

# Demystifying Artificial Intelligence, ChatGPT, and Large Language Models in the Legal Industry

Educational Webcast

10 | 25 | 2023

© 2023 Winston & Strawn LLP

WINSTON  
& STRAWN  
LLP

eDISCOVERY  
& INFORMATION  
GOVERNANCE

HAYSTACK®

# John Rosenthal

## ▶ WINSTON & STRAWN LLP – WASHINGTON, D.C.

- Antitrust and commercial litigator and trial lawyer.
- Substantial MDL and complex litigation experience.
- Chair, Winston eDiscovery & Information Governance Practice.
- National eDiscovery counsel for several Fortune 500 companies.

## ▶ THE SEDONA CONFERENCES

- Working Group 1 on Electronic Discovery (former steering committee member).
- Working Group 6 on International Privacy.
- Working Group 11 on Data Security.

## ▶ ADVISORY BOARD – GEORGETOWN LAW SCHOOL EDISCOVERY INSTITUTE

## ▶ D.C. BAR – E-DISCOVERY COMMITTEE

## ▶ CO-CHAIR FJC – JUDICIAL TRAINING PROGRAM (2017–2019)

## ▶ EDITOR-IN-CHIEF – EDISCOVERY ADVANTAGE

## ▶ AUTHOR OR CONTRIBUTOR TO SEMINAL WRITINGS IN THE FIELD, INCLUDING:

- Editor, Federal Judge’s Guide to Discovery (3<sup>rd</sup> ed.).
- Duke Conference’s Guidelines and Practices for Implementing the 2015 Discovery Amendments.
- Numerous Sedona Conference Commentaries (Principles).

WINSTON  
& STRAWN  
LLP

eDISCOVERY  
& INFORMATION  
GOVERNANCE



WINSTON  
LEGAL  
SOLUTIONS

“John Rosenthal is one of the best e-discovery lawyers in the US. One of his great strengths is how good he is in the courtroom and before judges about e-discovery matters.”

## LEGAL 500 US



# Ashish Prasad

- Vice President and General Counsel, HaystackID.
- Served, among other things, as a litigation partner, founder, and Chair of the Mayer Brown LLP Electronic Discovery and Records Management Group.
- Adjunct Professor of Law at Northwestern University Law School.
- Executive Editor of The Sedona Principles and Board Member of the Electronic Discovery Institute.
- Ashish has edited three treatises, authored more than two dozen articles, and given hundreds of continuing legal education seminars on topics of electronic discovery and complex litigation before judges, practicing lawyers, and industry groups in the United States, Europe, and Asia.
- Ashish is a graduate of the University of Chicago Law School, where he was a member of the Law Review, and the University of Michigan, where he graduated with high honors and high distinction.

HAYSTACK®

---

Ashish Prasad is widely regarded as among the leading experts on discovery in the United States and has been recognized and ranked in the Chambers Litigation Support Guide for six consecutive years.

---

# Agenda

- Introduction to ChatGPT
- Application to the Legal Industry
- Ethical Issues Around Its Use
- Guidance on Its Adoption
- Questions



# Flavors of AI in eDiscovery Until Now

## 31 Flavors of AI *Data Discovery*?

- Sentiment Analysis
- Portable AI Models
- Concept Clustering
- Anomaly Detection
- Custom AI Models
- AI Model Library



- Social Network Analysis
- TAR/CAL
- Computer Vision
- AI Translation
- AI Transcription
- Entity Extraction

# Arrival of Large Language Artificial Intelligence

TECH · A.I.

**Bill Gates predicts A.I. will change the world more and faster than his personal computing revolution**

## *In A.I. Race, Microsoft and Google Choose Speed Over Caution*

Technology companies were once leery of what some artificial intelligence could do. Now the priority is winning control of the industry's next big thing.

## **Now Google to shove its answer to ChatGPT into Gmail, Docs, apps via APIs, more**

PaLM springs forth

 [Katyanna Quach](#)

Tue 14 Mar 2023 | 21:25 UTC

Google has promised to offer API-level access to its large language model PaLM so that developers can build it into their apps and workflows, and thus make the ChatGPT-like text-emitting tech available to world-plus-dog.

**Letter signed by Elon Musk demanding AI research pause sparks controversy**

## **ChatGPT banned in Italy over privacy concerns**

6 days ago



GETTY IMAGES

OpenAI launched ChatGPT last November

WINSTON  
& STRAWN  
LLP

eDISCOVERY  
& INFORMATION  
GOVERNANCE

HAYSTACK

# Large Language Models and the Legal Industry

BROOKINGS



**A**rtificial intelligence (AI) is poised to fundamentally reshape the practice of law. While there is a long history of technology-driven changes in how attorneys work, the recent introduction of large language model-based systems such as GPT-3 and [GPT-4](#) marks the first time that widely available technology can perform sophisticated writing and research tasks with a proficiency that previously required highly trained people.

TECHNOLOGY

**New GPT-4 Passes All Sections Of The Uniform Bar Exam.  
Maybe This Will Finally Kill The Bar Exam.**


# Opps

Disrupted

## New York lawyers sanctioned for using fake ChatGPT cases in legal brief

By Sara Merken

June 26, 2023 4:28 AM EDT · Updated 4 months ago




A response by ChatGPT, an AI chatbot developed by OpenAI, is seen on its website in this illustration picture taken February 9, 2023. REUTERS/Florence Lo/Illustration [Acquire Licensing Rights](#)

HOME / MUSIC / BREAKING NEWS

## Fugees Rapper Pras Michel Claims Lawyer's Use Of AI Bungled Closing Arguments In Trial

By Bruce Haring

October 18, 2023 8:20pm



ADVERTISEMENT

Trending on Deadline



# Judicial Reaction – Court Mandated Disclosures


---

- **ND Tex** (Judge Brantley Starr) **Mandatory Certification Regarding Generative Artificial Intelligence**. All attorneys and pro se litigants appearing before the Court must, together with their notice of appearance, file on the docket a certificate attesting either that no portion of any filing will be drafted by generative artificial intelligence (such as ChatGPT, Harvey.AI, or Google Bard) or that any language drafted by generative artificial intelligence will be checked for accuracy, using print reporters or traditional legal databases, by a human being. Accordingly, the Court will strike any filing from a party who fails to file a certificate on the docket attesting that they have read the Court's judge-specific requirements and understand that they will be held responsible under Rule 11 for the contents of any filing that they sign and submit to the Court, regardless of whether generative artificial intelligence drafted any portion of that filing. Judge Brantley Starr, Judge Specific Requirements, *Mandatory Certification Regarding Generative Artificial Intelligence* (May 30, 2023).
- **Order on Artificial Intelligence**, U.S. Court of International, The Honorable Stephen Alexander Valden, Judge (June 8, 2023) U.S. Court Int. Trade (Judge Vaden) ORDERED that any submission in a case assigned to Judge Vaden that contains text drafted with the assistance of a generative artificial intelligence program on the basis of natural language prompts, including but not limited to ChatGPT and Google Bard, must be accompanied by: (1) A disclosure notice that identifies the program used and the specific portions of text that have been so drafted; (2) A certification that the use of such program has not resulted in the disclosure of any confidential or business proprietary information to any unauthorized party.
- **Standing Order for Civil Cases Before Magistrate Judge Fuentes**, U.S. District Court for the Northern District of Illinois (May 31, 2023), The Court has adopted a new requirement in the fast-growing and fast-changing area of generative artificial intelligence (“AI”) and its use in the practice of law. The requirement is as follows: Any party using any generative AI tool to conduct legal research or to draft documents for filing with the Court must disclose in the filing that AI was used, with the disclosure including the specific AI tool and the manner in which it was used.
- **Standing Order Re: Artificial Intelligence (“AI”) Cases Assigned to Judge Bayleson**, U.S. District Court for the Eastern District of Pennsylvania (June 6, 2023) Judge Michael M. Bayleson of the U.S. District Court for the Eastern District of Pennsylvania issued an order requiring attorneys and pro-se litigants to disclose the use of AI in drafting pleadings. His order, however, was not limited in scope to the use of GenAI tools; rather, it referenced AI tools in general.


# A Little Geek Speak

---


**Machine learning (ML)** is a branch of artificial intelligence (AI) that focuses on teaching machines to recognize patterns in data and make predictions or decisions based on those patterns. It involves feeding large amounts of data into a computerized algorithm, which then uses statistical techniques to identify patterns and make predictions or decisions about new data.



**Large Language Model (LLM)** is where an algorithm has been trained on a large amount of text-based data, typically scraped from the open internet, and so covers web pages and - depending on the LLM - other sources such as scientific research, books or social media posts.



**Generative AI** refers to a category of LLM that generates new outputs based on the patterns within data on which they have been trained. Like other forms of artificial intelligence, generative AI learns how to take actions based upon past data, but it also creates brand new content based on that training (such as images, videos, text, or computer code) rather than simply categorizing or identifying data.



**Natural Language Chatbot** is a software program that can understand and respond to human speech.

# Examples of Models



Source 2023 – EDI Summit

# ChatGPT

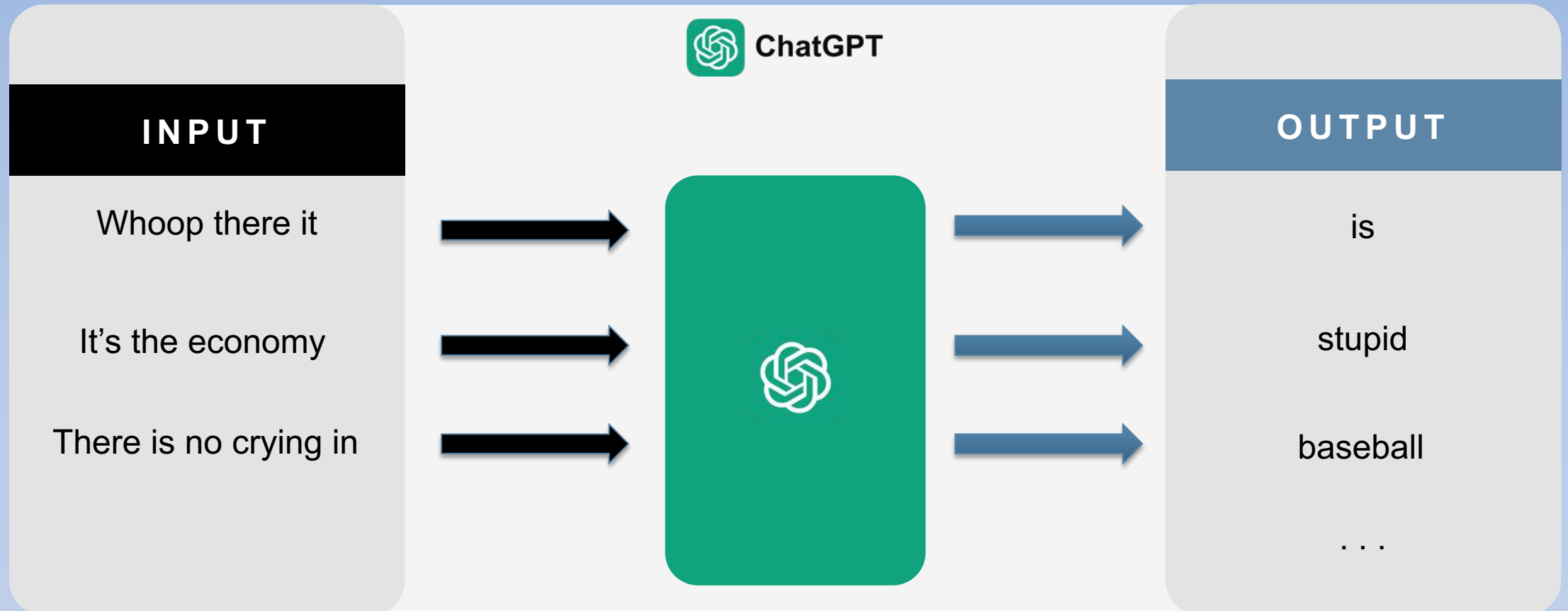
---



- **ChatGPT** is a Large Language Model designed to assist with generating human-like text in a chatbot or conversational setting.
- It has been **trained on a large dataset** of chat and conversation transcripts and is able to **generate responses** that are **appropriate and relevant** to the input it receives.

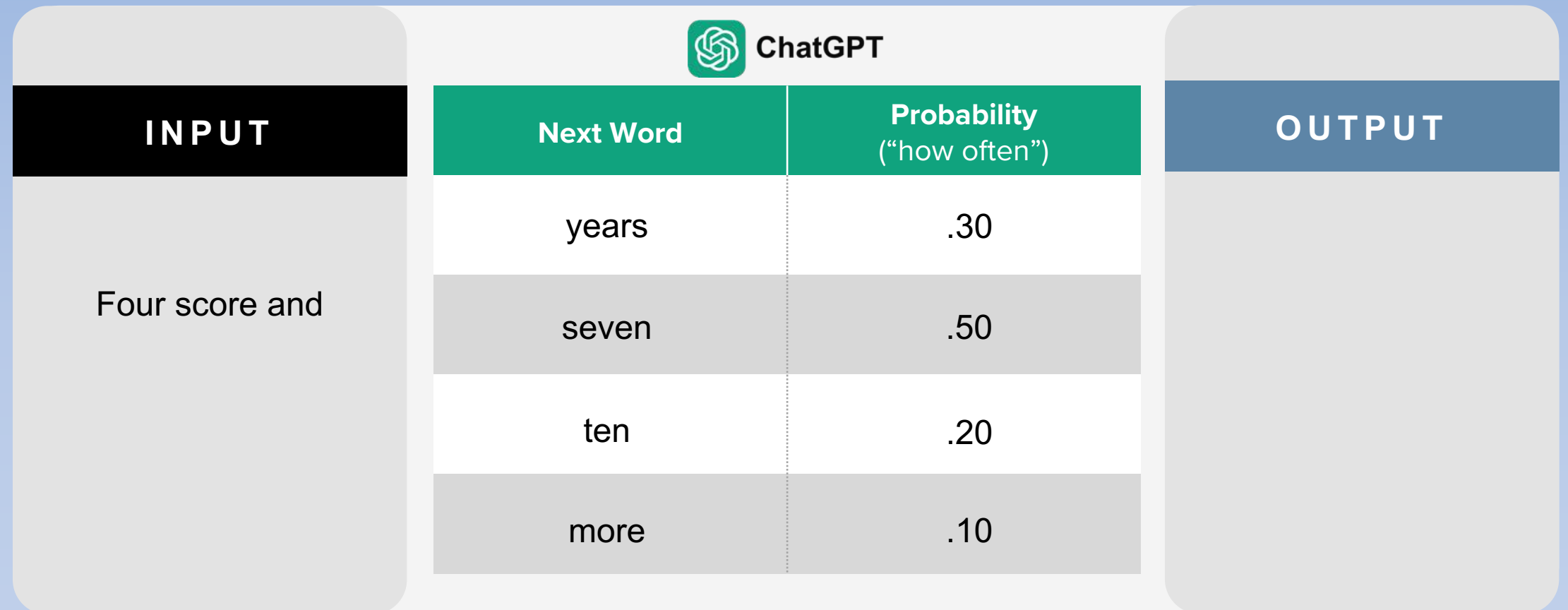
# How It Works

Predict the *next* word in a sentence using the previous words.



# How It Works

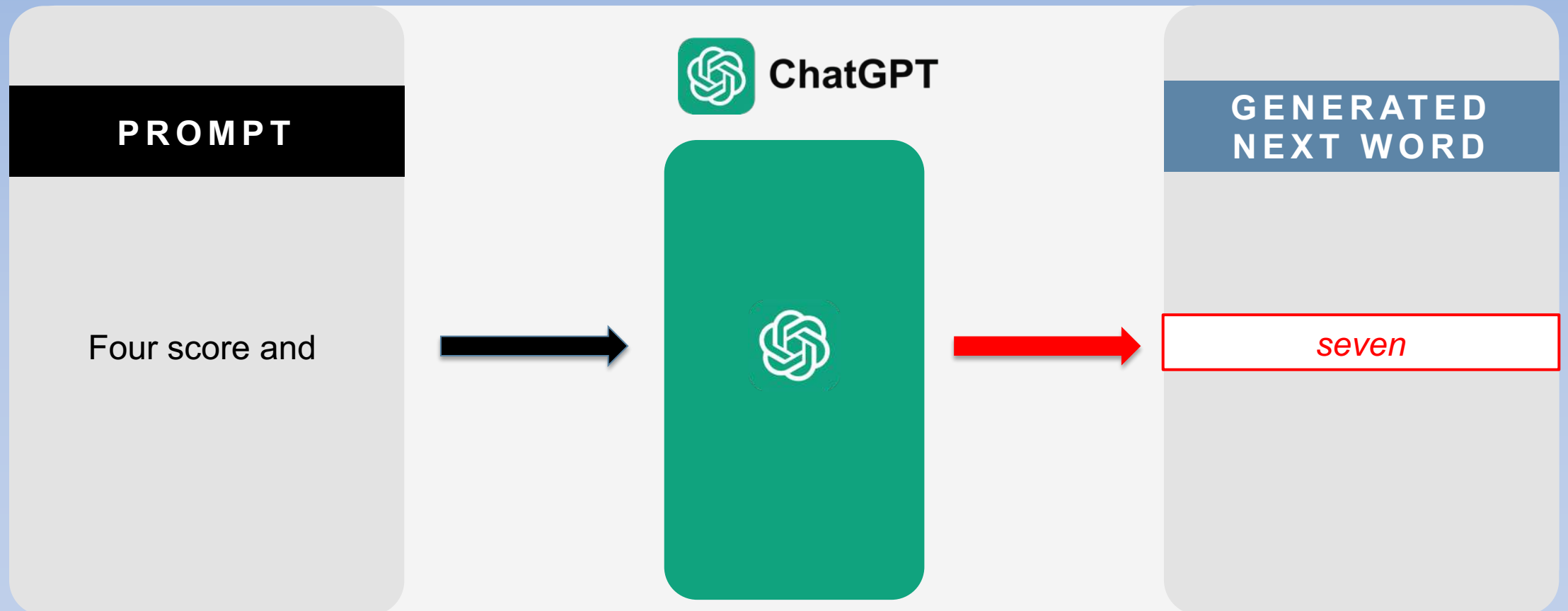
It calculates a probability table.



# How It Works

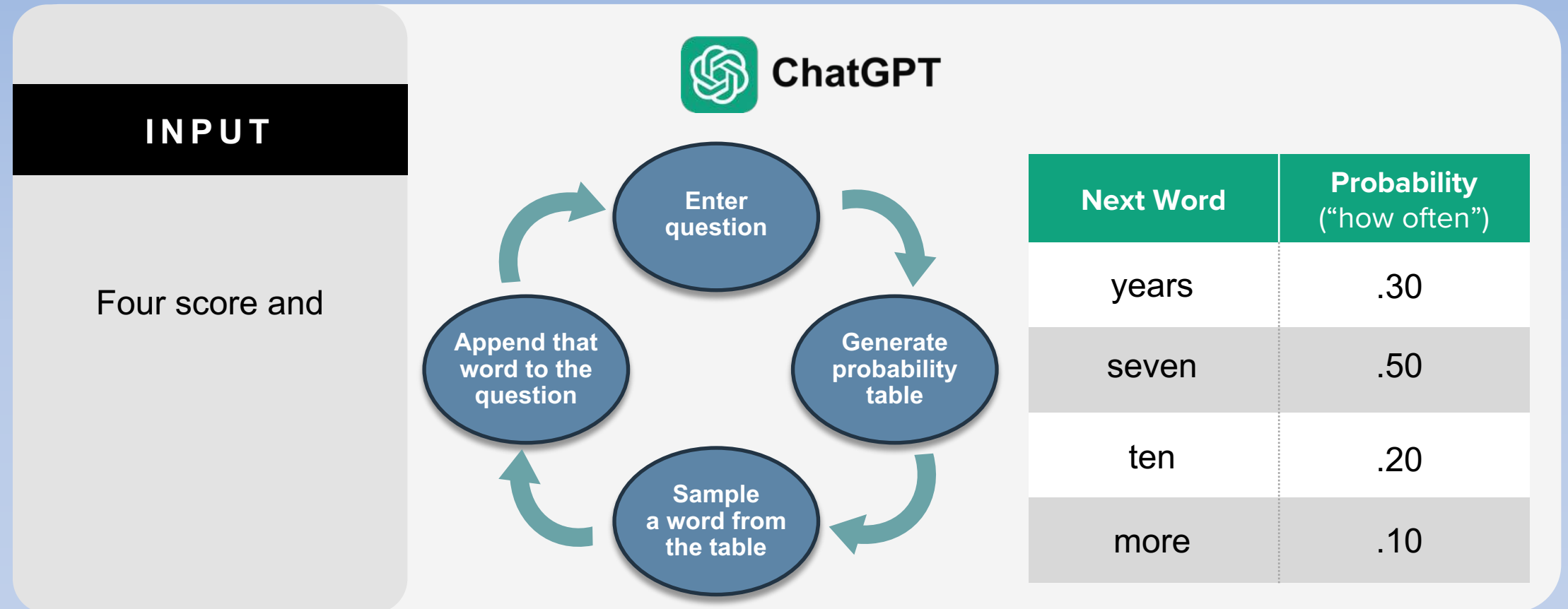
---

It calculates a table and picks a word from it.



# How It Works

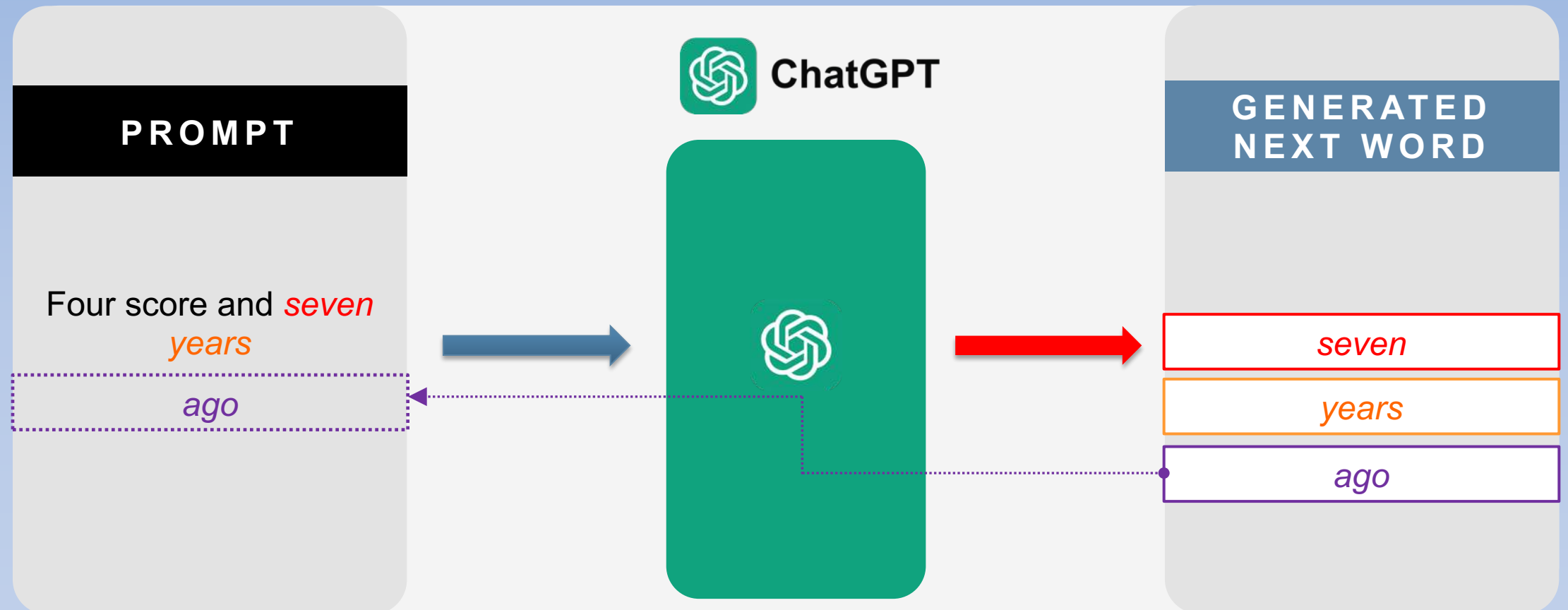
Run a loop 1000s of times to predict an answer.





# How It Works

---



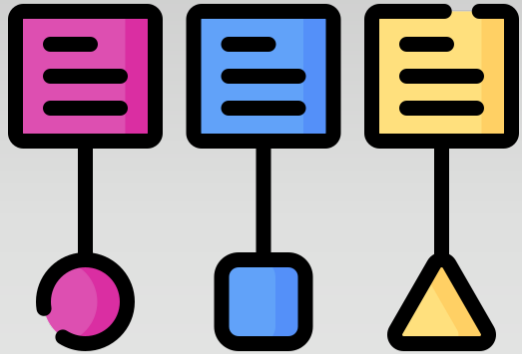
# Training

---

- Provided human-based answers to refine its learning (“Instruction Tuning”).
- Have GPT provide several answers and have humans rank the results.
- What is it trained against?
  - Public models trained against the entire internet.
  - Private models are trained against specific data sets.

# Large Language Models are Trained to Solve Common Language Problems

---



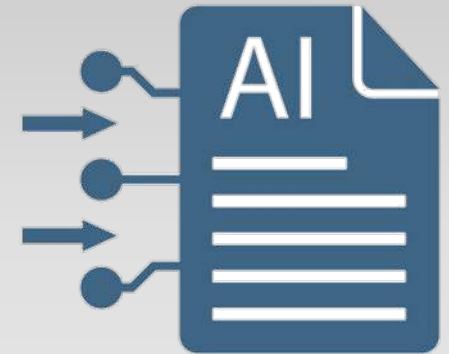
**Text  
Classification**



**Question  
Answering**



**Document  
Summarization**



**Text  
Generation**

# Practical Example – Microsoft Teams

Focus on what matters with AI-powered meetings, including ChatGPT from OpenAI.

The screenshot shows a Microsoft Teams meeting interface for a meeting titled "Sales Analysis Review". The meeting content pane is open, showing "Meeting content" with items like "Sales report Q4...", "Capacity stats list...", and "VanArsdePitktDe...". The "Notes" section is active, displaying "AI Notes" and "Mentions". A "Suggested notes" section is highlighted, listing three notes with timestamps: 5:00, 5:05, and 5:34. Below the notes is a "Suggested tasks" section with two tasks: "Beth will follow up on advertisement budget." and "Jon Shammass will double check with Amanda regarding the ETA for the release of the new product."

**Suggested notes**

- ▶ **Serena** wants to look at the sales report before she and **Beth** spend more budget on the campaign. 5:00
- ▶ **Beth** explains that they are on track for new product release in December. But they will need to keep an eye on advertisement budget. She will follow up on that. 5:05
- ▶ **Beth** explains that they are managing the capacity well. They could be a problem if they get more popular. **Babak** is going to double check on that. 5:34

**Suggested tasks**

- ▶ Beth will follow up on advertisement budget.
- ▶ **Jon Shammass** will double check with **Amanda** regarding the ETA for the release of the new product.

# Problems

---

- It gets things wrong (guardrails).
- It makes things up (hallucinations).
- Cannot understand emotion and empathy like a human.
- Bias (data, style, historical, user).
- Open Questions:
  - Copyright
  - Privacy
  - Defamation
  - Consumer Protection – Advertising

# Legal-Specific Use Cases

WINSTON  
& STRAWN  
LLP

eDISCOVERY  
& INFORMATION  
GOVERNANCE

HAYSTACK®

# Legal-Specific Use Cases for Legal Practice

---

**Generate correspondence:** Draft any correspondence, including demand letters and litigation holds.

**Generate timelines:** Quickly analyze documents and prepare a chronology of key events.

**Simplify:** Upload complex documents and receive succinct summaries.

**Review documents:** Get answers to questions about the contents of documents, including evaluating documents for privilege.

**Draft discovery responses:** Draft interrogatories, RFAs, and RFPs to propound in any case.

**Brief a case:** Generate a case brief and answer key questions about that case.

**Answer a complaint:** Admit or deny allegations in a complaint and recommend affirmative defenses and objections.

**Respond to discovery:** Prepare responses and objections to interrogatories, RFAs, and RFPs.

**Legal research:** Prepare a memo answering a legal research question.

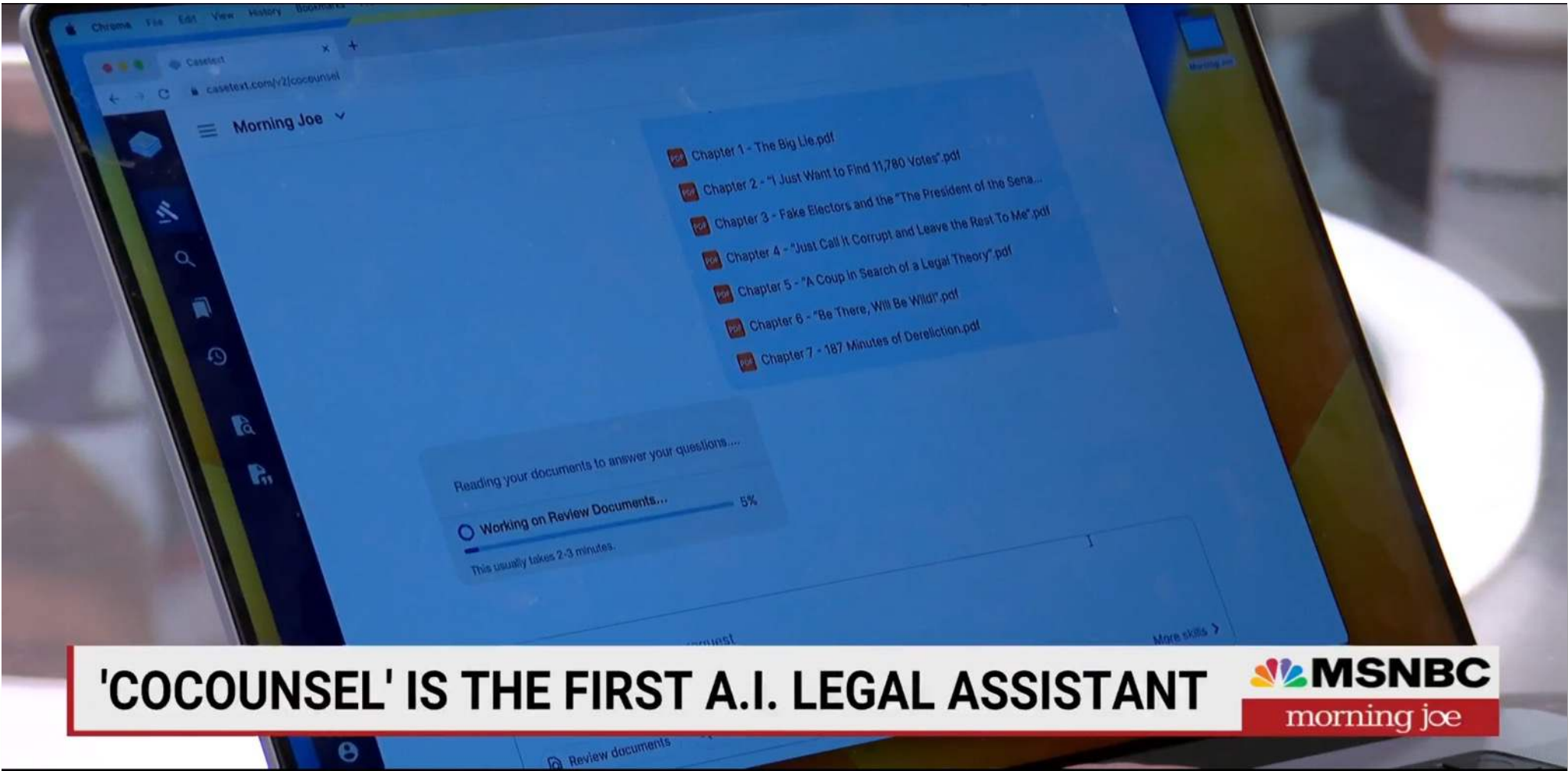
**Prepare for a deposition:** Generate an outline of deposition questions.

**Extract contract data:** Analyze contracts and answer your questions about contractual provisions.

**Analyze a policy:** Upload a policy document to see if it conflicts with other contracts.

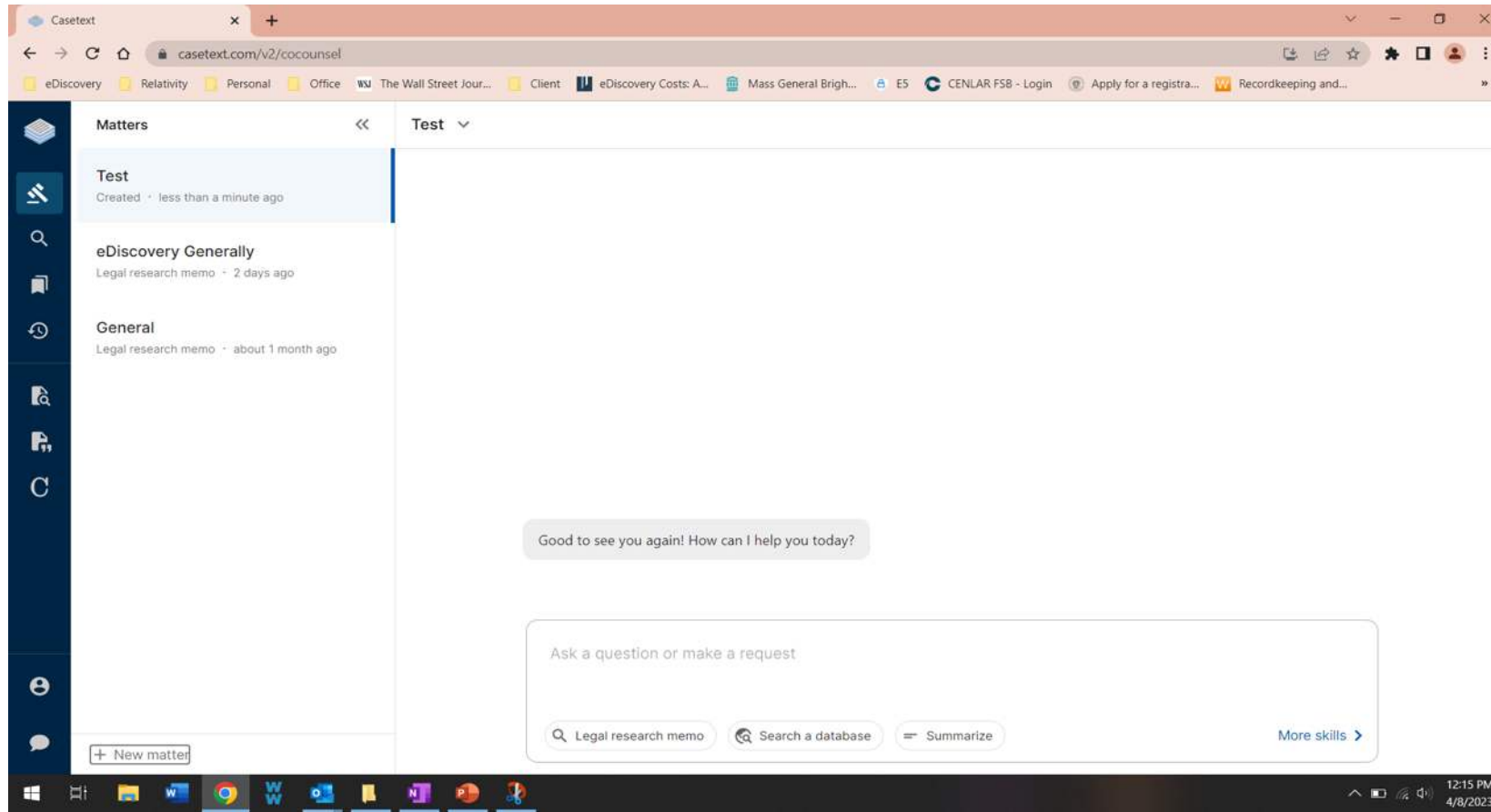
**What is market:** Assess “what is market” in a given set of contractual clauses.

**Edit a contract clause:** Get recommended edits to a contract clause.

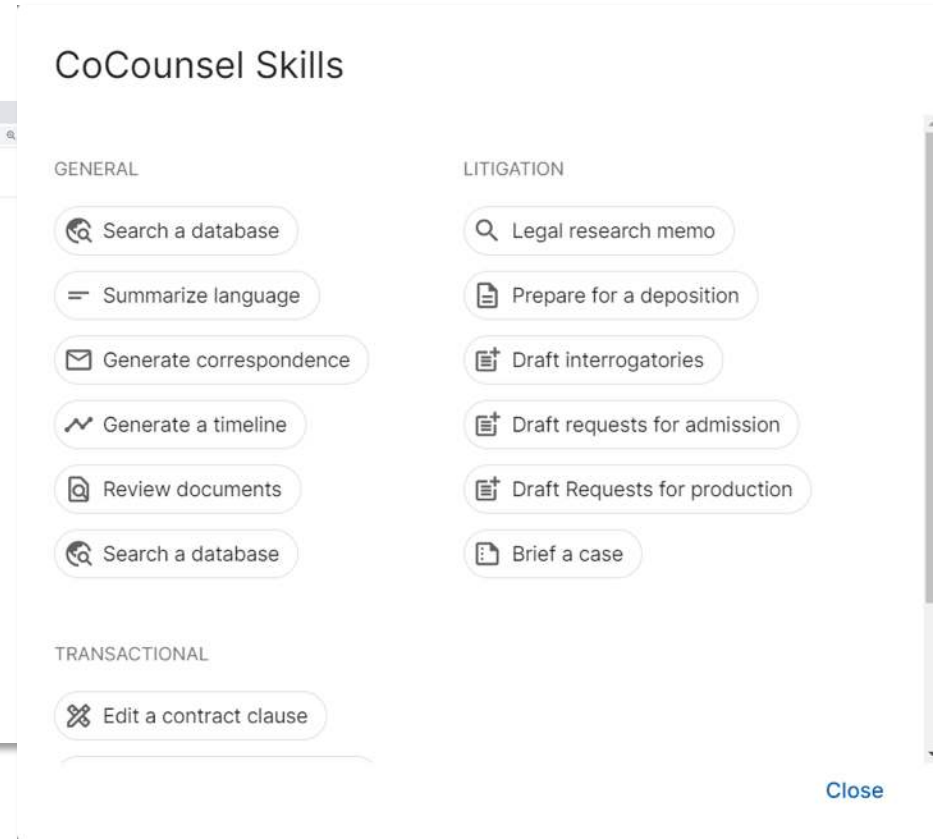
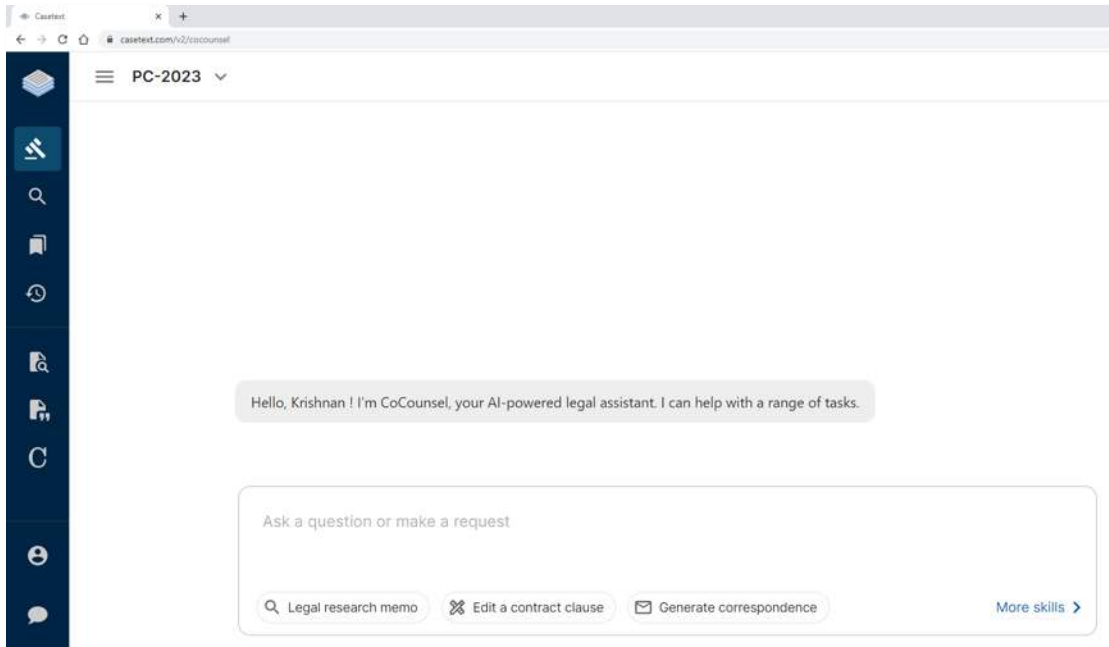




# Interface



# Skills

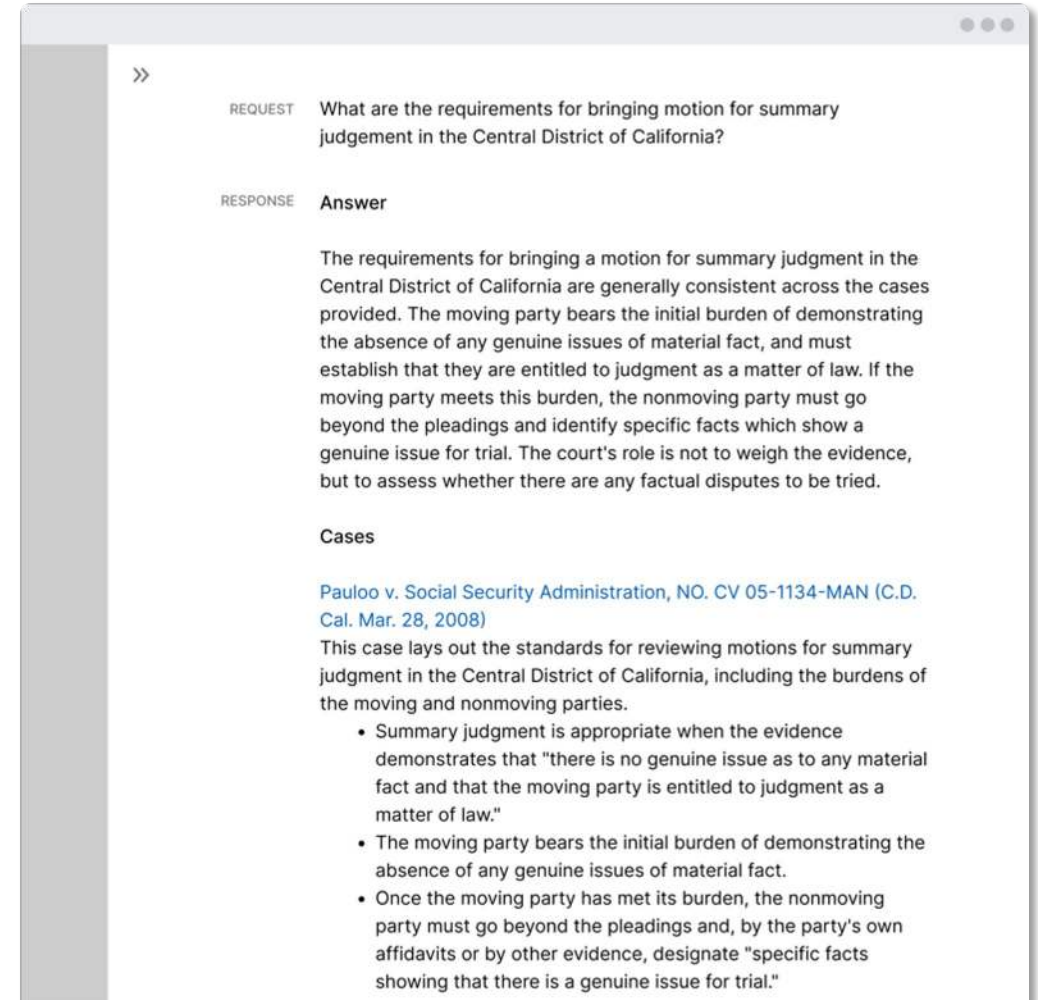


# Research Memos

---

Search and synthesize primary law.

It runs multiple searches simultaneously and then summarizes its findings in a memo, which includes links to the top cases as well as the most relevant quotes.



>>

REQUEST What are the requirements for bringing motion for summary judgement in the Central District of California?

RESPONSE **Answer**

The requirements for bringing a motion for summary judgment in the Central District of California are generally consistent across the cases provided. The moving party bears the initial burden of demonstrating the absence of any genuine issues of material fact, and must establish that they are entitled to judgment as a matter of law. If the moving party meets this burden, the nonmoving party must go beyond the pleadings and identify specific facts which show a genuine issue for trial. The court's role is not to weigh the evidence, but to assess whether there are any factual disputes to be tried.

**Cases**

[Pauloo v. Social Security Administration, NO. CV 05-1134-MAN \(C.D. Cal. Mar. 28, 2008\)](#)

This case lays out the standards for reviewing motions for summary judgment in the Central District of California, including the burdens of the moving and nonmoving parties.

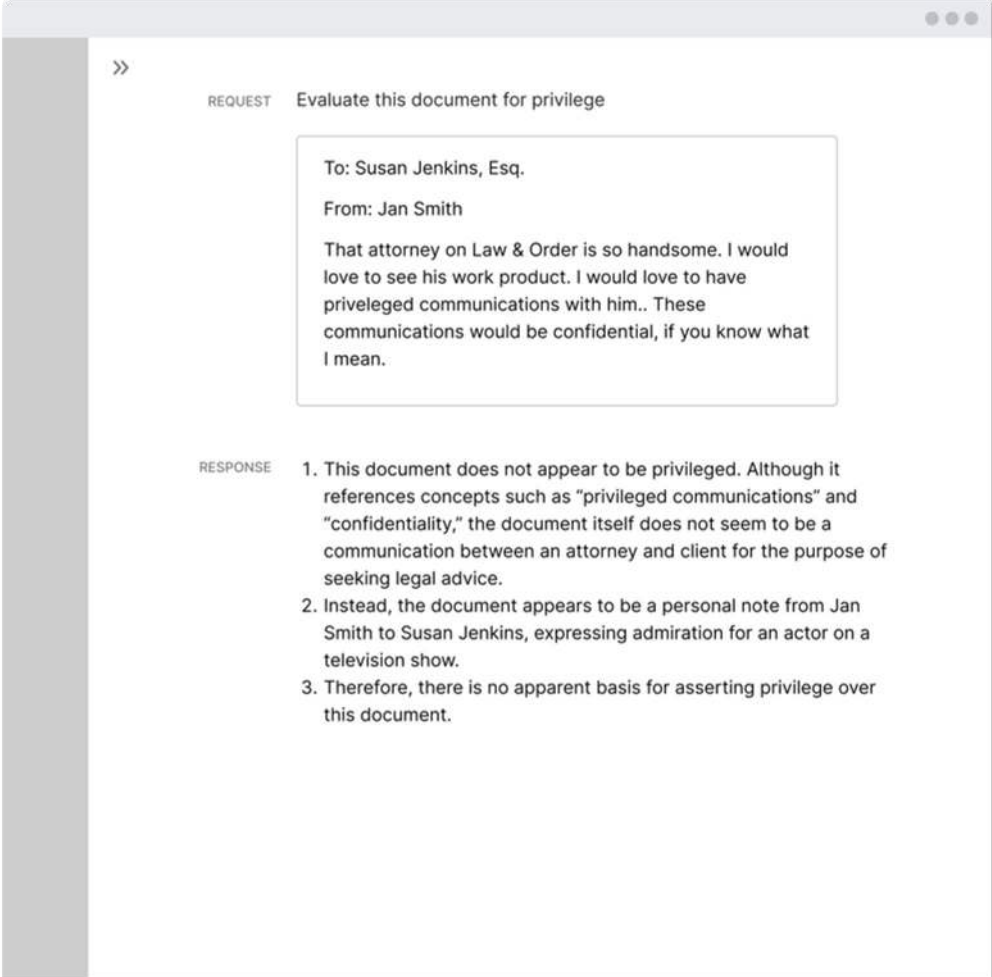
- Summary judgment is appropriate when the evidence demonstrates that "there is no genuine issue as to any material fact and that the moving party is entitled to judgment as a matter of law."
- The moving party bears the initial burden of demonstrating the absence of any genuine issues of material fact.
- Once the moving party has met its burden, the nonmoving party must go beyond the pleadings and, by the party's own affidavits or by other evidence, designate "specific facts showing that there is a genuine issue for trial."

# Document Review

---

Upload documents to ask questions about the contents of those documents.

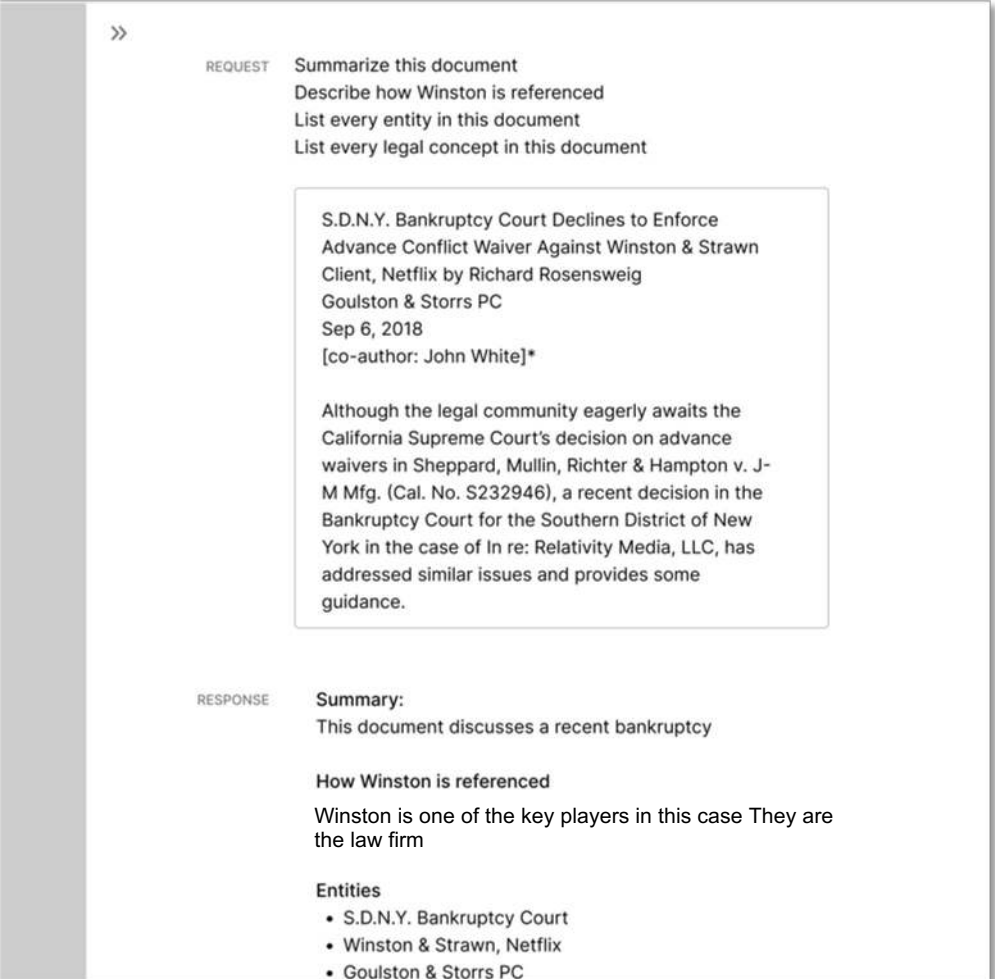
The ability to analyze documents for privilege.



The screenshot shows a user interface for document review. At the top, there is a navigation arrow and a 'REQUEST' section with the text 'Evaluate this document for privilege'. Below this is a text box containing an email snippet: 'To: Susan Jenkins, Esq.', 'From: Jan Smith', and a paragraph of text: 'That attorney on Law & Order is so handsome. I would love to see his work product. I would love to have priveleged communications with him.. These communications would be confidential, if you know what I mean.' Below the text box is a 'RESPONSE' section containing three numbered points: 1. This document does not appear to be privileged. Although it references concepts such as "privileged communications" and "confidentiality," the document itself does not seem to be a communication between an attorney and client for the purpose of seeking legal advice. 2. Instead, the document appears to be a personal note from Jan Smith to Susan Jenkins, expressing admiration for an actor on a television show. 3. Therefore, there is no apparent basis for asserting privilege over this document.

# Document Review (Another Example)

Capabilities for abstractive summarization and entity recognition.



>>

**REQUEST**

- Summarize this document
- Describe how Winston is referenced
- List every entity in this document
- List every legal concept in this document

S.D.N.Y. Bankruptcy Court Declines to Enforce Advance Conflict Waiver Against Winston & Strawn Client, Netflix by Richard Rosensweig  
Goulston & Storrs PC  
Sep 6, 2018  
[co-author: John White]\*

Although the legal community eagerly awaits the California Supreme Court's decision on advance waivers in Sheppard, Mullin, Richter & Hampton v. J-M Mfg. (Cal. No. S232946), a recent decision in the Bankruptcy Court for the Southern District of New York in the case of In re: Relativity Media, LLC, has addressed similar issues and provides some guidance.

**RESPONSE**

**Summary:**  
This document discusses a recent bankruptcy

**How Winston is referenced**  
Winston is one of the key players in this case They are the law firm

**Entities**

- S.D.N.Y. Bankruptcy Court
- Winston & Strawn, Netflix
- Goulston & Storrs PC

# eDiscovery and LLMs

---



# Ethical Rules

WINSTON  
& STRAWN  
LLP

eDISCOVERY  
& INFORMATION  
GOVERNANCE

HAYSTACK®

# The Ethical Rules of Legal Technology

Rule 1.1 (Competence), Comment 8 provides that a lawyer should “keep abreast of changes in the law and its practice, including the benefits and risks associated with relevant technology.”

- Lawyers must “have a firm grasp on how electronic information is created, stored and retrieved” and must “know how to make and respond to electronic discovery requests and to advise their clients regarding electronic discovery obligations.”
- The Seventh Circuit Pilot Program for Electronic Discovery requires that counsel must be familiar with the federal rules of electronic discovery (Education Principle 3.01) and engage in continuing education regarding electronic discovery (Education Principle 3.02).
- 37 states have adopted a technical competence rule arising from Rule 1.
- Implications:
  - Level of competence you will require of your outside and corporate legal staff.
  - Do you require subject matter expertise in this area?



# The Ethical Rules of Legal Technology

---

Rule 2.1 states, “In representing a client, a lawyer shall exercise independent professional judgment and render candid advice,” which potentially can involve referring “not only to law but other considerations as moral, economic, social and political factors, that may be relevant to the client’s situation.”

- Rule 2.1 directly addresses the ethical duty of a lawyer to avoid the temptation of entirely relying on the output of AI legal software.

# The Ethical Rules of Legal Technology

---

Rule 1.6 (a) “A lawyer shall not reveal information relating to the representation of a client unless the client gives informed consent...”

Rule 1.6(c) (Confidentiality) provides that a lawyer must take reasonable efforts to prevent the “unauthorized or inadvertent disclosure” of information relating to representation of a client.

- Rules implicate:
  - How is privileged or trade secret information stored, retrieved, and retained?
  - How is your privileged and trade secret information used by the AI provider?
  - Do you have the ability to track and delete your privileged and trade secret information?

# Bar Position on Cloud Computing

---

- State Bar Professional Ethics Committees have advised that a lawyer may utilize cloud computing services provided that the lawyer “take[s] reasonable precautions to ensure that confidentiality of client information is maintained, that the service provider maintains adequate security, and that the lawyer has adequate access to the information stored remotely.”
- The lawyer must also research the cloud computing service prior to use and assure that files stored electronically are “readily reproducible and protected from inadvertent modification, degradation or destruction.” (Florida Bar Ethics Opinion 12-3).

# Confidentiality Risk is Real!

## Samsung workers made a major error by using ChatGPT

By Lewis Maddison published 3 days ago

Samsung meeting notes and new source code are now in the wild after being leaked in ChatGPT



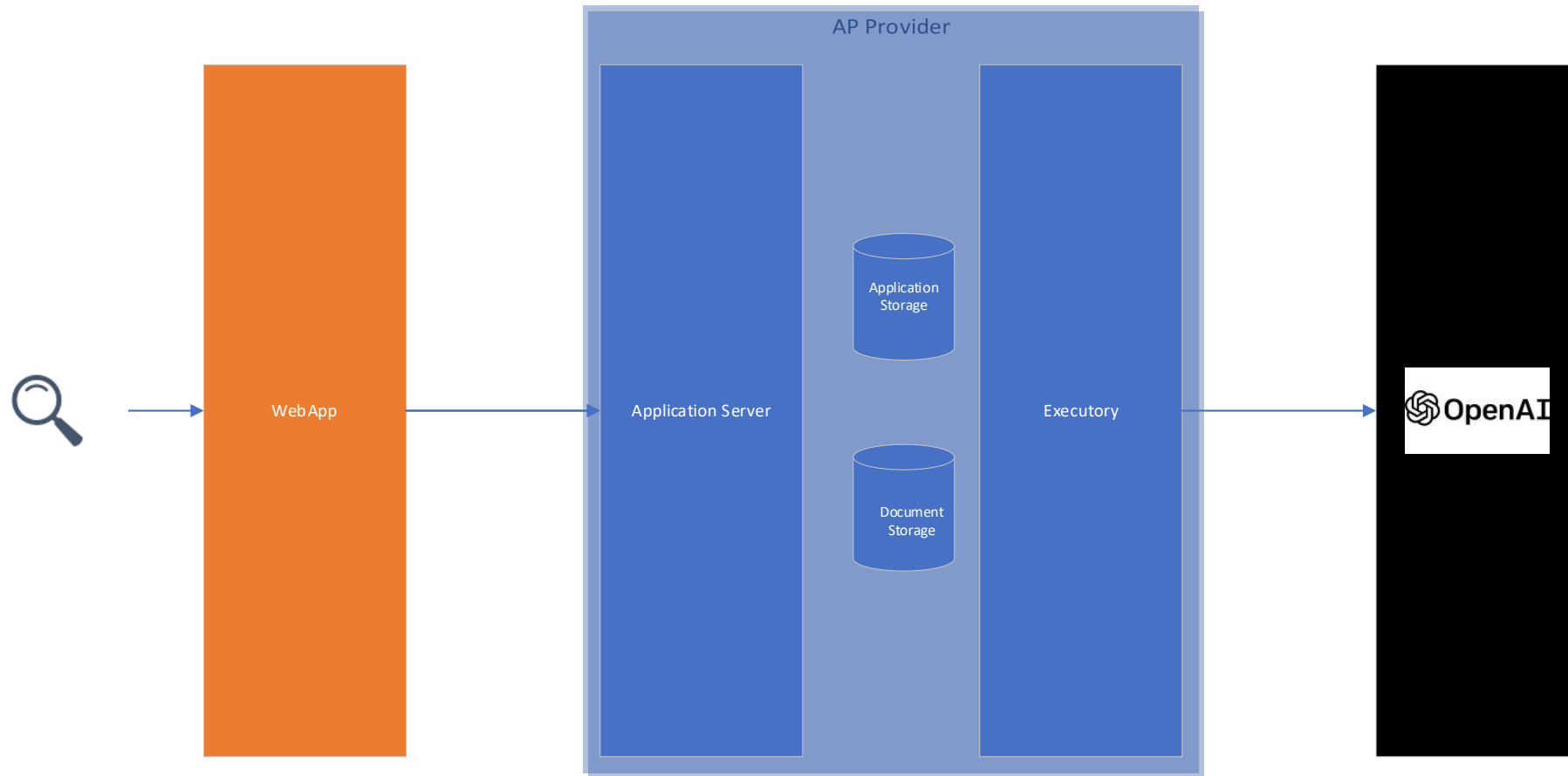
(image credit: Valeriya Zankovych / Shutterstock.com)

Samsung workers have unwittingly leaked top secret data whilst using ChatGPT to help them with tasks.

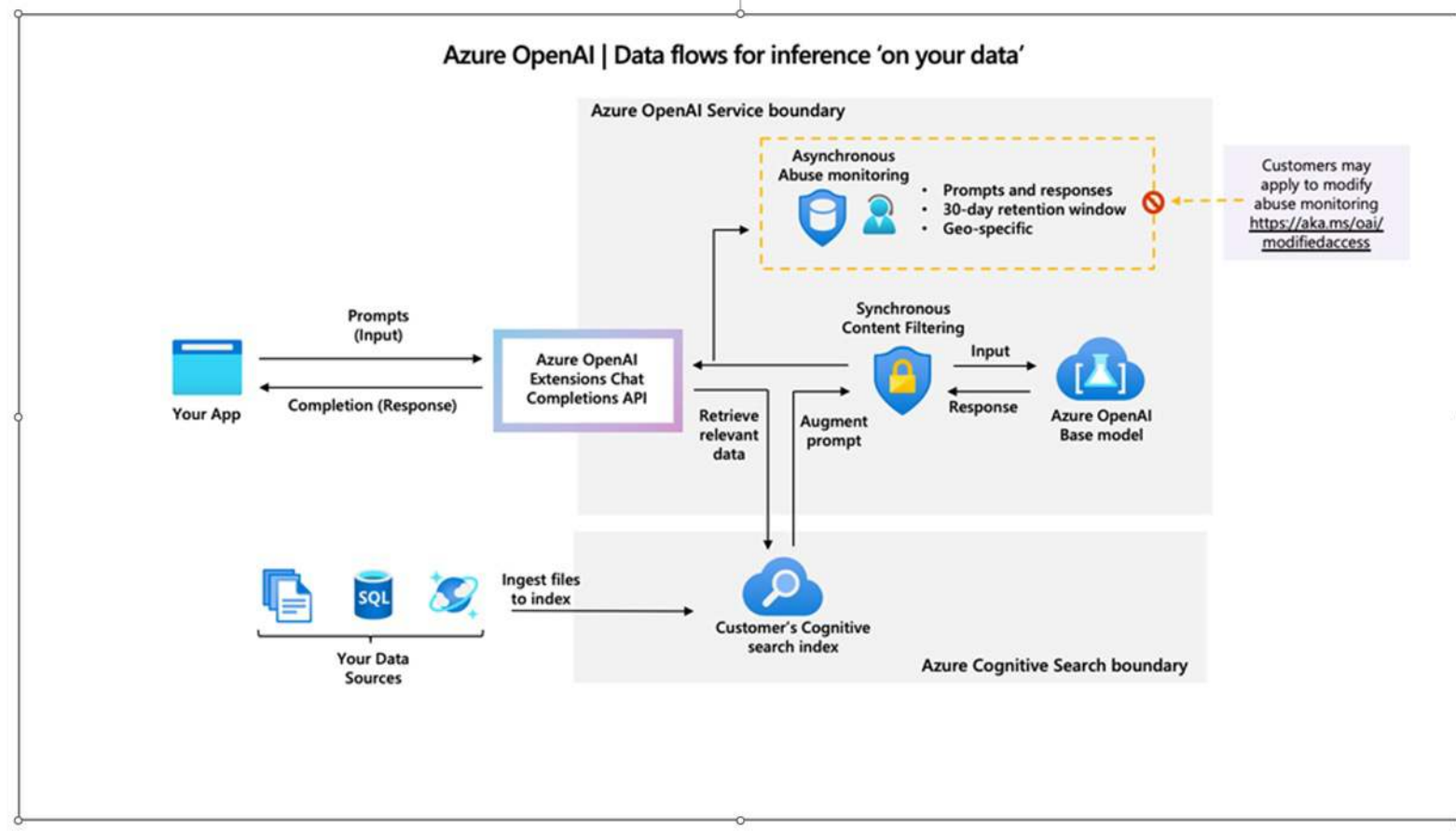
The company allowed engineers at its semiconductor arm to use the AI writer to help fix problems with their source code. But in doing so, the workers inputted confidential data, such as the source code itself for a new program, internal meeting notes data relating to their hardware.

The upshot is that in just under a month, there were three recorded incidences of employees leaking sensitive information via ChatGPT. Since ChatGPT retains user input data to further train itself, these trade secrets from Samsung are now effectively in the hands of OpenAI, the company behind the AI service.

# How it Works



# Microsoft Azure – OpenAI



# The Ethical Rules of Legal Technology

---

Rule 4-3.3 prohibits a lawyer from knowingly “making a false statement of fact or law to a tribunal or fail[ing] to correct a false statement of material fact or law previously made to the tribunal[.]” When representing a client, “a lawyer shall not knowingly... make a false statement of material fact or law to a third person[.]”

- AI programs do not always cite their sources, and some AI programs have been caught making “surprising mistakes’ with basic math” or even creating facts (or polling data) that do not exist.

# Will Artificial Intelligence replace lawyers?



# Implications for the Legal Industry

---

## LAW FIRM

- Staffing
  - Speed
  - Efficiency
  - Technological Expertise
  - Staffing Model
- Revenue and Profitability
- Risk Exposure

## CLIENT

- Use of Confidential Information
- Lowering Costs of Routine Tasks
- Insourcing vs. Outsourcing
- Engagement Guidelines

## BUSINESS

- Trade Secret Protection
- Ownership
- Trademark/Copyright
- Privacy
- Contracting

# Implications for Litigation

---

- Your opponents (e.g., class lawyers) will have access to the same capabilities, allowing them to do more with less human resources and investment.
  - Drafting Complaints
  - Conducting Document Review
  - Preparing for Depositions
- Likely to use requests for inspection to require opponents to use such technology to search for and find information?
  - Can you identify any document talking about the toxic risk of your products?
  - Can you identify any design defects in your products?
  - When did you first discover a latent defect in your design?
  - Can you identify when the first complaint arose about X?
  - Who are the most knowledgeable people in your company on Y?

# Best Practices Around Adoption

---

- Understand the technology.
- Understand the security risks (confidentiality).
- Mitigate confidentiality risks (prohibitions or limitations, technical protections, transparency, notice consent).
- Understand the quality of the results.
- Determine whether and what guardrails to establish (validation).
- Set a clear written policy (law firms, vendors).
- Education and training.
- Communication between law firm and clients (e.g., engagement letters, master service agreements).

# Questions?

John Rosenthal

Winston & Strawn LLP

JRosenthal@Winston.com

Ashish Prasad

HaystackID

APrasad@HaystackID.com