

How to Engineer Prompts for GenAI Document Review in eDiscovery

Generative artificial intelligence (“GenAI”) models, aka Large Language Models or ChatBots, have generated significant interest in the legal and eDiscovery field since released in 2023. These models improve the quality and utility of using AI tools in connection with legal document review. Interacting with GenAI models occurs when users prompt it by asking a question or providing instructions. These questions and/or instructions are referred to as ‘prompts.’ Prompts are text or other inputs sent to the GenAI models, which are utilized to generate textual responses and/or other content. Within the Lexbe eDiscovery Platform, users can submit a prompt that queries the content of a single specific document, or they can submit a prompt that queries the collective content of multiple documents. Both implementations have their place and utility, and use cases and examples for each are addressed below along with drafting tips.

Drafting Tips

As GenAI models work in textual ‘chat mode’, users need to understand how to draft, evaluate, modify, and test prompts to maximize the benefits and utility of these new tools. GenAI responses are dependent on the specific wording and context of the prompts. As a result, ‘prompt engineering’ involves the process of optimizing the wording of prompts to more effectively guide the models and produce useful responses. Chatting with these AI engines is designed to be no different than corresponding with a human. Generally, the more specific the request, and the more context provided, the more accurate the response is likely to be. While some requests may not require much particularity (e.g. “Translate this document into English”), more complex requests will benefit from additional context and specificity.

Improve the quality of the GenAI’s responses by being mindful of the following when crafting your prompt:

- ✓ **Talk to It Like a Human** - The simplest advice is to interact with the AI as you would an intelligent human, as GenAI models are designed to be most effective if used in this manner. In particular, write instructions as you would for an experienced attorney, paralegal, or other legal staff member on your team.
- ✓ **Be Specific** - Useful prompts are usually specific, so provide details, constraints, or examples to guide the AI model. Ambiguity in prompts can lead to irrelevant or off-topic responses.

Request a Free Demo or Consultation

One of our experienced eDiscovery Sales Consultants will be happy to answer your questions.

Call us today at
800.401.7809 x22,
or email us at sales@Lexbe.com

How to Engineer Prompts for GenAI Document Review in eDiscovery

- ✓ **Provide Context** - The more background information and context provided, generally, the better the response will be. Even redundancy can be helpful sometimes as it provides more background information for the AI engine to work with.
- ✓ **Treat It Like a Review Protocol** - The process of creating a summary for an insurance company, or summarizing the key issues of a case to define responsiveness and train internal law firm document reviewers or contract outside managed reviewers, is very similar to the process for developing GenAI prompts in a litigation context. These known processes can be adapted to develop more useful and reliable prompts.
- ✓ **Iterate Instructions and Test Like with Keywords** - Legal professionals are all familiar with keywords in legal research and eDiscovery search applications. Developing well-constructed prompts is not all that different, as it involves iteratively trying different approaches to achieve the best combination and trade-off between broadness and specificity.
- ✓ **Start Broad and then Narrow with Useful Qualifications** - When testing a prompt, one strategy is to begin with a broad request, such as from a Request for Production (e.g., provide all documents involving discussion of the 'XYZ Project'), and then qualify or narrow the broad statement as needed. But be careful. If the instructions are too broad (e.g., 'identify all documents relevant to an insurance case'), or require excessive legal judgment (e.g., 'identify all documents that could show negligence'), results may be inferior.
- ✓ **Broaden Instructions** - Another strategy is to state presumptive inclusions for key aspects of a query. E.g., "Identify all documents that discuss the XYZ Project, including anything involving meetings in Cincinnati in 2022 or Columbus in 2023."
- ✓ **Define Terms if Needed** - GenAI Models are amazingly well-informed about most persons, places, and things that might be involved in a litigation. However, if obscure language or entities are included in a prompt, then defining those elements may improve results. E.g., 'XYZ Project'
- ✓ **Specify Length** - The GenAI Models will provide a more or less detailed response based on the length specified in the prompt. For example, you can request a summary that is 20 words, 200 words, or within a range of 25-50 words depending on your requirements.
- ✓ **Formatting as Needed** - If needed, tell the AI model the specific format you want the response to be in. Requesting that a response include a list in bullet point or ordinal numbered format, or that a summary be written in a specific style or tone, can provide a more easily digestible output.
- ✓ **Use Simple but Complete Language** - Keep your language clear, concise, and straightforward, while still supplying sufficient context. Avoid jargon or overly complex terminology that might confuse the AI model or lead to misinterpretations.
- ✓ **Improve Results with Iterative Submissions** - GenAI models 'remember' the previous prompt and response and allow you to modify your response, taking your previous response into consideration.

How to Engineer Prompts for GenAI Document Review in eDiscovery

Prompts for Single or Multiple Document Review

In the eDiscovery context, GenAI models can be extremely helpful in answering questions or applying textual processes to a single document at a time, across a group of identified documents (subject to the context window of the applicable GenAI engine), or against all documents in a litigation database or repository via processing functionality.

In the Lexbe eDiscovery Platform, submitting a prompt to query the text of a single document can be done via the CoPilot icon in the Document Viewer. Users can also query the collective text of multiple documents by using CoPilot on the Browse page. Lastly, Lexbe's AutoPilot can run GenAI prompts with the touch of a button across all or a specified subset of documents in a case. This occurs as a processing step with responses automatically placed into a field of the user's choosing. This process could be used as a Level 1 review, or a full replacement for human review as part of a fully-automated review process.



Language Translation - GenAI models do an excellent job of translating documents on the fly. This can be particularly useful in connection with an email thread that might contain multiple non-English languages which can confound other automated translation technology and translators as well.

Example Prompt 1: "Translate this document into English."

The screenshot displays the Lexbe eDiscovery platform interface. The main window shows a document viewer with a German text document. The document content includes:

Aus der geplanten Fusion mit Dynegy wurde nichts. Stattdessen endete die 16-jährige Geschichte von Enron mit der größten Pleite seit Gründung der Vereinigten Staaten. Ein Überblick über die Ereignisse.

Juli 1985:
Das US-Unternehmen Houston Natural Gas fusioniert mit Internorth, einem Erdgas-Konzern aus Omaha (Nebraska). Es entsteht der Konzern Enron Börsen-Chart zeigen, der zunächst vor allem als Gaspipeline-Betreiber tätig ist.

Absturz: Der 3-Jahres-Chart der Aktie von Enron

1989: Enron steigt in den Handel mit Erdgas ein und wird innerhalb weniger Jahre der größte Gas-Händler in den USA und Großbritannien.

Dezember 2000: Der Konzern teilt mit, dass Jeffrey Skilling, bislang Präsident und Chief Operating Officer (COO), im Februar 2001 das Amt des Vorstandsvorsitzenden übernehmen wird. Ex-CEO Kenneth Lay soll Chairman bleiben. Die Aktie von Enron erreicht am 28. Dezember eine neues 52-Wochen-Hoch und schließt bei 84,87 Dollar.

August 2001: Nach nur sechs Monaten gibt Jeffrey Skilling wieder auf und reicht seinen Rücktritt ein. Der Konzern holt Kenneth Lay an die Spitze des Vorstands zurück und macht ihn wieder zum CEO.

15. Oktober 2001: Bei einem offiziellen Russland-Besuch führt Lay Gespräche mit US-Wirtschaftsminister Don Evans über ein Enron-Projekt in Indien. Nach offiziellen Angaben werden mögliche Schief lagen von Enron in diesem Gespräch nicht erwähnt.

16. Oktober 2001: Enron veröffentlicht seine Zahlen für das dritte Quartal. Nach Angaben von Finanzvorstand Andrew Fastow liegt der Verlust bei 638 Millionen Dollar. Gleichzeitig ist zu erfahren, dass das Eigenkapital von Enron um 1,2 Milliarden Dollar gesunken ist. Dies soll vor allem auf Beteiligungs-Verträge zurückzuführen sein, für die Fastow verantwortlich zeichnet.

22. Oktober 2001: Enron bestätigt, dass die US-Börsenaufsicht SEC (Securities and Exchange Commission) eine Vor-Untersuchung begonnen hat, um mögliche Konflikte

The right sidebar shows the 'CoPilot' interface with a 'Translate' prompt and its output:

Translate

From the planned merger with Dynegy, nothing came of it. Instead, the 16-year history of Enron ended with the largest bankruptcy since the founding of the United States. An overview of the events.

July 1985:
The US company Houston Natural Gas merges with Internorth, a natural gas corporation from Omaha, Nebraska. The Enron Corporation is created, initially operating primarily as a gas pipeline operator.

Crash: The 3-year chart of Enron's stock.

1989: Enron enters the trading of natural gas and becomes the largest gas trader in the United States and Great Britain within a few years.

December 2000: The company announces that Jeffrey Skilling, previously President and Chief Operating Officer (COO), will assume the position of CEO in February 2001. Former CEO Kenneth Lay will remain Chairman. Enron's stock reaches a new 52-week high on December 28 and closes at \$84.87.

How to Engineer Prompts for GenAI Document Review in eDiscovery



Language Translation (cont'd)

Example Prompt 2: "Translate this document into English using a reading level for a college level or above expert ready."

Example Prompt 3: "Translate this document into English and identify the applicable author and original language for all translated text."

Note that documents don't have to be translated in order for prompts to be effective. Since GenAI works on topics and not the words themselves, it is very effective at working across languages even if language translation is not specifically requested for the output.



Acronym Identification - Unknown acronyms can slow document reviewers and lead to review errors. Because GenAI models are trained on so much data, they can recognize significantly more acronyms than most reviewers, including those with extensive industry experience.

Example Prompt: "Please identify and provide the meanings of any acronyms found in this document."

The screenshot displays the eDiscovery software interface. The top navigation bar includes options for Native, PDF, HTML, Redaction, and Compare. The main document viewer shows an email with the following content:

```

From: Hall, Nora <Nora.Hall@mail.house.gov>
Sent: Thursday, February 17, 2000 9:30 AM
To: jeb@jeb.org
Subject: FW: SATS budget

I'm forwarding this to you for your enlightenment on the SATS that I spoke
to you about in your office. The e-mail is from Embry
Riddle Aeronautical University.

-----
> From: ZellwegA@cts.db.erau.edu[SMTP:ZellwegA@cts.db.erau.edu]
> Sent: Thursday, February 17, 2000 8:34 AM
> To: Hall, Nora
> Subject: SATS Budget
>
> Nora, Here's my assessment of the FY-01 SATS budget request. Hope
> that's the kind of info you need. Dres
>
> NASA Appropriation for Small Aircraft Transportation System (SATS)
> The President's budget for FY-01 contains a little less than $10M
> to
> start a new five-year "focus" programs for SATS. The objective of
> the
> program is to show "technology readiness" for the SATS enabling
> technologies. NASA plans to partner with states (state
> governments,
> local governments, airports, economic development groups, business
> etc.) to:
>
> (a) Demonstrate the SATS technologies, and
>
> (b) To show how SATS would work as a total system, used by people
> in
> the state as a new transportation alternative, and thus to
> demonstrate
> the economic viability SATS.
>
> NASA has stated that Virginia and Florida will be the two lead
> states
>
> for SATS demos. Virginia, because of its proximity to Washington,
> D.C. and to NASA Langley Research Center and because of the strong
> support from Virginia Secretary of Transportation, will very likely
> be
> selected as the "Technology Showcase" (item (a) above).
>
> We believe that Florida is in an ideal position for the total
  
```

The right-hand sidebar shows the 'Coding' and 'Details' panels. Under 'Responsiveness, Privilege, and Confidentiality', the 'Not Specified' radio button is selected. Below this, the 'CoPilot' section is visible, and the 'Acronyms' section is expanded to show the following list:

Acronyms and their meanings in the given document:

1. SATS - Small Aircraft Transportation System
2. FY-01 - Fiscal Year 2001
3. NASA - National Aeronautics and Space Administration

How to Engineer Prompts for GenAI Document Review in eDiscovery



Document Summarization - A key use case of GenAI in eDiscovery is to summarize long documents and highlight key points. Human reviewers fatigue reading documents dense with text which are often industry specific, technical in nature, and almost always boring. Summary types can be specified in the prompt, including the desired length of the summary, a focus on key case issues, specific custodians, date ranges, or other factors important to the litigation at issue.

Example Prompt 1: "Summarize this document in 100 to 150 words."

Example Prompt 2: "These documents contain medical records detailing treatment administered to a patient on specific dates. Your task is to extract key details and organize them into a list in chronological order based on the date of treatment. If a date of treatment isn't available, then sort the document to the end of the date-sorted list. Identify the medical provider's name, purpose of the visit, diagnosed conditions, a succinct treatment summary within 50 words, prescribed medications, and any mentioned medications. Your analysis will facilitate an efficient understanding of the patient's medical history."

The screenshot displays the Lexbe eDiscovery platform interface. At the top, a navigation bar includes 'Home', 'Case', 'Analysis', 'Browse', 'Discovery', 'Management', 'Account', 'Processing', and 'Admin'. A 'Filter Applied: Upload Batch is "Medical Records"' banner is visible. The 'CoPilot' AI engine window is open, showing the 'AI Engine' set to 'OpenAI'. The output is a 'Medical Records Summary' containing three entries:

- Date of Treatment:** 01/01/2008
Medical Provider: John Smith MD
Purpose of Visit: Followup cervical spinal stenosis
Diagnosed Conditions: Cervical spinal stenosis with associated right greater than left upper extremity radiculopathy including weakness
Treatment Summary: Ms. ABC's condition was assessed and discussed in detail. Nonoperative management was deemed insufficient, and surgical intervention was planned. A C4-5, C5-6 anterior cervical decompression and fusion was described, with surgery scheduled for after 07/20/07.
Prescribed Medications: Not listed
Mentioned Medications: Not listed
Supporting Documentation: [Document bf12f713-b4dc-ee11-ac6d-0aad7126d95f (link)](placeholder)
- Date of Treatment:** 04/01/2008
Medical Provider: John Smith MD
Purpose of Visit: Postoperative followup - Cervicalgia, cervical radiculopathy, and difficulty swallowing status post cervical fusion
Diagnosed Conditions: Cervicalgia, cervical radiculopathy, difficulty swallowing post cervical fusion C3-C7
Treatment Summary: The patient reported increased difficulty swallowing and pain post-fall. Dr. S ordered MRIs and x-rays. The patient's pain was constant, rated 10/10, and managed with hydrocodone. A previous surgery was scheduled but not completed.
Prescribed Medications: Hydrocodone
Mentioned Medications: Not listed
Supporting Documentation: [Document b912f713-b4dc-ee11-ac6d-0aad7126d95f (link)](placeholder)
- Date of Treatment:** 07/01/2009
Medical Provider: Jane Doe MD
Purpose of Visit: Neurologic examination
Diagnosed Conditions: Not specified
Treatment Summary: The patient was alert and oriented with no cognitive or language dysfunction. Cranial nerves, manual motor testing, deep tendon reflexes, sensory examination, and nail hit were all normal. A noncontrast CT scan showed no acute infarction.

At the bottom of the CoPilot window, it shows 'Go to page: 1 / 1' and 'Items per page: 25'.

How to Engineer Prompts for GenAI Document Review in eDiscovery



Deposition Summarization - Summarizing/digesting depositions can be time-consuming, expensive, tedious, or all three! Even for an experienced litigation professional it can take an entire working day to summarize a single deposition transcript. An excellent use case for GenAI tools is to handle this important but difficult task quickly, at lower cost, and with excellent quality.

Example Prompt 1: "List all referenced Exhibits in this transcript. Reference where the Exhibit is discussed, then include a 30-word summary of the applicable testimony related to each of the specified exhibits. Generate this list in numbered format."

CoPilot

AI Engine
OpenAI

Deposition Transcript/Exhibit Summary

1. Exhibit 1 (Page 13-14): This is a transcript from a deposition in the case of Russell-Stanley Holdings vs. Vincent Buonanno. Nadeau confirms his employment at NECC and his role as a truck driver.
2. Exhibit 2 (Page 15-16): Transcript from a deposition in Emhart vs. Home Insurance Company. Nadeau acknowledges his participation but cannot recall specifics.
3. Exhibit 3 (Page 16-17): Transcript from a deposition in Emhart Industries, Inc. vs. NECC. Nadeau recalls picking up drums from Otis Air Force Base and Quonset Point.
4. Exhibit 4 (Page 35-36): An affidavit with Nadeau's signature, mentioning the pickup of 2400 barrels with turbine oil residues from Otis Air Force Base.
5. Exhibit 5 (Page 37-38): A statement signed by Nadeau, discussing the pickup of barrels from Otis Air Force Base and Quonset Point, labeled with "turbine oil."
6. Exhibit 6 (Page 44): Department of Defense Sealed Bid Sale document. Nadeau does not recognize the document or the names listed as contacts for Otis and Quonset.
7. Exhibit 7 (Page 55-56): Aerial photograph of the Centredale site from 1965. Nadeau identifies the NECC and Metro-Atlantic buildings and the Texas Tower.
8. Exhibit 8 (Page 121-122): A Sanborn map of the Centredale site from 1965. Nadeau confirms the locations of NECC and Metro-Atlantic buildings, including the Texas Tower.
9. Exhibit 9 (Page 121-122): A larger version of the Sanborn map. Nadeau clarifies the locations of NECC buildings, the Texas Tower, and the loading doors for trucks.

How to Engineer Prompts for GenAI Document Review in eDiscovery



Deposition Summarization (cont'd)

Example Prompt 2: Review the following deposition transcript as an experienced litigation attorney focusing on the following key issues;

- A. Describe issue 1
- B. Describe issue 2
- C. Describe Issue 3
- D. Describe issue 4
- E. Describe issue 5

Review the transcript and create a comprehensive deposition summary focused on the key issues in the litigation. List each key issue as a topic header in bold and then list as bullet points what the deponent/witness answered questions under the topic in date order. For each item, list the page and line numbers.

Example Prompt 3: “Summarize from the following transcript regarding the issue of alleged negligence on the part of ABC Corporation, specifying the applicable testimony, the date range of activities discussed, and other individuals or entities mentioned, and the page and line ranges of the applicable testimony.”

Example Prompt 4: “List all referenced Exhibits in the following transcript and reference the page or pages in the transcript where the Exhibit is discussed and include a 30-word summary in each instance of applicable testimony about or related to the specified exhibit.”

Example Prompt 5: “In the following deposition transcript list all the objections that were made in numbered format, including who made it, a comprehensive summary of the objection, and a reference to the page number in the transcript where the objection was made.”



Responsiveness - Determining ‘relevance’ under the FRCP or similar state laws, or ‘responsiveness’ to Requests for Production, is usually the core and a key function of a document review. Modern search technology works well enough, and legacy technology assisted review (‘TAR’) software has made many improvements on the process, but GenAI TAR promises to make significant improvements in quality, functionality, and cost. GenAI TAR also can provide a reason for the responsiveness/relevance determination which provides a major improvement in utility.

How to Engineer Prompts for GenAI Document Review in eDiscovery



Responsiveness

(cont'd)

Example Prompt 1: “Is this document relevant to the following topic: All documents concerning a bid to host the Summer Olympic Games in Florida. Answer (yes/no), followed by the confidence level of your assessment on a scale of 1 to 10, with 1 being the least confident, and 10 being the most confident, and then state your reasoning for the responsiveness determination in 35-50 words.

Desired format:

Yes/No

Confidence Level:

Reasoning:“

The screenshot displays the eDiscovery software interface. The top navigation bar includes 'DISC', 'DOC', 'NOTES', 'FACTS', 'RELATED', and 'COPILOT'. The main document viewer shows an email with the following content:

050307 Viewing Doc 4 of 25

HITS ORIGINAL HTML PAGE PDF TEXT REDACTION ANNOTATION

9b11ff7f-9593-ee11-ac6d-0aad... 1 / 1 | 100%

From: Jeb Bush <jeb@jeb.org>
Sent: Sunday, May 14, 2000 12:44 PM
To: 'Russell Williams'
Subject: RE: Support Olympic Bid

Thanks for your email and the advice. I appreciate your hard work to bring the Olympics to Central Florida.

Jeb Bush

-----Original Message-----
From: Russell Williams [mailto:rkwilliams@oasistech.net]
Sent: Friday, May 12, 2000 1:52 AM
To: jeb@jeb.org
Subject: Support Olympic Bid

Dear Governor Jeb Bush,

Please ignore the negative people who don't realize that the Summer Olympics is the world's greatest event. I don't believe any other area in the United States is better suited for the Olympics, and Ed Turanchik and his Olympic Committee are doing a great job of promoting it. Since we have everything in our favor to make tremendous profits for worthwhile projects at the expense of Olympic advertisers, it is not unreasonable for the State of Florida to guarantee we will take

Olympic Bid Relevance

Yes
Confidence Level: 9
Reasoning: The document contains correspondence discussing efforts and support for bringing the Summer Olympics to Central Florida, which directly relates to a bid to host the Games.

Example Prompt 2: “Does this document include any content about hiring individuals for the following positions: [Title 1, Title 2, Title 3], answer (Yes/No). Any mention of job qualifications in hiring, or interviews for these specific job positions should result in a response of “Yes.” If the document only mentions job positions other than those specifically listed above, then the response should be “No.”

How to Engineer Prompts for GenAI Document Review in eDiscovery



Privilege or Attorney Work Product - Determination of privilege or attorney work product can, in many cases, be the most important part of a document review. Keyword searches can be effective to a large degree, but GenAI technology can serve as an additional check, identifying potentially privileged communications and/or work product that might be otherwise overlooked as it can classify on document aspects beyond text.

Example Prompt 1: "Carefully examine the document and identify any content that may be subject to legal privilege, such as attorney-client privilege or work product privilege. If privileged, answer (Yes/No). Specify the information you believe is privileged in list format, and then provide your reasoning in 35-50 words for the basis of their privileged status. Additionally, indicate whether any other legal privileges may apply to the document."

Desired format:

Yes/No:

Reasoning:

050451 Viewing Doc 5 of 25 < > ↓ SAVE CANCEL

HITS ORIGINAL HTML PAGE PDF TEXT REDACTION ANNOTATION

cb72f685-9593-ee11-ac6d-0a... 1 / 2 100%

From: Rex <ecurry@interaccess.net>
Sent: Friday, May 12, 2000 3:52 AM
To: Bush, Jeb
Subject: Stop government \$\$\$ for the 2012 Olympics!

Dear Governor Bush:
I object to the government giving tax support to the Olympics.

The government, which has the ability to use force to extract taxes from its citizens, has now expanded its responsibilities to include sporting events. I personally find this offensive. I think you have recently experienced the abuse of Federal Powers with the INS coming into our state, violating the rights of a citizen, all supposedly for a good reason. Well now the State of Florida has come up with its own good reason to use our resources, and I am sure displace and inconvenience many private citizens, to support the Olympic effort. This too is an abuse of the power given to the government, by its citizens, to be used for much more serious purposes.

The power of government should not be wielded for these types of purposes. When are our political leaders going to realize this and curb their insatiable urges to intrude into our daily lives unnecessarily.

Assessing Privilege

No

Reasoning:
The document is a letter from an attorney, Ellis Rexwood Curry, to Governor Jeb Bush expressing personal views on public funding for the Olympics and other civic issues. It does not contain confidential communication between attorney and client, nor does it discuss legal advice or strategy related to any specific legal matter. Therefore, attorney-client privilege or work product privilege does not apply. No other legal privileges appear to apply to this document.

Example Prompt 2: "Review the document thoroughly and identify any information that may be subject to attorney-client privilege and/or work product privilege. Specifically, look for communications or materials exchanged between you and your client that were made for the purpose of seeking or providing legal advice or representation. Additionally, consider any materials prepared in anticipation of litigation or for trial. List the identified content and provide a brief rationale explaining why you believe they are privileged under applicable law."

How to Engineer Prompts for GenAI Document Review in eDiscovery



Confidential, Highly-Confidential & AEO Documents - Classifying documents as confidential, highly-confidential, or attorneys eyes only (AEO) can be of critical importance. In the context of commercial litigation, identifying if a document contains any confidential information can be difficult in isolation as specifics of what the author considers to be proprietary, confidential, or trade secret can be difficult for an AI to determine without sufficient context. That said, GenAI classification can be helpful. Two example prompts are below. The first takes a broad approach using the types of documents typically identified as confidential in confidentiality, non-disclosure, or litigation protective orders, such as trade secrets, financial records, strategic plans, employee information, etc. The second approach is more specific, having the GenAI engine classify confidentiality by using known confidentiality entities or other terms. A third approach would be to combine the two approaches.

Example Prompt 1: "Carefully examine the document and identify if any content includes Confidential Information. If confidential, answer (Yes/No). Specify the information you believe is potentially confidential information. "Confidential Information" shall include, but is not limited to, information related to XYZ Company's business and its actual or anticipated business, research, and development or related to the XYZ Company's technology, clients, business plans, promotional and marketing activities, finances and other business affairs, inventions, product plans, designs, databases, computer software (including source and object codes), algorithms, models, methodologies, know-how, processes, trade secrets, ideas, patent applications, workflows, hardware configuration information, lists of or information related to employees and consultants, lists of or information related to the XYZ Company's suppliers and customers, pricing strategies and price lists, and other technical, business, financial, customer and product development plans, forecasts, strategies and information.

Example Prompt 2: "Carefully examine the document and identify if any content could be considered to include Confidential Information. If confidential, answer (Yes/No). Specify the information you believe is potentially confidential information. "Confidential Information" shall include, but is not limited to, and reference to the following: the ABC Software Project, meetings in Dallas in January or February 2023, the DEF R&D Facility, and John Smith or Mary Jones.



PII Identification - Classifying personally identifiable information ('PII') is critical as part of modern litigation and is required to conform with operative privacy requirements. GenAI classification can be helpful in identifying applicable PII as part of an efficient redaction workflow.

How to Engineer Prompts for GenAI Document Review in eDiscovery



PII Identification (cont'd)

Example Prompt 1: “Identify all documents that contain any bank or other financial account numbers, social security numbers, passport numbers, credit card numbers, email addresses, phone numbers, user names, passwords, or other personal information. Ensure the identified documents reflect the diversity of information that can be used to uniquely identify individuals. The output should resemble a structured list, each item clearly representing a distinct type of PII commonly found in documents.”

The screenshot shows the eDiscovery interface with a document viewer on the left and a results pane on the right. The document is a Chase bank statement for July 1, 2008 through July 31, 2008, for Primary Account: 00000988081483. The results pane lists the following PII:

- 1. Bank Account Numbers:**
 - Primary Account: 00000988081483
- 2. Names:**
 - John Doe
- 3. Addresses:**
 - 1234 Sample St., Whirlpool, FL 12345
 - JPMorgan Chase Bank N.A., Ohio/West Virginia Markets, P O Box 260180, Baton Rouge, LA 70826-0180
- 4. Bank Name:**
 - JPMorgan Chase Bank N.A.
- 5. Phone Numbers:**
 - Customer Service Information: 1-800-935-9935
 - Hearing Impaired: 1-800-242-7383
 - Para Espanol: 1-877-312-4273
 - International Calls: 1-713-262-1679
- 6. Website Information:**
 - WebSite: www.Chase.com

A note at the bottom of the results pane states: "Please note that the document does not contain any social security numbers, passport numbers, credit card numbers, email addresses, user names, passwords, or other types of personal information that are commonly used to uniquely identify individuals."

Example Prompt 2: “List any account numbers, email addresses, social security numbers, or phone numbers that appear in this document. List each one only once.”

The screenshot shows the eDiscovery interface with the same document viewer. The results pane lists the following PII:

- Account Numbers:**
 - Primary Account: 00000988081483
- Phone Numbers:**
 - Customer Service Center: 1-800-935-9935
 - Hearing Impaired: 1-800-242-7383
 - Para Espanol: 1-877-312-4273
 - International Calls: 1-713-262-1679

A note at the bottom of the results pane states: "No email addresses or social security numbers are listed in the document."

How to Engineer Prompts for GenAI Document Review in eDiscovery



Pre-Identified Key Issue Classification - An excellent use of GenAI is to identify documents that are probative to key issues in the case or investigation. Key issues in a case can be discerned from the pleadings, requests for production, client interviews, proposed jury instructions, and other sources. GenAI excels in this use case over CAL predictive coding algorithms and keyword searches, as GenAI better interprets context in prompts and documents.

Example Prompts for Multi Document Review or an Entire Case



Unknown Key Issue(s) Identification - Another strategy to identify key issues is to ask the GenAI engine to identify and summarize any issues discussed, rather than classifying known issues as stated above.

Example Prompt 1: “Analyze the provided group of documents and identify the key issues addressed within them. Consider the main themes, arguments, and concerns presented across the documents. Highlight any recurring patterns, conflicting viewpoints, or areas of consensus among the documents. Pay particular attention to any unresolved questions, points of contention, or areas requiring further exploration. Your analysis should offer insights into the core issues discussed within the document set and provide a comprehensive understanding of the topics covered.”

CoPilot

AI Engine
OpenAI

Unknown Key Issues (v1)

The provided group of documents revolves around the theme of restoration of civil rights for felons, particularly in the state of Florida. The key issues addressed within these documents include:

- **Restoration of Voting Rights for Felons**:** A significant portion of the correspondence is concerned with the process and policies regarding the restoration of voting rights for individuals with felony convictions. There are discussions about the efforts to assist inmates with restoration of civil rights forms (Document d8939892-9193-ee11-ac6d-0aad7126d95f [\[link\]](#)), as well as the mention of a lawsuit alleging insufficient assistance by the Department of Corrections in this regard (Document d8939892-9193-ee11-ac6d-0aad7126d95f [\[link\]](#)).
- **Clemency Process**:** Several documents touch on the clemency process, including specific cases where individuals are seeking clemency or restoration of rights (Document 0e9f9892-9193-ee11-ac6d-0aad7126d95f [\[link\]](#), Document 819f9892-9193-ee11-ac6d-0aad7126d95f [\[link\]](#), Document eebecfa8-9193-ee11-ac6d-0aad7126d95f [\[link\]](#)). There is also mention of a proposal for reforming the restoration of civil rights process (Document a60a7daa-9193-ee11-ac6d-0aad7126d95f [\[link\]](#)).
- **Public Opinion and Political Action**:** Some documents reflect public opinion and political action related to the rights of felons, including voting rights and clemency (Document 104c83a4-9193-ee11-ac6d-0aad7126d95f [\[link\]](#), Document 1ac2cfa8-9193-ee11-ac6d-0aad7126d95f [\[link\]](#), Document 47f4a08c-9193-ee11-ac6d-0aad7126d95f [\[link\]](#)). There are also references to political figures and their stances on these issues (Document 95399298-9193-ee11-ac6d-0aad7126d95f [\[link\]](#)).
- **Racial Disparities**:** Some documents address the racial disparities in the criminal justice system and how they relate to the disenfranchisement of felons, with specific references to the impact on African American communities (Document d2b85095-9193-ee11-ac6d-0aad7126d95f [\[link\]](#)).
- **Legal and Judicial Considerations**:** There are references to legal and judicial considerations, including court rulings and legal

How to Engineer Prompts for GenAI Document Review in eDiscovery



Unknown Key Issue(s) Identification (cont'd)

Example Prompt 2: “Carefully review these documents and provide the following information:

Document Overview:

Provide a brief summary of the documents under review, including:

- Total number of documents.
- Types of documents (e.g., contracts, correspondence, legal memos).
- Date range of the documents.

Identification of Unknown Issues:

1. Legal Issues: Identify any legal issues or inconsistencies within the documents.
2. Factual Discrepancies: Highlight any inconsistencies or contradictions in the factual information presented.
3. Missing Information: Note any gaps or missing pieces of information that could impact the case.
4. Relevance to Case: Determine the relevance of each document to the case at hand.
5. Potential Risks: Identify any potential risks or liabilities indicated by the documents.

Analysis and Recommendations:

Provide an analysis of the identified issues, including:

- Legal implications.
- Potential strategies for addressing the issues.
- Recommendations for further investigation or action.

Conclusion:

Summarize the key findings and conclusions drawn from the document review process.

****Formatting Instructions:**** Please use bullet points, headings, and clear subsections to organize your responses effectively. Ensure that your analysis is concise and to the point, providing sufficient detail to support your conclusions.”

CoPilot AI Engine OpenAI

Unknown Key Issues Identification

Document Overview:

- Total number of documents: 30
- Types of documents: The documents are primarily email correspondences, with some including attachments related to policy discussions, legal matters, and personal requests. There are also draft statements and reports attached to some emails.
- Date range of the documents: The documents range from March 10, 1999, to November 22, 2003.

Identification of Unknown Issues:

1. Legal Issues:
 - There are discussions about the restoration of civil rights for felons, which involves legal processes and potential changes to existing laws or policies.
 - The case of Samer Majzoub involves unpaid restitution, which has legal implications for the restoration of civil rights.
 - The document from Greg Munson mentions two lawsuits related to felon voting rights, indicating ongoing legal disputes.
2. Factual Discrepancies:
 - There are varying opinions and statements about the process and criteria for restoring civil rights to felons, which may indicate inconsistencies in understanding or application of the law.
 - The document from Tim Moore challenges the accuracy of a report on voter disenfranchisement, suggesting factual discrepancies in publicized data.
3. Missing Information:
 - Some documents refer to attachments or additional information that is not provided within the documents themselves.
 - The document from Greg Munson refers to a case (FCBSL v. Moore) without providing the full context or background information.
4. Relevance to Case:

How to Engineer Prompts for GenAI Document Review in eDiscovery



Name and/or Entity Identification - Identifying the individuals or entities discussed in a group of documents can be a valuable way of understanding the agents or key players in a dispute.

Example Prompt 1: “Generate a list of all entities across the set of documents provided. Entities may include, but are not limited to, companies, committees, people, organizations, groups, lobbyists, and other relevant entities mentioned within the documents. Please ensure that the list captures all distinct entities and their variations (e.g., full names, acronyms, aliases) mentioned throughout the documents.

Your task is to extract and compile a comprehensive inventory of entities present in the documents. This list should encompass both proper nouns and noun phrases that refer to entities within the context provided. Ensure that the list distinguishes between different types of entities and provides sufficient context to understand their relevance within the documents.

Please employ appropriate techniques for entity recognition, normalization, and disambiguation to ensure the accuracy and completeness of the entity list. Additionally, if there are any uncertainties or ambiguities regarding entity identification, provide annotations or notes to clarify the context and assist in further analysis.”

Example Prompt 2: “Analyze these documents and list all names of people mentioned more than twice within them. The documents may contain various formats of names (e.g., John Doe, J. Doe, Doe, John). Consider different representations of names, including those embedded within sentences or paragraphs. Ensure that all identified names are accurately extracted, listed separately in bullet point format, with duplicates removed.”



List of Email Addresses - Extracting all email addresses from a group of documents or across an entire case can be helpful with identifying subjects or custodians who might be witnesses in a matter.

Example Prompt 1: “Provide a list in bullet point format of each email address found across these emails. Only list each email address one time.”

Example Prompt 2: “Analyze these documents and list all email addresses mentioned within them. The documents may contain various formats of email addresses (e.g., user@example.com, user[at]example[dot]com) and names (e.g., John Doe, J. Doe, Doe, John). Consider different representations of email addresses, including those embedded within sentences or paragraphs. Ensure that all identified email addresses are accurately extracted, listed separately in bullet point format, with duplicates removed.”

How to Engineer Prompts for GenAI Document Review in eDiscovery



List of Email Addresses (cont'd)

The screenshot shows the CoPilot interface with a dark blue header. On the right side of the header, it says "AI Engine" and "OpenAI" with a dropdown arrow. Below the header, there is a search bar with the text "Email Addresses (Deduplicated)" and a play button icon. The main content area displays a list of email addresses, each preceded by a hyphen. The list includes:

- SB226@aol.com
- tcahill@mcn.org
- fmunson@lakeland.net
- Lhyvarinen@hotmail.com
- MARIAJSERRANO@aol.com
- RWilcox106@aol.com
- tdb1116@garnet.acns.fsu.edu
- tlistser@survivingthesystem.com
- arogerwilson@hotmail.com
- athena141@yahoo.com
- LeeC@cunninghamoil.com
- jeb@jeb.org
- Wendy.Berger@MyFlorida.com
- catislandcorp@hotmail.com
- TMCUIC@aol.com
- sjr@msn.com
- sevarddiamond@aol.com
- ROSHIDASAN@aol.com
- czickler@juno.com
- EMART51@aol.com
- Viclukis@aol.com
- Cassandra.L.Collins@dep.state.fl.us
- jciotti256@earthlink.net



Timeline - Generating a timeline of what has occurred across a group of documents is a common practice, and can be invaluable. GenAI can accomplish this in a way that could not be addressed by prior versions of AI.

Example Prompt 1: "Create a chronological timeline of the key events across these documents. Ensure each event is presented in sequential order, providing brief descriptions or context where necessary to illustrate the progression of events."

How to Engineer Prompts for GenAI Document Review in eDiscovery



Timeline (cont'd)

CoPilot

AI Engine
OpenAI

Timeline - Key Events

1. **March 10, 1999** - Tim Moore from FDLE sends an email to Governor Jeb Bush with an attachment discussing the issue of voter disenfranchisement and providing a report that counters the numbers publicized by the Washington Post. The report suggests that the number of disenfranchised individuals in Florida is significantly lower than reported. (Document: d2b85095-9193-ee11-ac6d-0aad7126d95f [\[link\]](#))
2. **October 30, 1999** - An email is sent to Jeb Bush expressing opposition to clemency for a three-time convicted drug felon, Mr. Mazziotti, and urging the Governor to uphold the law. (Document: ab2f9298-9193-ee11-ac6d-0aad7126d95f [\[link\]](#))
3. **November 10, 1999** - Jeb Bush receives an email from Stuart Strahl announcing the merger of the National Audubon Society and Florida Audubon Society, creating a unified conservation program in Florida. (Document: 95399298-9193-ee11-ac6d-0aad7126d95f [\[link\]](#))
4. **November 16, 1999** - An email is sent to various recipients, including Jeb Bush, discussing the Secret Oaks Association permit application and the opposition to it. (Document: dc389298-9193-ee11-ac6d-0aad7126d95f [\[link\]](#))
5. **November 29, 1999** - Brian S. Yablonski sends an email to Justin Sayfie, copying Jeb Bush, discussing various stories in the Washington Post about events in Florida, including a story about a lawsuit where drug dealers sued a hotel for negligence and won. (Document: 5e379298-9193-ee11-ac6d-0aad7126d95f [\[link\]](#))
6. **March 13, 2001** - Charles R Zickler sends an email to Jeb Bush expressing his opposition to HJR49, which would allow felons to have their voting rights reinstated. (Document: 1ac2cfa8-9193-ee11-ac6d-0aad7126d95f [\[link\]](#))

Example Prompt 2: "Create a timeline of significant events using information extracted from these documents. Consider the key milestones, developments, and noteworthy occurrences to construct a comprehensive timeline. Ensure the timeline reflects the chronological order and provides insights into the progression of events. Format the timeline with clear dates and concise descriptions of each event for clarity and coherence."

How to Engineer Prompts for GenAI Document Review in eDiscovery



Comparing Document Draft Versions - Comparing multiple versions of draft versions of documents is essential for understanding changes, intent, and potential impacts on a commercial litigation case. Beyond the simple red-line comparisons or simple text differential analysis that software has done well for years, deeper insights can be gleaned by leveraging specific, targeted GenAI prompts that aim to uncover more nuanced analysis and insights into a witnesses' motivation, intent and actions.

Example Prompt 1: "Identify and interpret the significance of added, removed, or altered clauses and provisions in the different versions of the following document, grouped by legal issue"

Example Prompt 2: "For the following contractual versions, identify and compare changes in key contract elements such as liabilities, indemnities, payment terms, and termination rights across multiple versions of a contract. Highlight added, removed, or modified clauses, and group them by their legal implications on the parties involved. Provide a summary of how these changes affect the overall balance of obligations between the parties.

Example Prompt 3: "Review the following multiple versions of a supply chain contract to analyze changes in risk allocation clauses, including indemnities, warranties, and insurance requirements. Group these changes and include the date of the change and summarize the changed terms.



For More Information on Using GenAI in eDiscovery View the Following YouTube Videos

