

The Art of AI: Using Large Language Models in Economic Consulting

Lisa Pinheiro and **Jimmy Royer** — 6 MIN READ

When using generative AI to help experts analyze massive amounts of text, correctly focusing and interpreting a large language model (LLM) requires a combination of data science and domain knowledge.

If you're not a data science expert, it's easy to think of generative AI as a "black box." It appears to the layperson that there's no way of knowing how a generative AI model comes up with its answers – you plug in your questions, and the model mysteriously spits out answers.

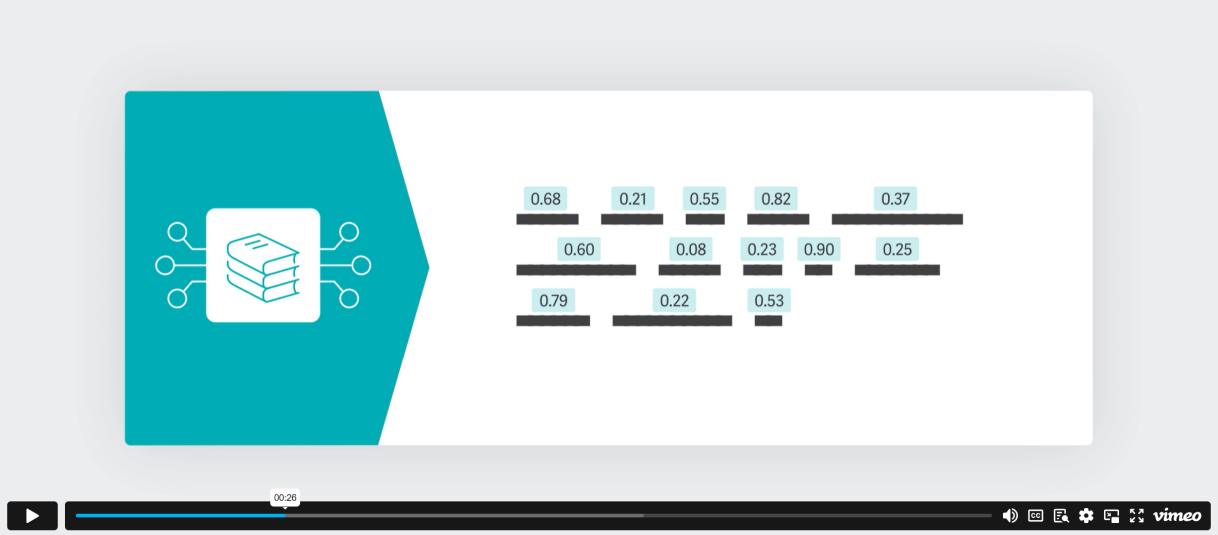
Fortunately, this perception is not accurate. Our economic and data science experts understand how generative AI and LLMs work and, importantly, how to use them correctly. In litigation, this understanding is crucial for knowing how to extract and use AI-generated information that is relevant to your case.

Making LLMs work for you

At the most basic level, LLMs can be viewed as a type of AI used to generate summaries from massive amounts of information – hence the term "generative AI."

A common misconception about LLMs, however, is that they are black boxes, making it impossible to understand how they work. This misconception can be dangerous if it leads to the assumption that an LLM can "think" like a human, resulting in its outputs being accepted without critical examination.

In reality, LLMs are best managed by teams of experts who combine a deep understanding of the statistical nature of the model's operation with deep domain, or subject matter, knowledge. This expertise can support and enhance the quality of LLM-assisted research by updating and focusing inputs, and accurately and critically interpreting the output.



Using a retrieval-augmented generation (RAG) or similar process to lower the number of documents an LLM considers can help generate a more detailed and cohesive summary. However, it can also increase the risk of leaving out important information. Depending on the desired trade-off between cohesiveness (and level of detail) and exhaustiveness for a given task, this type of calibration needs to be adjusted by data scientists working in conjunction with domain experts and, in the case of litigation, a legal team.

There's more to the process than curation of source material

Another important part of the process for focusing the LLM on the right data is writing an effective and appropriate prompt based on domain knowledge. In practice, this can mean having a prompt that is multiple pages long, but one that provides significant context to help the algorithm focus on the most relevant information.

In this type of targeted approach, the interaction with the model is not an open-ended conversation between a person and a machine, as is sometimes represented in the general media or in social contexts. Instead, in a targeted data aggregation process, data scientists and domain experts apply their technical and subject matter knowledge to instruct the LLM and calibrate it to ensure that the right words are used for the right purpose.

On the back end, the process of information extraction similarly relies on expert analysis and domain knowledge. Here, the use of an RAG process allows the experts to validate and, if necessary, correct the model's output. The experts may critically examine the output by testing it against a series of questions such as:

- Is the output current? Does it reflect the most recent data available from public and (potentially) private sources?
- Does the output align with what we already know about the topic (domain knowledge)?
- Can we identify the sources the LLM used to generate the output?

How might an LLM be used in litigation?

Properly managed LLMs incorporate and expand on previous natural language processing (NLP) approaches, such as:

- Topic analysis, which is used to identify key themes in large amounts of text – often as a first stage for classifying documents within those themes – and weed out irrelevant sources (essentially, sorting signal from noise)
- Sentiment analysis, which tells you whether the information is expressed or viewed either positively or negatively, often by comparing word use with domain-specific dictionaries that have been compiled for this use

LLMs are also used to identify relevant information or documents in datasets that are too large to review manually – often hundreds of thousands or millions of entries.

These capabilities allow LLMs to be used effectively in a variety of types of litigation. For example, an LLM might be used to assist in aggregating large amounts of historical corporate statements and news reports in a way that can help contextualize or even quantify the impact of specific statements on share prices. An LLM can support the analysis of the statements and reports in the context of everything else that came before, after, or simultaneously across a broad range of relevant sources.

Similarly, an LLM applied to large amounts of customer reviews over time can help quantify hard-to-measure features such as product quality or customer satisfaction along various dimensions. This capability can then help answer specific questions such as how much value is generated by the specific quality improvements of products, and how much should be attributed to other factors such as marketing.

Using generative AI as a complement to expert analysis, not a replacement for it

It's not unusual to hear about LLMs "writing" essays or articles, when in fact they are producing lengthy and detailed responses to prompts they are given, based on statistical analysis of millions of pre-existing sources. In reality, LLMs are neither more nor less than powerful tools that, when employed skillfully and calibrated appropriately, can be extremely useful to experts when analyzing vast amounts of information.

In recent presentations, representatives from the US Securities and Exchange Commission have put corporate executives and their counsel on notice that generative AI cannot be used as a scapegoat. Companies will be held liable for the content they publish and trades they make, for example, even if recommended by generative AI. Therefore, it may become crucial for companies to be able not only to understand what content an LLM produces, but also to explain the processes and models used to generate that content and ensure its accuracy. The right combination of technical knowledge and domain expertise can be employed to improve the curation and calibration of sources, the efficacy of prompts, and the quality of output. ■