent Profession <https://youtu.be/S-e33ibYFIU>
- *epi* Insight podcasts <https://www.epi-learning.org/course/view.php?id=129> episodes:
 - epi* Guidelines: Use of Generative AI in the Work of Patent Attorneys
 - AI in Patent Practice: A Reality Check
 - AI at the EPO: Learning from the Best
- IPKat articles on AI <https://ipkitten.blogspot.com/search/label/AI>
- Bastian Best Software Patent Attorney blog: <https://bestpatent.eu/>
- IP Lawyer Tools by Martin Schweiger (section “Robot Patent Drafting”) <https://ip-lawyer-tools.com/>
- Russell IP – IP Tool Demo Day: Exploring Nine IP Tools (May 2025) <https://www.russellip.com/ip-tool-demo-day> & other posts <https://www.russellip.com/ai-in-ip-potential-risks-and-rewards/>
- Responsible AI UK (RAI UK) – research on ethics and regulation <https://rai.ac.uk/about-us/>

Author Note

John Gray is a European Patent Attorney and UK Patent Attorney, and director of John Gray IP Limited. Among other roles, he is Chair of *epi*'s Online Communications Committee and a member of the Disciplinary Committee of *epi*.

The views expressed are those of the author alone and do not reflect the official position of the *epi* or any other organisation. Contact: www.johngrayip.com

This article is based on a presentation delivered to AIPPI Grupo Español on February 2026. Thanks to the group for this stimulating invitation.

²²Bastian Best Software Patent Attorney blog, <https://www.bastian.best>; IP Lawyer Tools by Martin Schweiger (section 'Robot Patent Drafting'), <https://www.iplt.io>; Russell IP, 'IP Tool Demo Day: Exploring Nine IP Tools'; EPO Podcast, 'Exploring AI in the Patent Profession', <https://www.epo.org/en/news-events/podcasts>; *epi* Insight Podcast; IPKat articles on AI, <https://ipkitten.blogspot.com>; Responsible AI UK, <https://www.rai.ac.uk>