



PERSPECTIVES

Next Stop on the AI Train: eDiscovery

Our perspectives feature the viewpoints of our subject matter experts on current topics and emerging trends.

The material in this article was researched, compiled, and written by J.S. Held. It was originally published in The Legal Technologist in May 2024.

INTRODUCTION

Most everyone with whom we speak in the legal profession believes Artificial Intelligence (AI) will transform how we work and the solutions we offer for our clients. We often hear: “If you’re not on board the AI train, you will get run over by it.” But which way is that train heading and how do we ride it to the right destination?”

The Grand Conductors

Microsoft, Amazon, and Google have thus far played the role of conductors in the world’s AI journey. Simply put, very few companies have the resources to build out technology on the scale required to power machinery. Each of these companies have invested billions of dollars in the last year into Generative AI (GenAI) developers like OpenAI and Anthropic. The developers, in turn, have committed significant funds toward their investors’ cloud platforms, which power the AI’s operation.

Providing AI services, such as OpenAI’s ChatGPT, is also expensive. It takes intensive hardware, networking resources, and a skilled team of technologists to power the engine. It is estimated to cost up to \$21 million per month to keep ChatGPT online. As OpenAI CEO Sam Altman puts it, “the compute costs are eye-watering.”

The costs for providing these Large Language Model (LLM) services are recouped largely by charging consumers by the token, where a token is 4 English characters (100 tokens is roughly 75 words). With all of the research & development spending, the per token price also has increased in the latest generation of technology. For example, GPT-4 tokens cost four times as much as tokens created by its predecessor. But with the additional costs have come advancements that have caught the eye of the legal technology industry.

eDiscovery Full Steam Ahead

Although there are many use cases for AI in legal

technology – legal research, contract management, writing assistance, and more – one of the most exciting applications is in the area of ediscovery. Deemed to be among the most expensive parts of the legal process, (a famous Rand study 10 years ago put the cost of producing data at \$18,000 per GB) it has the largest potential for cost efficiencies by replacing human effort with AI computing. And while technology-assisted review (TAR) has been used for over a decade to help find relevant documents faster and with more accuracy, newer GenAI technologies have recently advanced to contribute even more to the discovery process. In 2022, GPT-3.5 – OpenAI’s state-of-the-art GenAI – failed the bar exam finishing in the 10th percentile. A year later, the better-trained GPT 4 passed the bar exam finishing in the 90th percentile. Now armed with its “JD,” GenAI has the potential to be a valuable asset to a legal team.

As a result, ediscovery software companies have begun to build the new and improved GenAI into their platforms. Some have used the technology to summarize longer documents into a paragraph or two, allowing reviewers to get a sense of each document quickly before deciding if further attention is warranted. Others have built legal assistant-type chatbots into the user interface, which train themselves on all of the searchable information within the database. Reviewers can ask the chatbot questions in plain English without having to learn technical search syntax. In seconds reviewers receive not only summarized answers, but also a reference to all the documents in the database used as sources for the answers. This use of GenAI will be a powerful tool in large investigations where the goal isn’t to find every relevant document, but to quickly understand the legal matter’s key issues.

However, the most ambitious implementation of GenAI in ediscovery software is the creation of **review bots** – AI assigned to seek out and identify all documents responsive to an issue, or all documents that need to be flagged for privilege. At its full potential, review bots will change the paradigm of ediscovery, with a single piece of software replacing teams of 50 or more contract attorneys billing hourly for months on end for a large matter. This advancement is being rolled out for widespread use in 2024.

As with any new technology, it won’t get it right at the beginning. Users will have to effectively craft the right

prompts, validate results to avoid misinterpretations and hallucinations, and iteratively refine the input provided to the review bots. But the real challenge comes as each iteration incurs another round of costs. (Remember, new tokens being requested equal new costs). Given current pricing dynamics the AI review bots theoretically would be cheaper than employing a large team of contract attorneys, if the bots get things right on the first attempt. After several rounds of training and refinement, the pendulum swings back toward human review as most cost-effective.

A Grand Age Of Exploration

While AI cost concerns are currently a major factor for legal teams, I don't see them as a permanent roadblock. Technology costs will come down. Remember when we measured the processing and hosting of data in the thousands of dollars per GB? Now we talk in single digits. Already in 2024, OpenAI has announced reducing per-token pricing to compete with Anthropic and Google Gemini AI platforms. The competition will foster faster innovation and better economics across a wide range of legal technologies. We humans can now put theory into practice, testing GenAI in real matters against actual data.

Other AI developments will accelerate growth and adoption as well. Open source frameworks and development toolkits are available for different use cases. Companies are building their own enterprise LLMs to make data-driven decisions in every aspect of their business. Legal teams are training models based on years of case data to *get smart* on specific types of investigations (e.g. anti-money laundering or healthcare fraud).

CONCLUSION

While AI models should be built and used with ethics as the primary consideration, our ability to share, refine, and grow these models will have a tremendous impact on how legal teams approach data. We will all have a part to play in keeping the train on track.

Acknowledgments

We would like to thank [Mike Gaudet](#) for providing insight and expertise that greatly assisted this research.

[Mike Gaudet](#) is a Managing Director in J.S. Held's [Digital Investigations & Discovery](#) group within the [Global Investigations Practice](#). He has more than 20 years of experience providing solutions for corporations, legal teams, and government agencies related to data discovery and governance challenges. He is an expert eDiscovery practitioner and technologist, with a master's in computer science. He has proficiency in leveraging the right tools to quickly gain insight from data, and to efficiently achieve project goals on time and under-budget. He has experience executing ad-hoc projects as well as designing and implementing Software-as-a-Service (SaaS) solutions.

Mike can be reached at mike.gaudet@jsheld.com or +1 281 415 5742.

This publication is for educational and general information purposes only. It may contain errors and is provided as is. It is not intended as specific advice, legal or otherwise. Opinions and views are not necessarily those of J.S. Held or its affiliates and it should not be presumed that J.S. Held subscribes to any particular method, interpretation or analysis merely because it appears in this publication. We disclaim any representation and/or warranty regarding the accuracy, timeliness, quality, or applicability of any of the contents. You should not act, or fail to act, in reliance on this publication and we disclaim all liability in respect to such actions or failure to act. We assume no responsibility for information contained in this publication and disclaim all liability and damages in respect to such information. This publication is not a substitute for competent legal advice. The content herein may be updated or otherwise modified without notice.