Understanding Discovery Obligations In Era Of Generative AI

By **Rose Jones, Kassi Burns and Meredith Perlman** (November 29, 2023, 12:36 PM EST)

While the legal profession grapples with heady topics like appropriate use cases, potential bias, privilege and confidentiality considerations, and the application of legal ethics when using OpenAI's ChatGPT and other generative artificial intelligence tools, it is important to look ahead to contemplate how the use of these tools will be addressed in discovery.

This article discusses best practices and strategic insights litigators should consider in federal court litigation when dealing with discovery produced by generative AI tools.

It also contemplates the development of corporate generative AI policies that will address a wide variety of issues related to the use of generative pretrained transformer — or GPT, the large language models underlying many chatbots — tools.

Preservation Obligations

Over the past few years, litigators have increasingly seen issues around obligations and discoverability of other varied and more modern data sources — such as mobile data, hyperlinks and collaboration tools — become an area of focus in discovery disputes.

Court orders around relevancy and proportionality have been issued with regularity as the judiciary addresses how these obligations should apply to emerging technologies.

Preserving generative AI prompts presents its own challenges.

Dynamic Responses

One of the unique challenges GPT tools present is the dynamic nature of the responses. These models generate text responses based on the input prompts.
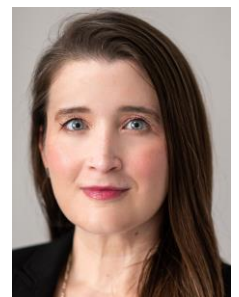
However, these responses are not static and can vary depending on the model's parameters, context and even the same prompt entered at separate times.

This dynamic nature makes it challenging to capture and preserve a specific response tied to a particular prompt.

No Tracking Functionality

Rose Jones

Kassi Burns

Meredith Perlman

GPT service providers typically do not offer native prompt logging or archiving.

This lack of built-in logging capabilities means users must rely on their own recordkeeping methods.

Ephemeral Interactions

GPTs often use real-time or ephemeral online interactions. Preserving such interactions can be logistically complex and may require immediate preservation.

Best Practices

Accordingly, once the duty to preserve is triggered and you determine that generative AI prompts are within the scope of the litigation, you should observe preservation best practices for these tools, such as considering how you will authenticate and verify GPT prompts.

You should develop procedures to ensure content's integrity, which may include documenting time stamps and any potential edits. You will also want to maintain comprehensive records of your use of such tools, including the specific prompts you fed into the tool and the responses generated. This documentation can aid in identifying relevant content.

In addition, you should implement regular backup procedures for GPT prompts to help ensure their preservation, especially in dynamic, collaborative environments. Capture and preserve metadata associated with these prompts, which can provide valuable context and authenticity.

Issue legal hold notices to relevant personnel, including those using GPT tools, to communicate the duty to preserve and prevent inadvertent deletions.

Be sure to provide sufficient explanation addressing the complexity of generative AI tools, especially where self-preservation of prompts is the only option.

This preservation and legal hold notice process should align with any existing corporate generative AI policies.

Production Obligations

In the context of federal civil litigation, production obligations pertain to the process by which parties must provide electronically stored information in response to discovery requests.

These obligations are integral to ensuring transparency and fairness in legal proceedings, as they dictate how parties exchange relevant evidence.

As generative AI tools become more prevalent in business and personal use, it is essential to understand how production obligations apply to the content generated by these systems.

Parties should be prepared to address the unique challenges associated with GPT prompts, including relevance, proportionality, privilege and confidentiality, while employing best practices to fulfill their production obligations.

Relevance

Information subject to production obligations must be relevant to the claims and defenses in the litigation under Federal Rule of Civil Procedure 26(b)(1). GPT prompts, like any other electronically stored information, must meet this relevance standard.

Given the complexities surrounding generative AI technologies, there may be many instances where a given matter does not meet this standard. Parties may dispute the relevance of GPT prompts, particularly if their connection to the case is not immediately apparent.

Also consider the impact of any corporate generative AI policies in place and the requirements such policies may impose on employees to ensure they refine and validate GPT outputs, and confirm the outputs reflect the information and message the employee intends to relay.

Proportionality

As lawyers navigate the production of GPT prompt data in litigation, it is crucial to apply the principle of proportionality. Balancing the relevance and importance of this data with the costs and burden of production is essential.

Courts should consider the volume and significance of this prompt data in proportionality assessments. Federal Rule 26(b)(1) limits discovery to matters that are relevant to any party's claim or defense and proportional to the needs of the case.

You should weigh the significance of GPT prompts and how important they are to proving or defending against claims, as high-impact data is more likely to be proportional for production; examine the volume of GPT prompts as large volumes may require more stringent proportionality considerations than data limited in scope; and calculate the potential costs and burden associated with producing GPT prompts.

Factor in data privacy and confidentiality concerns, as some GPT prompts may contain sensitive or privileged information.

Keep records of your proportionality assessment, documenting the criteria considered, the basis for your decisions and any negotiations with opposing parties.

Note that proportionality is not static, so be sure to reevaluate your assessment as the case evolves and more information becomes available.

By following these practice tips and strategies, you can make informed decisions regarding the proportionality of GPT prompts in litigation, ensuring that they maintain a balance between relevance and burden while complying with legal obligations and ethical considerations.

Form of Production

The form in which you can produce electronically stored information can vary. Courts may specify the format, such as native files, PDFs or structured data. With GPT prompts, choose a format that ensures its usability and understandability by all parties involved.

Like short-form messages, there is no standardized format for production, so the parties should

negotiate in good faith toward a result that is equitable under the Federal Rules of Civil Procedure.

Producing GPT prompts can be challenging. For example, parties must take measures to ensure that all relevant GPT prompts are produced. Additionally, GPT prompts may contain sensitive or confidential information. Balancing the obligation to produce with data privacy and confidentiality concerns is essential.

You should implement quality control measures to ensure that parties produce GPT prompts accurately and completely. If necessary, redact sensitive information from GPT prompts to protect confidentiality.

Lastly, ensure the authenticity and integrity of GPT prompts through proper documentation and verification.

Practical Considerations for Business and Individuals

Businesses and individuals must develop effective data management protocols, implement legal and technological strategies to meet their obligations, and ensure that employees using generative AI tools are aware of their responsibilities in maintaining compliance with data-related regulations.

Be prepared to effectively guide these efforts through your legal advice and consultation.

Ensure that employees have training and awareness of their responsibilities as it relates to the proper use of GPT tools. Here are some examples.

Training Programs

Develop comprehensive training programs for employees using generative AI tools. Ensure they understand their responsibilities in data management and compliance.

Data Ethics, Privacy and Privilege

Educate employees on data ethics, privacy regulations, privilege and best practices for handling sensitive information when using GPT tools.

Privilege training should, at a minimum, address attorney-client privilege and the work-product doctrine, but depending on the industry, it may also need to address other kinds of privilege, e.g., bank examiner's privilege.

Clear Communication Channels

Establish clear channels for employees to seek guidance or report issues related to GPT-generated data and prompts. Encourage transparency in data management practices.

Internal Guidelines and Policies

Create internal guidelines and policies that address the appropriate use of generative AI tools, data handling procedures and legal compliance.

Regular Updates and Refreshers

Periodically refresh training materials and conduct refresher courses to keep employees informed about changes in data management protocols and legal requirements.

Preparing for Our Generative AI Future in Litigation

In the context of federal civil litigation, addressing preservation and production obligations regarding GPT and prompts is essential for maintaining the integrity of legal proceedings. The evolving landscape of generative AI, particularly GPT tools, introduces unique challenges that legal professionals, businesses and individuals must navigate.

Preservation obligations encompass the duty to protect and retain relevant electronically stored information in anticipation of or during litigation. GPT-generated content and prompts, as a form of electronically stored information, require special attention due to its characteristics, such as identification challenges and data integrity concerns.

Production obligations come into play when parties need to produce electronically stored information as part of discovery requests. GPT-generated data and prompts may be subject to these obligations, necessitating careful consideration of relevance, proportionality and production format.

Conclusion

The integration of GPT tools into litigation practices requires a nuanced understanding of preservation and production obligations, and it calls for the development of best practices.

Legal professionals, businesses and individuals must adapt to the unique challenges presented by GPT-generated content and prompts while upholding the principles of fairness, transparency and compliance with legal obligations in federal civil litigation.

Rose J. Hunter Jones is a partner and the director of the e-discovery project management and technology practice at King & Spalding LLP.

Kassi R. Burns is a senior attorney at the firm.

Meredith A. Perlman is an attorney at the firm.