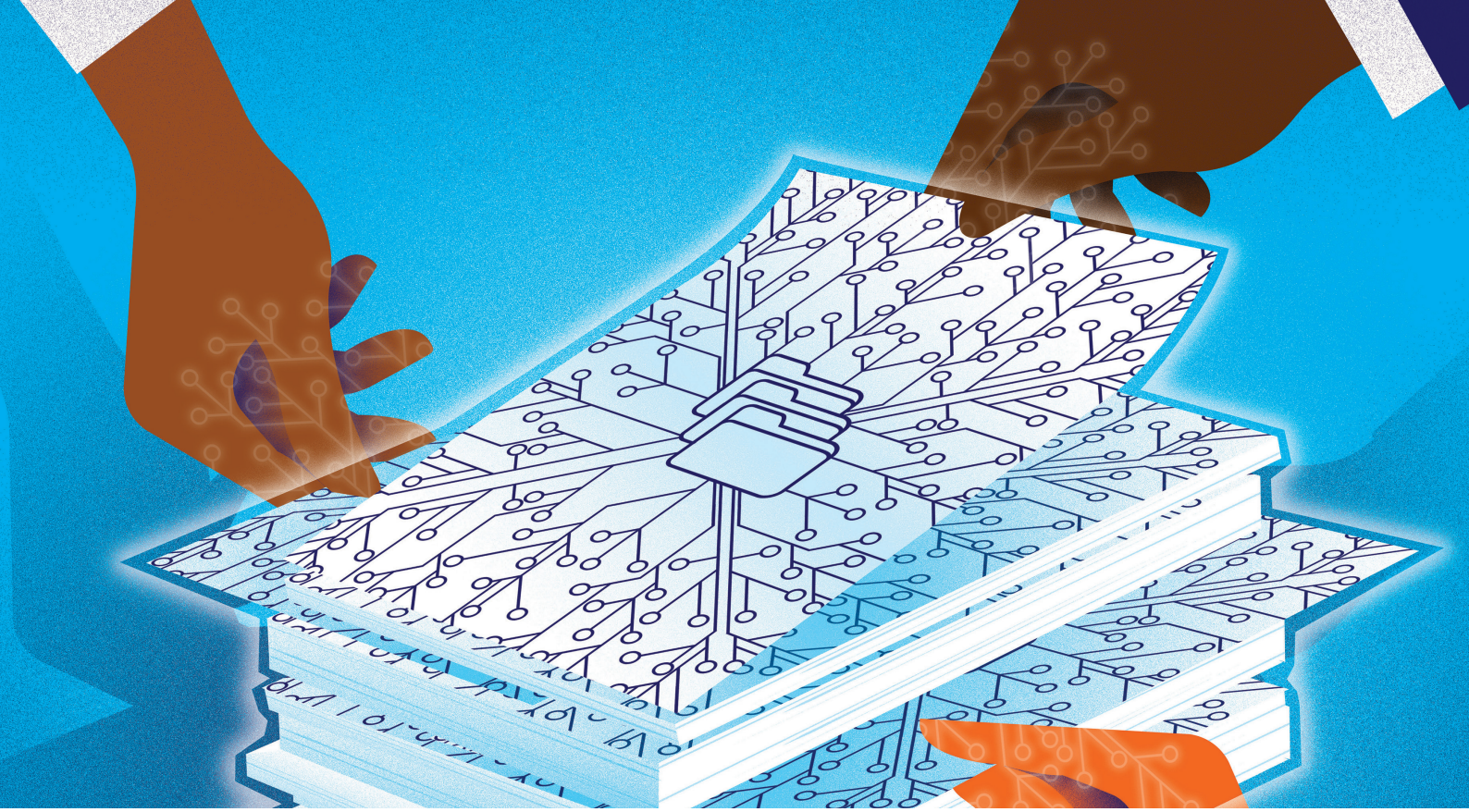

USING AI IN DISCLOSURE
A ROUNDTABLE DISCUSSION

Fiona Campbell of Fieldfisher LLP, Dan Heinrichs of Sky Discovery, Andrew Moir of Herbert Smith Freehills Kramer LLP, and Nathalie Baker and Jason Keenan of FTI Consulting discuss the evolving role of AI and GenAI in the context of disclosure exercises in litigation.

Large-scale disclosure exercises have become a defining feature of modern litigation, driven primarily by the sheer volume of electronic documents that are now created in the ordinary course of business. The use of email, shared drives, mobile devices and collaboration platforms such as Microsoft Teams generate a constant stream of potentially disclosable documents. As a result, disclosure has shifted from a finite paper-based exercise into an open-ended electronic process requiring the identification, search and review of extensive material held across multiple systems.

Against this backdrop, AI is increasingly positioned as a practical way to navigate these challenges. The use of AI, as distinct from generative AI (GenAI), is not new in the context of disclosure. Predictive coding has been used for many years to identify relevant documents within a volume of electronic documents and was explicitly approved by the High Court in 2016 in *Pyrrho Investments Ltd and another v MWB Property Ltd and others* and *Brown v BCA Trading Ltd* (*[2016] EWHC 256 (Ch), see Briefing "E-disclosure: lift off for predictive coding technology?", www.practicallaw.com/8-628-4545; [2016] EWHC 1464 (Ch), see News brief "Predictive coding: an essential consideration?", www.practicallaw.com/0-630-1983*).

A decade on, there has been a rapid emergence of GenAI tools. However, the use of GenAI is an evolving area and there is an inherent tension between its potential efficiency gains and concerns around transparency, confidentiality and accuracy. In particular, GenAI can produce plausible but incorrect output, known as hallucinations, requiring robust governance and human verification.

In this article, the authors discuss the benefits, challenges and potential future developments in this rapidly evolving field (*see box "Author profiles"*).

PREDICTIVE CODING

Predictive coding is a form of machine-learning based technology-assisted review (TAR) where, in broad terms, lawyers review and code a sample set of documents as relevant or irrelevant and an algorithm uses the patterns it learns from those coding decisions to predict the degree of relevance of the remaining documents.

Current adoption

We asked our panel to assess the current state of predictive coding adoption: how widely is it now being used and what type of case is it best suited for?

Author profiles

Fiona Campbell is a Director at Fieldfisher LLP and a UK and Ireland qualified commercial litigator specialising in complex, high-value disputes and technology-driven disclosure. She is a recognised leader in AI governance and digital evidence strategy, advising both on the responsible use of AI within legal practice and on compliance frameworks for clients navigating emerging regulatory obligations. Fiona is an award-winning AI innovator and frequent speaker on the legal and regulatory implications of artificial intelligence in dispute resolution.

Dan Heinrichs leads the Consulting, Solutions and Client Engagement departments at Sky Discovery UK. In this role, he oversees both the practical application and technical development of AI solutions on client matters and advises legal teams on their proportionate and defensible use in disclosure.

Andrew Moir is a Partner in the Intellectual Property Group at Herbert Smith Freehills Kramer LLP and is also Global Head of the firm's cyber and data security practice. With a background in physics and software engineering, Andrew has throughout his career sought to bridge the gap between law and technology. Most recently, this has included building software to manipulate and interrogate data in bulk using generative AI (GenAI), as well as building (from scratch) bespoke machine-learning algorithms to intelligently classify data. Andrew is also the author of Practical Law's e-disclosure practice notes.

Nathalie Baker is the Lead Data Scientist for D&A EMEA at FTI Consulting. She has extensive experience building robust machine-learning solutions for clients, particularly in the litigation space, including applications of natural language processing and GenAI. She has more than six years of experience implementing and fine-tuning large language models. Additionally, Nathalie has conducted various AI assessments of third-party machine-learning pipelines to ensure robustness, adherence to best practices and technical alignment with regulations.

Jason Keenan is a Managing Director at FTI Consulting and a California-qualified lawyer turned eDiscovery specialist. Based in London, he advises global legal teams on technology-assisted review, analytics and GenAI-enabled discovery workflows. He emphasises defensibility, validation and clear documentation, supporting efficient review on complex disputes and investigations.

The authors all note that Practice Direction (PD) 57AD (Disclosure in the Business and Property Courts) requires parties to consider using this type of technology where there are datasets of more than 50,000 documents, which is not uncommon.

Fiona Campbell of Fieldfisher LLP and Dan Heinrichs of Sky Discovery agree that predictive coding, in the form of continuous active learning (CAL), has become almost routine in the context of large-scale disclosure exercises (*see box "Key terms"*). For cases with substantial volumes of documents (say, anything over 20,000 documents), Fiona sees CAL being used as a matter of course as it helps to surface the most relevant documents quickly; it is used in most of Fieldfisher's significant cases.

At Sky, CAL is used in over 90% of litigation matters (and more than 50% of arbitration matters) involving a significant volume of documents. However, Dan comments that CAL is less helpful in the context of investigations, which he views as a separate category, where documents are reviewed in a different way to a disclosure exercise.

Andrew Moir of Herbert Smith Freehills Kramer LLP agrees that the use of predictive coding (generally now CAL) is common in large-scale commercial litigation and regulatory investigations with high data volumes and is increasingly standard in cases under PD 57AD, while noting that adoption does vary by firm, jurisdiction and the circumstances of the claim. Some still prefer manual review where the claim value justifies it or due to the perceived risks of, or lack of familiarity with, the technology.

Cost-effectiveness

We asked our panel to identify the types of cases where predictive coding is most cost-effective.

Fiona has found that CAL comes into its own where there is a huge corpus of documents (hundreds of thousands or millions) and it is not cost-effective to have reviewers reading every one. CAL allows teams to prioritise what is most likely to be relevant. As the review progresses, there will be diminishing returns: the system will indicate that the remaining material has a very low chance of being relevant, so continuing to review it would be disproportionate in terms of both time and cost.

Andrew agrees, commenting that while there is an initial overhead in setting it up, this is outweighed by the economies of scale that come from semi-automated review. He adds that it is particularly useful where relevance criteria are well-defined and consistent across the dataset. It is less suitable for small-scale disputes, where document context is nuanced or there are a significant number of different issues, or where the review requires human judgement.

Dan shares these views, observing that a core application of CAL in document-heavy litigation is using it to prioritise documents for review, therefore reducing the manual review required. He notes that predictive coding technology is now incorporated into all major e-disclosure tools at no additional

cost, so the costs are directly related to the human effort required.

Andrew adds that, historically, the pricing of predictive coding could sometimes make it uneconomical just when it might be needed the most; it used to be disproportionately expensive to use the bigger the case became. However, pricing models have matured over the years.

The courts' stance

As predictive coding has become increasingly prevalent in disclosure, we asked our panel: how proactive have the courts been in finalising the terms of orders providing for the use of predictive coding and what judicial attitudes have emerged toward these technologies?

Andrew observes that there have not been many cases relating to the use of predictive coding, but the decisions that do exist support its use. *Pyrrho*, in 2016, was the first case in which the High Court approved the use of predictive coding to assist with electronic disclosure. Later that year, for the first time, the court ordered the use of predictive coding despite one party's objections to its use in *Brown*.

In *Isbilen v Turk*, the High Court endorsed a two-stage process to check for privilege and the relevance of electronic documents and other imaged information (*[2022] EWHC 697 (Ch)*). The first stage was for keyword searches to be run to exclude potentially privileged documents. The second stage was for the non-privileged documents to be reviewed using predictive coding to establish relevance. The court was satisfied that predictive coding was an "appropriate way to proceed". Andrew suspects that the absence of cases is because the use and understanding of predictive coding has become sufficiently widespread, and the court's view on it sufficiently clear, that parties now rarely challenge its use.

Fiona agrees that, having set the tone back in 2016 in *Pyrrho* and *Brown*, the courts have shifted from cautious approval to routine acceptance of predictive coding. The use of TAR is now standard practice and is expressly recognised and built into the disclosure review document (DRD) template under PD 57AD. In her experience, the courts take a pragmatic view, expecting parties to co-operate and agree an approach rather than have the court dictate it. While the courts do not usually impose TAR, they do expect practitioners to have considered it, so arriving at court with a plan to manually review hundreds of thousands of documents without justification may raise eyebrows. In practice, most DRDs either include TAR from the outset or leave room for it to be applied as part of a sensible, technology-driven workflow.

This mirrors Dan's experience. He has found that most parties have sophisticated e-disclosure advisers who will agree that the use of technology generally and CAL (and, increasingly, GenAI) specifically is effective for document review. Communications between the parties tend to centre on how CAL is deployed and validated, rather than disputing its use. Sky has only had one matter where the other party declined to agree to use CAL and requested an extension to complete its manual, linear review. In that case, the court declined the extension and preferred the use of CAL, which would enable the parties to meet the relevant timelines. As parties usually reach agreement, the courts generally do not need to be proactive about encouraging them to use technology. In practice, agreeing specific keywords usually involves far more heated discussions.

Key terms

Absolute relevance determination. A decision made by AI about whether a document is relevant or not relevant to a case, without further human review or validation.

Auditability. The capability of an AI system to be independently inspected, traced and verified in terms of how it operates, how it makes decisions and the data it uses.

Continuous active learning (CAL). A machine-learning approach where the AI system continuously learns and improves its document classification as human reviewers provide feedback on relevance decisions throughout the review process.

Defensibility. The ability to justify and explain an AI-assisted review methodology, decisions and results if challenged by opposing parties or the court.

Non-deterministic. A generative AI system may produce different outputs when given the same input, rather than generating identical results each time.

Reproducibility. The ability to achieve the same or similar results using the same dataset and algorithm within the same environment.

Retrieval augmented generation (RAG). The process of enhancing the output of a large language model by enabling it to reference a wider authoritative knowledge base beyond its original training data, before generating a response.

Transparency. The ability to understand and explain how an AI system works, makes decisions and produces its outputs.

Validation. The process of evaluating an AI model's performance on part of a dataset to assess its accuracy, reliability and generalisability before deployment across the wider dataset.

EVOLUTION OF GenAI

Whereas predictive coding is primarily used to classify documents by their likely relevance, GenAI is increasingly being used to support substantive review by working directly with the content of documents.

GenAI functionality

We asked our panel: what is different about GenAI in the context of disclosure?

Jason Keenan of FTI Consulting explains that GenAI represents a shift from classification to comprehension. While predictive coding classifies documents based on examples, GenAI can interpret content, summarise themes and answer context-specific questions with justification and reasoning provided (*see box "Is there still a place for predictive coding?"*). Andrew adds that predictive coding is entirely mathematical

Is there still a place for predictive coding?

As generative AI (GenAI) rapidly transforms the legal technology landscape, we asked our panel whether there is still a place for predictive coding.

Fiona notes that GenAI has not yet received judicial or legislative endorsement but, for now, practitioners can agree to use GenAI within the disclosure review document and stand by the validation if they are confident in it. However, predictive coding remains the defensible, transparent benchmark whose output can be fully scrutinised if challenged.

Dan points out that the statistical validation underpinning predictive coding, which enabled courts to approve its use in *Pyrrho Investments Ltd and another v MWB Property Ltd and others*, is the same methodology now being applied to validate AI results (*[2016] EWHC 256 (Ch)*). Importantly, predictive coding can complement GenAI rather than oppose it. By using predictive coding to prioritise documents, lawyers can direct GenAI to review the most likely relevant material first, then decide whether to cut off the review or continue. He cautions that until there is case law approving the use of GenAI in document review, some firms may prefer predictive coding for its greater certainty and reduced risk of scrutiny.

Andrew agrees that, rather than replacing predictive coding entirely, GenAI may complement it, particularly in early-stage review or document summarisation. Predictive coding remains valuable for classification tasks (such as relevance), is well-established and court-approved, is iterative (so that its accuracy inherently improves the more human validation is applied) and is often more transparent and predictable than GenAI.

Jason observes that predictive coding remains useful for now, although, in the UK, it is generally confined to larger commercial disputes where volume justifies the investment. Newer technologies such as continuous active learning have already been widely adopted, offering more dynamic ranking and lower upfront costs. He feels that, as GenAI workflows mature and costs fall, predictive coding may struggle to compete with the next-generation models.

and deterministic, using only the dataset in question. In contrast, GenAI uses language models that have been trained on external materials alongside the dataset, giving them the ability to "understand" and answer questions, and is non-deterministic.

Jason notes that GenAI allows legal teams to move from identifying what is relevant to understanding why it matters, which supports fact-finding, issue analysis and narrative building. Andrew highlights its ultimate future potential to answer general questions across entire datasets; for example, if the prompt is along the lines of "explain how the fraud was perpetrated" a GenAI tool could review the entire document set and answer the question as a narrative, as distinct from using the prompt "identify whether this document relates to fraud" where the tool assesses this question for each document separately. However, technical challenges remain with the former, meaning that its utility, capability and reliability are still evolving. The issues principally relate to "context window" limitations; that is, how much data GenAI tools can consider at once. Some of these issues are being overcome through techniques such as retrieval augmented generation (RAG) (*see "Key use cases" below*).

Fiona points out that the regulatory treatment of predictive coding and GenAI also differs. Predictive coding is expressly recognised under PD 57AD and has had judicial backing since *Pyrrho*, while GenAI lacks formal footing (*see "Court approval and guidance" below*). The use of GenAI depends on practitioner co-operation under PD 57AD's collaborative framework. It is not yet part of the standard DRD, although parties can amend the DRD to include agreed provisions on the use of GenAI and can adopt the International Legal Technology Association's (ILTA) Generative AI Best Practice Guide (ILTA guidance) on a case-by-case basis (*see below*).

Key use cases

We asked our panel: what purposes can GenAI be used for in a disclosure context; that is, where can it help?

Andrew references the ILTA guidance, which suggests a number of possible use cases in e-disclosure (*see box "ILTA potential use cases"*). However, he cautions that these applications carry significantly different risk profiles; for example, absolute relevance determinations or privilege assessments will carry higher risk than document summarisation.

Dan points out that GenAI technology is evolving rapidly, with applications that seemed like the "holy grail" only a year ago now available. Unsurprisingly, due to the potential costs savings, e-disclosure platforms initially focused on using AI for first-level review (in place of a large document review or paralegal pool). However, numerous other applications have emerged, such as:

- Early case assessment, including identifying trends, drafting dramatis personae, building chronologies, identifying key documents, creating categories and identifying possible missing documents through gap analysis.
- Chatbots which can summarise documents, answer specific questions (such as "is there any discussion about pricing or general commercials?") and be used for objective coding (that is, "extract the date from this document") or translations.
- Tools to improve the quality of data, such as AI-backed optical character recognition (OCR) tools that can read handwriting and poor quality scanned documents.
- Facial, image and colour recognition, which can remove hundreds of manual hours of work by searching thousands of images with text or keywords; for example, a search for all photographs featuring boats.
- RAG, which Dan views as the most exciting technology at the moment. RAG makes a case securely searchable so that lawyers can ask specific questions, such as about a witness's involvement in

particular transactions or what job role a particular custodian had on a specific date, enabling teams to build their case strategy or maybe even facilitate settlement.

Jason has found that GenAI allows legal teams to gain insight earlier, validate findings faster and focus human review on the most meaningful material. This is achieved through more adaptive and context-aware TAR workflows, generating chronologies and summaries, performing deep analysis across complex datasets, interpreting relationships between entities and improving data extraction.

Fiona considers GenAI to be valuable for time-consuming administrative work that is sometimes done by juniors or paralegals, such as summarising datasets, building chronologies, mapping relationships and surfacing anomalies or patterns that are not obvious to the human eye, as well as "interrogating" the documents on particular points. It is especially powerful in complex matters such as fraud investigations, where it can draw out recurring themes, inconsistencies or unusual activity "holistically" rather than document by document, reducing reviewer bias and connecting the dots faster.

ILTA potential use cases

The International Legal Technology Association's (ILTA) Generative AI Best Practice Guide suggests a number of possible use cases in e-disclosure (*www.iltanet.org/resources/active-learning-best-practices*). These include:

- Data processing and enhancing text recognition.
- Converting documents to text for search and analytics.
- Issue and document identification, categorisation and prioritisation.
- Multi-document summarisation.
- First-pass relevance review, with the intention to entirely replace first-level review.
- Redaction assistance.
- Privilege identification.
- Quality checking and sampling.
- Sentiment analysis.
- Chain of inquiry analysis.
- Anomaly and pattern detection.
- Foreign language review.

COURT APPROVAL AND GUIDANCE

The landmark decision in *Pyrrho* provided judicial endorsement for predictive coding in disclosure. We asked our panel whether similar express approval is required from the courts regarding the use of GenAI in the context of disclosure, and whether new court rules or guidance would be useful.

Current frameworks sufficient for now

Fiona does not consider that a *Pyrrho*-style judgment is needed for GenAI. PD 57AD encourages co-operation and agreement between parties on the use of technology, rather than pushing the courts to rule on it. To the best of her knowledge, no case has yet required judicial approval for GenAI and the courts appear open to its responsible use within this collaborative framework. However, practitioners should be aware that failing to verify any AI output that is relied on in court may amount to professional misconduct. The courts have criticised lawyers who submitted AI-generated material that contained fabricated authorities (*Taiwo v Homelets of Bath Ltd and others [2025] EWHC 3173 (KB), see News brief "Warnings on AI-generated false authorities: risks for practitioners", www.practicallaw.com/w-049-2864; Ayinde v Haringey LBC [2025] EWHC 1383 (Admin), www.practicallaw.com/w-047-7186*).

Ayinde prompted both the Solicitors Regulation Authority and the Bar Standards Board to review their existing frameworks to ensure that verification duties apply equally to AI-generated work as they do to work produced by a junior or trainee (*see box "Guidance from professional bodies"*).

In terms of court rules or guidance, Fiona favours a wait-and-see approach as the rapid evolution of technology could make any "hard-coding" of GenAI into the rules become outdated quickly. For now, more rules would risk duplication; it is better to apply the existing guidance rigorously and revisit it if practice outpaces it. Inevitably, there will eventually be a dispute over the use of GenAI, which will give the courts an opportunity to provide guidance on best practice and the required standards of defensibility.

Judicial scrutiny likely

Andrew agrees with Fiona that the current rules are sufficiently general to cover GenAI. The rules do not single out other disclosure technologies, so it is not necessarily the case that GenAI needs special treatment. As disputes over the use of GenAI are likely to come before the courts, he thinks that there will be court judgments. Although, given the variety of tools and approaches, there is likely to be some element of approval on a case-by-case basis, he considers that a reported judgment would be helpful, as *Pyrrho* was for predictive coding. Andrew also notes that, given the judicial criticism in *Taiwo* and *Ayinde*, parties might wish to bring any use of GenAI to the attention of the court even where they agree on its use and approach.

Judicial approval and increased use of GenAI

Dan believes that the lack of court precedent is a challenge. Sky has various matters where parties are seeking to agree on the use of AI in document review. Practical questions include whether it will be sufficient to provide the same metrics for AI tools as for predictive

Guidance from professional bodies

As generative AI (GenAI) becomes increasingly embedded in legal practice, UK professional bodies have issued guidance for practitioners, including:

- The Courts and Tribunals Judiciary updated AI guidance, dated 31 October 2025, which warns practitioners to use AI cautiously, preserve confidentiality and verify outputs (*www.judiciary.uk/guidance-and-resources/artificial-intelligence-ai-judicial-guidance-october-2025/*).
- The Law Society's October 2025 guidance on GenAI, which sets out practical guardrails for solicitors (*www.lawsociety.org.uk/topics/ai-and-lawtech/generative-ai-the-essentials*).
- The Solicitors Regulation Authority's November 2023 risk outlook report, which flags risks around privilege, confidentiality and data protection (*www.sra.org.uk/sra/research-publications/artificial-intelligence-legal-market/*).
- The Bar Council's updated guidance on GenAI, published on 26 November 2025, which warns explicitly about fabricated outputs and the duty to verify (*www.barcouncil.org.uk/resource/updated-guidance-on-generative-ai-for-the-bar.html*).

coding, whether human validation will be required, how large the validation set needs to be and whether each issue coded by the AI needs its own validation set. There are also challenges in terms of balancing the need to explain the underlying technology supporting the specific tool used while not compromising intellectual property rights. Another point concerns whether prompts are privileged or should be shared with the other side (similarly to agreeing keywords for searches). Dan concludes that the ILTA guidance is helpful but legal judgments on some of the key issues would be beneficial.

Standards for auditability

Jason argues that while express judicial acceptance of GenAI would be welcome, its use in disclosure is already gaining momentum. What is needed now is not judicial approval of the technology, but clear expectations around documentation, validation and transparency. The goal should not be to regulate specific technologies but to establish standards for auditability, reproducibility and human oversight. Clear guidance will give practitioners the confidence to innovate responsibly while giving the courts a common framework for assessing proportionality and defensibility.

SCOPE OF SEARCHES

We asked our panel to consider what the challenges are in terms of satisfying the opposing parties or the court regarding the scope of searches undertaken where GenAI tools are used.

Jason highlights that the main challenge is transparency. As GenAI is non-deterministic, parties should be prepared to evidence how results were validated. Courts and opposing parties will expect clear documentation of how the tools were applied, what data was accessed and how outputs were confirmed. Maintaining audit trails and engaging early with the opposing party on methodology will be important in establishing defensibility.

Fiona agrees that defensibility is key and recommends maintaining clear records of prompts and outputs. She adds that disputes over methodology are less likely where both sides use comparable GenAI tools. However, Dan notes that when the question of which underlying technology has been used arises, it can be challenging for the parties to evaluate and compare different GenAI tools when developers are understandably cautious about disclosing their intellectual property and large language models (LLMs), the underlying technology used in GenAI, are updated every few months.

Andrew agrees that consistency will be under scrutiny. GenAI does not always give the same answer twice, even based on the same prompt applied to the same data. This can particularly be an issue where, due to context window limitations, running a GenAI prompt over a large set of disclosure documents typically involves running the same prompt multiple times across different batches of documents. GenAI document review must include robust quality control and verification processes, such as dip-sampling or full second-level review of responsive documents. Successful dip-sampling verification may reduce arguments around whether the underlying prompt should be disclosed.

Andrew adds that future court decisions on GenAI use will likely balance potential cost and efficiency savings against the risk of missing relevant documents. We can expect significant focus on the prompt used to define relevance, similar to the scrutiny of keyword searches in conventional approaches, with challenges in more complex cases around fitting a description of the issues into a sufficiently short prompt. Each party's strategic approach will also be important; for example, if a party's focus is on finding everything that is relevant, it may be less interested in cost or efficiency savings.

Jason and Fiona both observe that most of the current use of GenAI is on already-filtered datasets, after search parameters have been agreed. As oversight frameworks mature, use cases are likely to extend further upstream into searching and culling before review.

COST-BENEFIT ANALYSIS

Costs remain a key consideration in disclosure and the adoption of new technologies. We asked our panel several questions about the cost-effectiveness of using predictive coding and GenAI in disclosure.

Cost comparison between AI and GenAI

We asked our panel how the costs of GenAI and predictive coding tools compare, and whether this is a relevant comparison.

Maturity and pricing models. Fiona and Andrew both note that predictive coding is a mature technology that is now built into most e-disclosure platforms' overall costs (*see box "Charging practices"*). GenAI initially used per-document or per-prompt pricing, but some providers are starting to bundle AI costs into platform fees. Fiona cautions that the comparison is not strictly like for like, as predictive coding is established and accepted while GenAI is still emerging. Ultimately, what matters is the overall proportionality of disclosure costs, which must still be approved by the court through Precedent H.

Dataset size. In Dan's experience, datasets of fewer than 50,000 documents are unlikely to be good candidates for the use of GenAI for document review. Drafting, iterating and testing prompts requires considerable time for both legal and technical teams, which may be prohibitive for a comparatively small dataset. Statistical sampling for 92.5% to 97.5% accuracy requires approximately 1,500 documents to be reviewed. With three or four iterations, senior members may review more than 6,000 documents before prompts are accurate enough and then must perform quality checks against the resulting AI coding decisions. Predictive coding, by comparison, can be worthwhile for datasets as small as a few thousand documents, even if it is just used for prioritisation.

Conversely, Jason observes that predictive coding remains cost-effective for certain large datasets, such as US regulatory reviews where upfront model-training investment by senior counsel or subject matter experts can be justified by a large reduction in traditional linear review costs. Disclosure review exercises tend to be smaller in the UK than the US, with use generally limited to large commercial disputes. CAL and GenAI require far less set-up time and can begin delivering results almost immediately, even for smaller datasets. Over time, the overall cost profile will increasingly favour GenAI tools that combine speed, flexibility and contextual depth.

Other factors. Dan notes that, for document review, key factors include the size of the document set, the type of document and the complexity of the issues but also the legal team's experience of drafting prompts, which can have a significant impact on costs. Jason agrees, noting that, while predictive coding offers linear efficiency gains, GenAI provides rapid qualitative insight and richer analytical output. The decision point will depend on project objectives, such as classification at scale versus contextual understanding. Andrew comments that while GenAI may have higher upfront costs, it offers broader functionality and may reduce downstream costs, such as review, summarisation and quality control.

Where GenAI works

We asked our panel whether there is a volume threshold where it might be cost-effective to use GenAI tools to assist with e-disclosure and whether there are particular types of cases where GenAI is likely to be particularly useful.

Complexity and nuance. Andrew notes that GenAI becomes cost-effective in cases with complex issues or large volumes of unstructured data, such as emails and chat logs, provided that costs do not scale disproportionately with volume. Some vendors suggest a range for data volume where the use of GenAI is most effective.

GenAI is particularly useful in cross-border matters involving multilingual datasets because it can be "language agnostic". Importantly, the threshold is not just volume but complexity: GenAI excels where traditional TAR struggles to capture nuance, particularly where single documents contain content that is relevant to more than one issue simultaneously, since TAR algorithms effectively only look at relative word prevalence.

Fit for purpose, not just volume. Dan agrees that the question is less about document volume and more about the problem being solved. It would also be reductive to assume that GenAI only has applications in document review. His team has used AI to compile a detailed overview of an employee's job, based on 20 lengthy documents, in the context of an employment dispute, and to create detailed product development chronologies, based on a few hundred documents, all within hours.

High-complexity matters. Jason agrees that GenAI delivers the best value where interpretation and insight are critical, such as investigations, competition or regulatory inquiries, and cases with highly technical or diverse datasets. It excels at understanding relationships, behaviours and context, rather than simply classifying relevance. However, the time and cost involved in refining prompts and validating results can make GenAI less practical for small or straightforward datasets, where traditional analytics may remain more efficient.

No strict volume threshold. Fiona adds that, as GenAI is now integrated into most

Charging practices

As generative AI (GenAI) tools become more common in disclosure, billing models may need to evolve. We asked our panel how they expect the costs for GenAI-assisted work to be charged.

Andrew expects to see a similar evolution to the billing models for predictive coding technologies. Initially, the costs will probably be treated as disbursements, as the services will most likely be provided by third-party vendors. However, as firms bring technology in-house or "white label" it, they will set the pricing models. Where tools charge on a per-document or text volume basis, these costs will likely be passed on to clients.

Jason notes that most providers are moving toward usage-based models, paying per prompt, token, or query volume. Service providers will likely blend time-based professional services with separate AI utilisation fees, although some technology companies are beginning to offer AI-inclusive packages with unlimited use. Clients are already asking for transparency about AI-assisted work and its impact on time and cost, so billing models will evolve accordingly.

Dan observes that AI work is currently charged using billable hours plus per-document or per-token (about three to four words) rates. If there is consolidation and AI pricing reduces (or becomes included in platform fees, for example), he believes that AI costs will be "baked in" with possibly higher hourly rates or platform fees; similar to how email replaced letters but the billable hour model remained the same, just with increased productivity.

Fiona is also already seeing GenAI costs being absorbed into broader platform licences, making adoption easier without materially affecting overall disclosure costs. If anything, it enhances transparency by making reviews more auditable and efficient. The only real consideration is equality of arms; the fact that some parties may have access to more advanced tools than others needs to be factored in when agreeing an approach in the disclosure review document.

e-disclosure platforms, there is no strict volume threshold. However, it is best used on filtered, issue-specific datasets rather than raw data, particularly in complex or fast-moving matters such as fraud, regulatory or investigatory cases where themes or connections must be identified quickly.

Potential impact of GenAI

As firms seek greater efficiency in disclosure, we asked our panel about the impact, in terms of the overall cost and time, of using GenAI in a large disclosure exercise.

Dan believes that the cost of large disclosure exercises should shrink dramatically with the use of GenAI. Human document review is the main cost driver, and this can be significantly reduced by applying GenAI at the first level. It also reduces timelines: his team has successfully prepared barristers to argue in court that nine months for a disclosure review exercise of 15 million documents at source was unnecessarily long, even using predictive coding rather than GenAI, which would shorten timeframes further.

Fiona agrees that, when deployed properly, GenAI can materially reduce time and costs by identifying key documents and themes early, and completing tasks that once required extensive first-level review more efficiently. As platform pricing becomes more inclusive, its deployment improves proportionality by reducing unnecessary review time.

Andrew sees significant time savings in first-level review and issue identification, reducing reliance on large review teams who can be redeployed to quality control and other tasks. For law firms that adopt this technology in-house, costs may shift from manual labour expenses to technology investment. While this requires upfront capital, overall costs should decrease over time.

Jason notes that the combination of CAL and GenAI has already reduced review timelines and improved quality of insight. GenAI eliminates much of the manual summarisation and narrative construction that is traditionally done late in a project, while CAL streamlines document ranking. As confidence grows, these technologies will further compress review timelines, allowing teams to focus on substantive legal analysis rather than mechanical review.

IN PRACTICE

While the potential of GenAI in disclosure is increasingly recognised, translating theory into practice raises numerous operational questions. From managing the rapid pace of technological change to determining issues around the drafting of prompts, legal teams face practical challenges in implementing these tools effectively.

Pace of change

We asked our panel: how quickly are things developing and are firms already using GenAI in a significant way to assist with disclosure?

Dan describes the current rate of adoption as "breathtaking". Unlike predictive coding, which took seven years from availability to the *Pyrrho* precedent and was not standard practice even five years later, GenAI has progressed remarkably faster. Just two years after the public release of ChatGPT and a year since mainstream platforms launched their own AI tools, parties are already proposing, and agreeing, to use GenAI for disclosure review. The general societal acceptance of GenAI tools outside of the legal arena has encouraged this rapid adoption. Providers are rapidly expanding AI-backed tools, and Sky is updating its demonstrations of GenAI tools and features every three to four weeks to reflect enhancements.

Andrew agrees that adoption is accelerating, particularly among firms with strong innovation teams or partnerships with legal technology providers. Some firms are experimenting with GenAI in "sandbox" environments before live deployment. Disclosure providers are rapidly integrating GenAI into existing platforms, making it more accessible to firms without in-house AI expertise. However, client attitudes vary, with some maintaining a "no GenAI" policy in their supply chain due to perceived risks, although this is often navigable for specific uses and products.

Jason also observes that adoption is advancing rapidly, with many law firms and service providers already piloting GenAI tools for summarisation, privilege review and disclosure preparation, as FTI has seen across hundreds of matters over the past two years. The technology is moving faster than procedural frameworks, raising questions about whether formal consultation is needed to evolve court rules.

Fiona has found that, while early uptake was cautious, GenAI is now being used more widely, particularly in regulatory investigations and inquiries where speed and insight are critical. She uses GenAI for early case assessment, data familiarisation and issue-driven analysis before promoting documents for full review, but expects that first-level review will eventually become largely automated.

Prompts

Some firms will have in-house expertise while others will engage with e-disclosure providers. We asked our panel how, in their experience, things are likely to work; for example, who drafts the prompts and might these need to be agreed or protected from disclosure?

Prompt drafting. Jason has found that prompt engineering is typically handled by disclosure specialists or data scientists working with legal teams to ensure that the instructions align with case strategy and privilege considerations. Andrew agrees, suggesting that legal teams will principally write prompts with guidance from disclosure providers or in-house data scientists but adds that, if GenAI is being used for relevance review, the prompt will equally need to cater for the opponent's case. Fiona notes that most systems now guide you through prompt creation automatically. In practice, her team still agrees on the framing of prompts at the outset, but the platforms largely refine them.

In Dan's experience, e-disclosure providers are heavily involved in drafting prompts and take on the administrative work involved in using the various AI tools. The legal teams have the final say on the specific prompts and, as they have the most thorough understanding of a matter, will complete the bulk of the drafting. As Sky's team have developed their prompting skills, over the last 12 months the consulting time required to run an AI review has halved. In addition, lawyers who are now working regularly with AI tools are becoming increasingly sophisticated and accustomed to the requirements of prompt drafting, adding this to their skillset and further reducing the time required to run an AI review.

Andrew envisages that prompts will principally be written by the legal teams with guidance from disclosure providers or in-house data scientists.

Privilege. Fiona views prompts as work product that should be protected by privilege,

similar to review protocols. However, the courts may eventually require prompts to be disclosed to explain a review methodology, so she recommends maintaining a clear repository of all prompts used, for good governance and transparency. Jason agrees that documenting prompts and treating them as potentially privileged is prudent. Whether courts will eventually expect parties to agree prompts (as with search terms) remains an open question, but clear governance around prompt design, approval and validation is essential to maintain defensibility. Andrew adds that, in contentious matters, there will likely be arguments around whether prompts should be disclosed or agreed to ensure transparency and defensibility.

Dan shares these views, noting that the key issue is whether prompts should be treated like keywords, which must be agreed with opposing counsel, or as review instructions, which would be privileged. Defendants will not want to share prompts that may reveal their strategy, while claimants will be very interested in seeing them. He thinks it likely that the courts will need to set a precedent on this issue.

RISK MANAGEMENT

We asked our panel about the potential risks for law firms when using GenAI in the context of e-disclosure, and how these risks can be managed or reduced.

Key risks

Andrew identifies accuracy as the main risk. Accuracy is highly dependent on the quality and thoroughness of the prompt, which requires know how within the legal team: GenAI platforms will not always refuse to answer a question on the basis of insufficient information but instead become more likely to guess or hallucinate. Hallucinations are potentially an issue more generally, particularly when GenAI is summarising or answering questions in relation to a set of documents. Hallucinations may also feed into GenAI's reasoning. Nathalie Baker of FTI Consulting agrees, adding that these hallucinations can appear plausible, especially when context is incomplete, so subsequent human reviewers should treat AI outputs as guidance, not conclusions.

Dan agrees that hallucinations are likely the most prominent risk, but notes that humans are also inherently fallible, especially after hours, weeks or months of reviewing documents. He has heard estimates of 60% to 70% accuracy for some paralegal reviews.

While Fiona acknowledges that there are always risks with workflows that are not independently verified, she adds that, unlike open models, most licensed GenAI systems are designed to return no output rather than false information. Her greater concerns are around bias, confidentiality and privilege, warning that the technology is only as reliable as the controls built around it. Andrew agrees that there is a particular risk of using GenAI to assess privilege, which is difficult and highly nuanced even for humans. Over-reliance on GenAI could be contrary to a lawyer's duty under paragraph 3.2(5) of PD 57AD to undertake a review to satisfy themselves that any claim to privilege is properly made and the duty under paragraph 16.2 of PD 57AD to review redactions to ensure that they are properly made.

Managing the risks

Our panel identified a number of ways that law firms can manage these risks.

Validation. The panel agrees that validation is the key to addressing the accuracy risks of using GenAI in disclosure. Andrew says that, while validating every GenAI output would erode its usefulness, legal teams should use dip-sampling or other approaches, and keep an audit trail of the entire process and prompts used.

Fiona cautions that every GenAI output should be verified, tested and documented before it is relied on. This could include both machine validation and a human validation review. Testing generally entails conducting elusion tests on the documents that are deemed non-responsive.

Nathalie highlights the non-deterministic nature of GenAI and explains that quantifying how often it produces different results can help to understand how this affects the response. However, where GenAI is used to generate text instead of distinct categorical outcomes, validation becomes more complex. Although there are proposed methodologies, the data science industry has yet to create a widely accepted set of measurement metrics that are similar to more traditional machine learning approaches.

Performance measurement. Dan emphasises that, for both human and AI review, defensible and repeatable quality assurance is key. The statistical sampling used in CAL can measure AI performance and Sky has consistently seen accuracy rates above 85%, as checked by a senior lawyer. Nathalie agrees, explaining that the performance of GenAI needs to be measured over time and across different types of documents to mitigate the risk of performance degrading. It is also vital to understand the performance metrics; for example, when using GenAI to classify documents as relevant or not relevant, looking only at the document reduction percentage would be misleading. Combining this metric with the rate of missed relevant documents provides a complete picture of the costs and benefits of GenAI.

Defensibility. Fiona highlights that documentation requirements remain unchanged from traditional disclosure practice: the full methodology and approach must be recorded in case the process is later challenged. The workflow must be defensible and transparent, with proper checks and balances built in. As long as the human-in-the-loop model is maintained and outputs are validated before being put before the court, the risks are low and the process is transparent. The difficulty is getting the balance right: if there is a human in the loop on every GenAI powered decision, there is no point in having used it at all.

Confidentiality and data security. Andrew feels that confidentiality concerns are manageable; for example, for third-party GenAI systems, it will be necessary at the contracting stage to ensure that the system runs in an isolated environment (private instance) and that content is not used to train the model in a way that compromises confidentiality. Nathalie agrees that it is essential to have a secure system in place to reduce the risk of data being exposed to the wrong parties. Ensuring that tools and vendors are vetted and that employees are educated on what tools they should be using also helps to mitigate data security risks.

Dan adds that it is important to speak to the GenAI provider about how they handle data and try to understand how the technology works, at least at a high level, before using it. Information security checks such as checking for ISO27001 or Cyber Essentials Plus (CE+) certification, asking detailed questions about the provider's data-handling practices and which AI models they use, and confirming that they will not train their models on confidential documents are good starting points.

Related information

This article is at practicallaw.com/w-049-3102

Topics

Other links from uk.practicallaw.com/

Topics

AI and dispute resolution	topic/w-048-1294
Disclosure: litigation	topic/4-103-1193
Privilege	topic/7-635-9367

Practice notes

AI and machine learning: overview	w-039-5275
Disclosure: electronic disclosure: some questions to ask clients about electronic documents	3-205-5555
Disclosure: electronic disclosure: what is it and how do I deal with it?	6-205-5554
Disclosure: reviewing the client's documents	9-204-1042
Disclosure: scoping and conducting a search for documents	1-204-1041
GenAI: some points of caution for litigators	w-047-8988
Legal professional privilege in civil litigation: an overview	2-205-2977
Privilege: some common issues	9-205-3997

Previous articles

Contracts for AI products: negotiating AI-specific clauses (2024)	w-043-8385
Privilege for in-house lawyers: navigating the challenges (2024)	w-043-2783
AI governance, risk and compliance: shaping an unknown future (2023)	w-040-0428
AI and automated decision making: to regulate or deregulate? (2022)	w-033-8467
Privilege in investigations: the journey continues (2021)	w-030-9877
The disclosure pilot scheme: views from the ground (2020)	w-025-3284
AI and data protection: balancing tensions (2019)	w-020-9713

For subscription enquiries to Practical Law web materials please call +44 0345 600 9355

Best practice. Fiona concludes that, ultimately, the solution to these risks is simple: human oversight. While AI can support disclosure, it cannot replace professional judgement. AI must be integrated into the broader disclosure framework rather than operating in isolation. Dan suggests treating GenAI output as the product of a very eager, but very inexperienced, trainee and to check everything in detail. GenAI output should be used as a starting point, not as a completed product.

Client use of GenAI

We asked our panel about the potential risks of clients using GenAI to identify relevant documents during a disclosure exercise.

Privilege and professional responsibility

Fiona answers that, if clients are relying on GenAI alone (without human oversight) to identify relevance or assess privilege, there is a real danger of disclosing material that should not be disclosed (or vice versa). This would also be a breach of the procedural rules, as the solicitor with conduct of the matter retains ultimate responsibility for the process and must certify that disclosure is accurate, including the correct application of privilege. When GenAI is deployed within a structured, validated workflow with human review, privilege checks and auditability, those risks are substantially reduced. Current judicial guidance already makes clear that AI use must be disclosed to the court and lawyers remain accountable for any output or decision submitted in their name.

Andrew agrees, adding that law firms would need to have oversight of any client use of GenAI, to the extent that they are asked to validate the process or present it in court, consistent with their overarching duty not to mislead the court.

Validation and defensibility

Nathalie warns that there is a risk of over-reliance on GenAI outputs without sufficient validation. While this can be managed with clear validation and sampling protocols, as already seen applied in traditional predictive coding or CAL workflows, the non-deterministic nature of GenAI makes it particularly difficult to validate the output. This means that where clients use GenAI to inform disclosure decisions, opposing parties may challenge the defensibility of the process. Practitioners should therefore ensure that any client-driven use of AI is supervised by counsel, logged, and validated through sampling. Robust documentation and independent testing by disclosure experts can provide further assurance if the AI outputs are later questioned in court.

Dan takes a cautious view and would advise against using GenAI to identify the relevance of documents at source, such as on an email server. He believes that the technology is not yet sufficiently robust to be defensible in legal proceedings and adds that any identification of data sources should be done by human qualified forensic professionals.

LOOKING AHEAD

We asked our panel where they think we will be in terms of the approach to e-disclosure in five years, and to suggest two or three key changes that they expect to see.

Agentic AI and flexible regulation

Fiona notes that, five years ago, no-one expected GenAI to exist, let alone reshape legal practice, so predicting the next phase is difficult. But she expects to see the rise of agentic AI, with systems autonomously handling defined disclosure tasks such as privilege review, redaction, issue tagging and confidentiality checks. She also thinks that validation techniques for GenAI will mature to a point where they gain formal judicial acceptance. On the legislative front, she expects more flexible, principles-based language that is designed to adapt to new technologies, rather than trying to regulate each one specifically.

Normalisation of GenAI

Jason predicts that GenAI will replace predictive coding as the dominant methodology for data assessment. The emphasis will shift from training static models to deploying adaptive, conversational tools that are capable of reasoning across entire datasets. AI-assisted disclosure will be fully normalised, with courts focusing on validation, transparency and proportionality rather than on the technology itself. This means that disclosure will evolve into an interactive,

insight-driven process, where lawyers use GenAI to interrogate data dynamically, surface key facts and understand narrative context far earlier in the case lifecycle.

Integration and industry standards

Andrew expects that GenAI will be embedded in mainstream disclosure platforms, with courts being more comfortable approving its use. However, he predicts that there will be some high-profile mishaps with the use of GenAI in disclosure exercises without suitable validation, checks and balances. It is likely that industry standards or protocols for GenAI used in litigation will emerge, in a similar way to e-discovery and data exchange protocols.

"Over-disclosure" and the end of linear review

Dan anticipates an increase in disclosure volumes. There has already been some limited case law ordering the "over-disclosure" of documents, such as in *Genius Sports Technologies Ltd and others v Soft Construct (Malta) Ltd and others* and *Kent v Apple Inc and another* (*[2022] EWHC 2637 (Ch), www.practicallaw.com/w-037-7167; [2023] CAT 20*). He sees a future where opposing sides perform a cursory privilege and confidentiality review, and then provide all of the documents that could potentially be relevant, as both sides in a dispute are able to use AI to review documents, removing the threat of a "document dump".

Maybe not in five years, but certainly in ten years, Dan believes the concept of a linear first-level review will have disappeared. AI tools will be able to complete full reviews of millions of documents with little instruction and provide the evidence to lawyers to act on within days, rather than many months, of review.

Fiona Campbell is a director at Fieldfisher LLP, Dan Heinrichs is an associate director of consulting at Sky Discovery, Andrew Moir is a partner at Herbert Smith Freehills Kramer LLP, Nathalie Baker is the lead data scientist for D&A EMEA at FTI Consulting and Jason Keenan is a managing director at FTI Consulting.