

Five Ways to Use ChatGPT for Investigations and Ediscovery¹



Image created by John Tredennick using Midjourney

By John Tredennick and Dr. William Webber

Since its release in late 2022, ChatGPT has captivated the world, drawing a million users in its first five days and totaling over 100 million today. Whether you are fearful or excited, there is little doubt we have entered a transformational era, much like the invention of the printing press, the steam engine, cell phones, and the Internet itself. The extent and speed of ChatGPT's progress (and of competitor systems) remains the only open question.

Like everyone else, we have read lots of flowery articles about the potential of ChatGPT to redefine the professional world. In this article, we intend to go beyond generalities by demonstrating five specific ways legal professionals can

¹ This article first appeared in the May 24, 2023 issue of [LegalTech News](#) from [Law.com](#), a publication of American Lawyer Media.



use ChatGPT (and its underlying GPT engine) to make ediscovery more efficient and cost-effective.

Using GPT and a basic search engine, we can show how legal professionals will soon take advantage of Large Language Models like GPT to streamline their ediscovery processes and better understand the key documents in their cases.

Prepare to be amazed.

Our Topic for This Exercise

Most ediscovery efforts start with a Rule 34 Request for Production. For this exercise, we will use one of the legal track topics from NIST's annual [Text Retrieval Conference \(TREC\)](#), which was created in part to test the ability of different machine learning algorithms to find relevant documents:

"2000 Recount – All documents concerning the contested result of the 2000 presidential election."

Our discovery collection consists of approximately 290,000 emails that Jeb Bush made public from his two terms as Governor of Florida. TREC used these emails for its programs and they provide a good test bed for the integration of GPT and traditional ediscovery search and review methods. Our GPT system makes these documents available for interactive analysis and interrogation through GPT.

1. Using GPT to Help with Keyword Searches

One of the first steps in addressing an RFP is to identify appropriate keywords and build a Boolean search. While we can lift terms from the topic request, experienced ediscovery professionals know that the process can require hours of time testing keywords and analyzing search results to come up with a comprehensive list of terms to include in the search.

Let's start by asking GPT to help us determine keywords for this topic with the following prompt (a question or instruction sent to GPT). Watch as we have a conversation with GPT to flesh out our keyword list and turn it into a Boolean search:

Prompt:

Here is my topic: "2000 Recount – All documents concerning the contested result of the 2000 presidential election." Can you suggest keywords to help make my search better?

GPT:

Keywords that could improve the search for "2000 Recount – All documents concerning the contested result of the 2000 presidential election" include: election manipulation, voting machines, Diebold, security, paper trail, ballot review, consortium, statewide hand count, overvoted ballots, polls, voting, election, fraud, revote, federal law, media projection, electoral college, standard ballot, personnel, troubleshooting, felony list, and legitimate voters.

These keywords are supported by the following documents:

- Document 018716: Mentions "election manipulation," "voting machines," "Diebold," "security," and "paper trail."
- Document 099018: Mentions "ballot review," "consortium," "statewide hand count," "overvoted ballots," and "outcome."
- Document 064808: Mentions "polls," "voting," "election," "contested result," "legal maneuvers," "ambiguous conclusions," "procedures," "previous elections," and "precedent."
- Document 124142: Mentions "contested result," "fraud," and "revote."
- Document 065969: Mentions "federal law," "media projection," "polls," "electoral college," "standard ballot," "voting machine," "recount," and "Florida votes."
- Document 123697: Mentions "voting equipment," "personnel," "troubleshooting," "felony list," "legitimate voters," and "right to vote."
- Document 145857: Mentions "touch screen voting machines," "security," "open election records," "Perfect Voting System," "investigation," "Georgia," "Maryland," "audit," and "public inspection."
- Document 123823: Mentions "recount," "contested result," and "AP Election Documentary."

- Document 064796: Mentions "recount," "ballots," "Secretary of State," and "deadline."
- Document 113386: Mentions "recount," "contested result," and "political careers."
- Document 113389: Mentions "recount," "contested result," "political career," and "Katherine Harris."

Prompt:

Can you suggest any other keywords I might use to improve my search?

GPT:

Here are some additional keywords that could improve your search:

- Hanging chads
- Butterfly ballot
- Supreme Court
- Gore v. Bush
- Florida Supreme Court
- Bush v. Palm Beach County Canvassing Board
- Bush v. Gore
- Absentee ballots
- Voter disenfranchisement
- Voter suppression
- Partisan politics
- Legal challenges
- Election reform
- Public opinion
- Media coverage
- Political polarization

These keywords are supported by the documents mentioned earlier, as well as other documents related to the 2000 presidential election recount.

Prompt:

Thanks. Can you build me a Boolean search based on all of the keywords you

have suggested?

GPT:

Certainly! Here's a Boolean search string that includes all of the keywords suggested:

("election manipulation" OR "voting machines" OR "Diebold" OR "security" OR "paper trail" OR "ballot review" OR "consortium" OR "statewide hand count" OR "over voted ballots" OR "polls" OR "voting" OR "election" OR "fraud" OR "revote" OR "federal law" OR "media projection" OR "electoral college" OR "standard ballot" OR "touch screen voting machines" OR "open election records" OR "Perfect Voting System" OR "investigation" OR "Georgia" OR "Maryland" OR "audit" OR "public inspection" OR "hanging chads" OR "butterfly ballot" OR "Supreme Court" OR "Gore v. Bush" OR "Florida Supreme Court" OR "Bush v. Palm Beach County Canvassing Board" OR "Bush v. Gore" OR "absentee ballots" OR "voter disenfranchisement" OR "voter suppression" OR "partisan politics" OR "legal challenges" OR "election reform" OR "public opinion" OR "media coverage" OR "political polarization")

You can use this search string in combination with your preferred search engine or database to find relevant documents related to the 2000 presidential election recount.

We are not suggesting that GPT has nailed every possible keyword one might use in a search. But it sure provides a good start at this tricky task. We would use it every time we had to build a Boolean search.

2. Using GPT to Summarize Documents

In this case, GPT linked its keyword suggestions to specific documents, which can be quite helpful. As a next step in the process, let's ask GPT to provide more details about these documents and tell us how they relate to our search objectives.

Prompt:

Show me summaries of the key documents you found on our topic.

GPT:

Certainly! Here are summaries of some of the key documents related to the 2000 presidential election recount:

- 1. Document 018716:** An email from Jeb Bush to Glenda Hood discussing concerns about the security and reliability of electronic voting machines, specifically those made by Diebold Inc. The email suggests that Bush and Hood should discuss the issue.
- 2. Document 099018:** An email from Mark Silva to Jeb Bush requesting Bush's thoughts on the results of the ballot review conducted by the consortium. The review found that if a statewide hand count had been conducted as ordered by the state Supreme Court, George W. Bush would still have won the election. Silva asks for Bush's thoughts on the relevance of this information and the idea that over voted ballots could have changed the outcome.
- 3. Document 064808:** An email from Sue Zipay to Jeb Bush regarding the recount for the 2000 presidential election. Zipay argues that doing another recount by hand for specific districts in Florida would be unfair to other voters and the system in place.
- 4. Document 123697:** An email from a Florida voter to the Florida Department of State expressing concerns about voting equipment, personnel, troubleshooting, felony list, legitimate voters, and the right to vote.
- 5. Document 145857:** An email from David Biddulph, inventor of the "Perfect Voting System," to Jeb Bush discussing concerns about the security of touch screen voting machines and suggesting Biddulph's invention as a solution.
- 6. Document 123823:** An email from Katie Muniz criticizing an AP Election Documentary story, stating that it sounds more like a conspiracy theory than a documentary. She also expresses her confidence that the film does not

mention the true election reform that took place under the leadership of Governor Jeb Bush after the November 2000 election.

These documents provide insight into various aspects of the 2000 presidential election recount, including concerns about voting equipment, the role of the state government in ensuring fair elections, and public opinion on the issue.

We can also ask GPT to provide extended summaries of these documents, asking it to identify the people mentioned (and their roles) and to provide key dates. Here are several examples:

Docid: 018716

Date: December 4, 2003

The document is an email from Jeb Bush to Glenda Hood, dated December 4, 2003, regarding a New York Times editorial by Paul Krugman titled "Hack the Vote." The editorial discusses concerns about the security and reliability of electronic voting machines, specifically those made by Diebold Inc. The author argues that the lack of a paper trail and the company's history of security issues and cover-ups could potentially lead to election manipulation. The email suggests that Bush and Hood should discuss the issue.

People mentioned:

- Jeb Bush (Sender; Former Governor of Florida)
- Glenda Hood (Recipient; Secretary of State of Florida; gehood@mail.dos.state.fl.us)
- Paul Krugman (Mention; New York Times Op-Ed Columnist)

Docid: 113386

Date: March 15, 2002

The document is an email sent by William Yardley to Jeb Bush on March 15, 2002. Yardley is writing a story about the key local players of the 2000 recount and how it affected their political careers. He asks Jeb Bush how he perceives the recount as a turning point in his political career and if he thinks the general

state political landscape has changed because of the recount or if its impact was eclipsed by Sept. 11. There are no specific keywords suggested in the document.

People mentioned:

- William Yardley (Sender; Author of the email; WYardley@herald.com)
- Jeb Bush (Recipient; Former Governor of Florida)
- Katherine Harris (Mention; Florida Secretary of State during the 2000 election)
- Carol Roberts (Mention; Palm Beach County Commissioner and candidate for Congress)
- Feeney (Mention; Tom Feeney, former Florida Speaker of the House)
- Scott Maddox (Mention; Former Mayor of Tallahassee and candidate for Governor)
- Lois Frankel (Mention; Former Florida State Representative and candidate for Congress)

Docid: 113389

Date: March 15, 2002

Summary: William Yardley, a journalist, sent an email to Governor Jeb Bush asking for a five-minute interview or email response regarding how the 2000 recount affected his political career and if it changed the state political landscape. Yardley is writing a story about key local players of the recount who are now running for higher office. The email also mentions other individuals on Yardley's list, including Katherine Harris, Carol Roberts, Feeney, Scott Maddox, and Lois Frankel.

People:

- Sender: William Yardley (WYardley@herald.com)
- Recipient: Jeb Bush (jeb@jeb.org)
- Mentioned: Katherine Harris, Carol Roberts, Feeney, Scott Maddox, Lois Frankel.

Docid: 145857

Date: August 19, 2003

Summary:

The email discusses concerns about the security of touch screen voting

machines and the need for open election records. The sender, David Biddulph, suggests his invention, the Perfect Voting System, which makes elections work like the state lottery and allows for public inspection and audit of election records. He also mentions that the governors of Georgia and Maryland are calling for an investigation into the security of touch screen voting machines. The email includes articles from the Atlanta Journal Constitution and the Washington Post on the voting machine security issue. Jeb Bush is asked for his opinion on the matter.

People:

- Sender: David Biddulph
- Receiver: Glenda Hood Secretary
- Mentioned: Jeb Bush, Clmoore7@aol.com, JohnsonB@HillsboroughCounty.org, Governor Sonny Perdue, Secretary of State Cathy Cox, Representative Tom Feeney, Hillsborough County Supervisor of Elections Buddy Johnson, Paul Craft

Docid: 123823

Date: September 20, 2002

The document is an email sent by Katie Muniz on September 20, 2002, to Kathleen Shanahan and the press office. Muniz criticizes an AP Election Documentary story, stating that it sounds more like a conspiracy theory than a documentary. She also expresses her confidence that the film does not mention the true election reform that took place under the leadership of Governor Jeb Bush after the November 2000 election.

People mentioned:

- Katie Muniz (Sender; Katie.Muniz@MyFlorida.com)
- Kathleen Shanahan (Recipient)
- Jeb Bush (Mention; Governor of Florida at the time of the 2000 election)
- Press Office (Recipient)

Docid: 099018

Date: November 11, 2001

This document is an email sent by Mark Silva to Jeb Bush on November 11,

2001. Silva is requesting Bush's thoughts on the results of the ballot review conducted by the consortium. The review found that if a statewide hand count had been conducted as ordered by the state Supreme Court, George W. Bush would still have won the election. The advantage for Gore only emerged when overvoted ballots were included, but these were not being counted. Silva asks for Bush's thoughts on the relevance of this information and the idea that over voted ballots could have changed the outcome.

People mentioned:

- Mark Silva (Sender; Reporter for the Orlando Sentinel; MSilva@orlandosentinel.com)
- Jeb Bush (Recipient; Brother of George W. Bush and Governor of Florida at the time of the 2000 election)

Docid: 064796

Date: November 13, 2000

The document is a letter from Morgan Lamb to Governor Bush, dated November 13, 2000, regarding the recount of ballots in the 2000 presidential election. Lamb suggests that there is nothing to be gained from incessant recounting of ballots, as they have been handled too much at this point. Lamb also suggests that the Secretary of State should not extend the deadline for reporting voting results from any county past the Tuesday deadline.

People mentioned:

Morgan Lamb (Sender; magnifique@ispchannel.com)
Governor Bush (Recipient; jeb@jeb.org)

Docid: 064841

Date: November 11, 2000

The document is an email sent on November 11, 2000, from Kathy and Joe Blais to Jeb Bush, with a copy to info@georgewbush.com. The email expresses confusion about the legality of recounting only "Democratic selected" precincts in Florida and not all precincts. The email also criticizes the Democrats for dividing the country and inciting people to file lawsuits. The email ends with a message of support for Jeb Bush and George W. Bush.

People mentioned:

Kathy & Joe Blais (Sender; kbbblais@email.msn.com)

Jeb Bush (Recipient; jeb@jeb.org)

info@georgewbush.com (CC)

3. Using GPT to Answer Questions About Specific Documents.

Once we find interesting documents for our analysis, we can ask GPT questions about their contents. This can help us better understand the issues underlying our search topic and move forward with our investigation and document search.

Here is an example of asking GPT questions about the specific documents we have found so far:

Prompt:

Can you tell me what is being discussed in the following documents? Document 018716, Document 099018, Document 064808, Document 124142, Document 065969. Who was speaking to whom and what were the discussions about?

GPT:

Sure! Here's a summary of the discussions in each of the documents you mentioned:

1. [Document 018716](#): An email from Jeb Bush to Glenda Hood discussing concerns about the security and reliability of electronic voting machines, specifically those made by Diebold Inc. The email suggests that Bush and Hood should discuss the issue.
2. [Document 099018](#): An email from Mark Silva to Jeb Bush requesting Bush's thoughts on the results of the ballot review conducted by the consortium. The review found that if a statewide hand count had been conducted as ordered by the state Supreme Court, George W. Bush would still have won the election. Silva asks for Bush's thoughts on the relevance of this information and the idea that overvoted ballots could have changed the outcome.

3. [Document 064808](#): An email from Sue Zipay to Jeb Bush regarding the recount for the 2000 presidential election. Zipay argues that doing another recount by hand for specific districts in Florida would be unfair to other voters and the system in place.

4. [Document 124142](#): An email from an unknown sender to Jeb Bush expressing concerns about the contested result of the 2000 presidential election and suggesting that a revote be held to ensure a fair outcome.

5. [Document 065969](#): An email from a constituent to Jeb Bush expressing concerns about the voting process in Florida and asking for Bush's help in ensuring that every vote is counted. The email discusses issues such as the use of voting machines, the recount process, and the role of the state government in ensuring fair elections.

In each of these documents, the sender is reaching out to Jeb Bush to express concerns or ask for his thoughts on various aspects of the 2000 presidential election recount. The discussions cover topics such as the security and reliability of voting machines, the fairness of the recount process, and the role of the state government in ensuring fair elections.

We also have the option to query the larger document population using a combination of keyword search, the relevance algorithm and GPT. The user starts by entering a question in the search box, perhaps in freeform mode. The question can be natural language with as many helpful terms as possible.

Prompt:

I am looking for documents based on the following Boolean Search: ("election manipulation" OR "voting machines" OR "Diebold" OR "security" OR "paper trail" OR "ballot review" OR "consortium" OR "statewide hand count" OR "overvoted ballots" OR "polls" OR "voting" OR "election" OR "fraud" OR "revote" OR "federal law" OR "media projection" OR "electoral college" OR "standard ballot" OR "touch screen voting machines" OR "open election records" OR "Perfect Voting System" OR "investigation" OR "Georgia" OR "Maryland" OR "audit" OR "public inspection" OR "hanging chads" OR "butterfly ballot" OR "Supreme Court" OR "Gore v. Bush" OR "Florida Supreme Court" OR "Bush v. Palm Beach County

Canvassing Board" OR "Bush v. Gore" OR "absentee ballots" OR "voter disenfranchisement" OR "voter suppression" OR "partisan politics" OR "legal challenges" OR "election reform" OR "public opinion" OR "media coverage" OR "political polarization").

Please find the most relevant ones and give me a summary of the main points you found.

GPT:

Document 013083:

David Biddulph, inventor of the "Perfect Voting System," sent an email to Governor Jeb Bush of Florida, urging him to take the lead on making election records open to public inspection and audit. Biddulph's invention makes elections work like the state lottery, where the voter never receives a hard copy of their ballot and their vote is secret unless they wish to prove the official vote tabulation is in error. Biddulph suggests that citizen confidence in the efficacy of elections would be comparable to their confidence in the lottery. The cost of licensing the invention is \$0.10 per vote cast, and voting equipment would need to have a printer to prepare the recount ballot and voter "stub." Biddulph estimates that the cost per vote cast would increase about 3-5%. Biddulph also mentions that the Governors of Georgia and Maryland are calling for an investigation into the security of touch screen voting machines, and he includes an article from the Atlanta Journal Constitution on the issue. (Document 013083)

Document 145857:

David Biddulph, inventor of the "Perfect Voting System," sent an email to Governor Jeb Bush and others, urging them to consider making election records open to public inspection and audit. He suggested that his invention, which makes elections work like the state lottery, could address the questions being raised about paperless voting systems. Biddulph's system would cost \$0.10 per vote cast and would require voting equipment to have a printer to prepare the recount ballot and voter "stub." He estimated that the cost per vote cast would increase about 3-5%. Biddulph also included an article from the Atlanta Journal Constitution about the security of touch screen voting machines and a recent Washington Post article on the voting machine security issue. Governor Bush forwarded the email to Secretary of State Glenda Hood

and asked for her response. (Document 145857)

Document 018716:

Jeb Bush received an email from ODemocracy@aol.com containing an editorial by Paul Krugman published in the New York Times on December 2, 2003. The editorial discusses the issue of paperless voting machines, specifically those produced by Diebold Inc. Krugman argues that the lack of a paper trail and the unreliability of the software used in these machines could potentially lead to election manipulation. He cites examples of security breaches and cover-ups by Diebold, as well as leaked internal emails suggesting that the company knew their system was flawed. Krugman calls for legislation requiring digital voting machines to leave a paper trail and for their software to be available for public inspection. Jeb Bush suggests discussing the issue with Glenda Hood Secretary. (Document 018716)

Document 170585:

The email is from David to Jeb Bush and others, asking if they are using the same voting machines and if they have a paper trail. The email is dated May 3, 2004. The email mentions a decision made by California Secretary of State Kevin Shelley to ban touch-screen voting machines in the state due to concerns over security. The decision affects 14 counties and millions of voters. The touch-screen machines, specifically those made by Diebold, have been accused of being vulnerable to computer hackers who might change votes. To regain certification, the touch-screen makers will have to create a paper trail or put in place a list of 23 security rules. Shelley has also asked the Attorney General to pursue criminal and civil actions against Diebold. The email is relevant to the search query as it mentions "voting machines," "security," and "paper trail." (Document 170585)

Document 170525:

The document is an email from Jeb Bush to Glenda Hood, dated May 3, 2004, regarding a concern raised by David about the voting machines being used in Florida. The email includes an article about California Secretary of State Kevin Shelley's decision to ban touch-screen voting machines in the state due to concerns over security and the vulnerability of the machines to computer hackers. Shelley's decision goes beyond the recommendations of an advisory panel and means that 14 counties and millions of voters in California will have

to find another voting system. The touch-screen makers will have to create a paper trail or put in place a list of 23 security rules to get back their certification by November. Shelley has also asked the Attorney General to pursue criminal and civil actions against Diebold, a company that makes touch-screen machines. (Document 170525)

Overall, these documents discuss concerns about the security and reliability of voting machines, particularly touch-screen machines produced by Diebold, and the need for a paper trail and public inspection of election records. They also mention the decisions made by California Secretary of State Kevin Shelley to ban touch-screen voting machines and the actions taken against Diebold. The documents suggest that there is a need for election reform and increased public confidence in the electoral process.

Prompt:

What were the discussions about "hanging chads" What are they and why do they matter?

GPT:

The discussions about "hanging chads" refer to the controversy surrounding the punch-card ballots used in the 2000 Presidential election in Florida. The issue arose when some of the ballots had not been fully punched through, leaving a "hanging chad" that made it unclear which candidate the voter had intended to select. This led to a recount and legal battle that ultimately ended with the Supreme Court ruling in favor of George W. Bush. The issue of "hanging chads" is mentioned in documents 154395, 065969, and 099027.

Prompt:

What were the main arguments pro or con about the fairness and accuracy of the election process?

GPT:

The main arguments about the fairness and accuracy of the election process were:

- Concerns about the accuracy of voting equipment and polls, as well as reports of legitimate voters being falsely placed on a felony list during the 2000

Presidential election and not being cleared to vote in the upcoming election. The authors argued that every legitimate vote counts and that the right to vote should not be jeopardized. They suggested that the entire state should be recounted if necessary to ensure fairness. (Documents 123697, 123959, 099605, 124162)

- Criticism of the electoral college and the need for all candidates to be elected by popular vote. (Document 065969)

- Concerns about the potential for computerized voting to be manipulated and the lack of a paper trail. (Document 127510)

- Disagreement about the need for a recount of Florida votes even after Governor Bush takes office as President of the United States of America. (Document 065969)

- Criticism of the idea of doing another recount by hand only for specific districts in Florida, which would be grossly unfair to every other voter in the country and to the system that was in place before the first vote was cast. (Document 064808)

- Arguments that the Florida election controversy was not just about technical issues like dimples and chads but also about the disenfranchisement of millions of citizens with past felony convictions. (Document 099605)

- Governor Jeb Bush's argument that 65 out of 67 counties in Florida got it right and that local supervisors of election are responsible for making elections work. He also provided \$30 million to implement reforms and stated that the failure of Dade and Broward Counties to get it right is a local problem. (Documents 124341, 124344)

Cited documents: 123697, 123959, 065969, 127510, 064808, 099605, 124341, 124344, 124162.

The system analyzes the question and then finds relevant documents responsive to the question. These are sent to GPT for analysis along with the question from the user.

4. Using GPT to Create an Investigation Report

The next step is to use GPT to create investigation reports. Assume that an investigator has found a number of important documents through any of a variety of methods (e.g. keyword searches,, witness meetings). Rather than asking a series of questions about the document, the investigator could simply ask for an “Investigation Report.”

The report would include the options to include this kind of information:

- Who is involved?
- What are their roles?
- With whom do they speak?
- When do they speak to each other?
- What do they say?
- Where are they when communicating?

We can develop standard report options and even control report formatting. The report might look something like this:

Prompt:

Please Prepare an Investigation Report for me. Here are the questions to be covered with the answers to each in a separate section of the report.

1. Who is involved?
2. What are their roles?
3. With whom do they speak?
4. When do they speak to each other? Create a timeline for this section.
5. What do they say? Summarize any back and forth conversations.
6. Where are they when communicating? Report only known information.

GPT:

Investigation Report: Communication and Roles in the Context of Election Manipulation and Voting Machines

1. Who is involved?

The following individuals are involved in the communication related to election manipulation and voting machines:

- David Biddulph, inventor of the "Perfect Voting System"
- Jeb Bush, former Governor of Florida
- Glenda Hood, former Secretary of State for Florida
- Clmoore7@aol.com
- JohnsonB@HillsboroughCounty.org
- Representative Tom Feeney
- Hillsborough County Supervisor of Elections Buddy Johnson
- Nora Austin Hall
- ODemocracy@aol.com
- Kevin Shelley, former California Secretary of State
- Kim Alexander, California Voter Foundation spokeswoman
- Diebold Inc.

2. What are their roles?

David Biddulph is the inventor of the "Perfect Voting System" and is advocating for open election records and audits. Jeb Bush is the Governor of Florida and is receiving emails and forwarding them to Glenda Hood, the Secretary of State for Florida. Clmoore7@aol.com, JohnsonB@HillsboroughCounty.org, Representative Tom Feeney, and Hillsborough County Supervisor of Elections Buddy Johnson are recipients of an email from David Biddulph. Nora Austin Hall is sending an email to Jeb Bush discussing an editorial about the vulnerability of electronic voting machines to election manipulation. ODemocracy@aol.com is sending an email to Jeb Bush containing an editorial by Paul Krugman published in the New York Times on December 2, 2003. Kevin

Shelley is the former California Secretary of State who made a decision to ban touch-screen voting machines in the state due to concerns over security. Kim Alexander is a California Voter Foundation spokeswoman. Diebold Inc. is a company that produces touch-screen voting machines.

3. With whom do they speak?

David Biddulph speaks to Governor Jeb Bush of Florida, Clmoore7@aol.com, JohnsonB@HillsboroughCounty.org, Representative Tom Feeney, and Hillsborough County Supervisor of Elections Buddy Johnson. Jeb Bush speaks to Glenda Hood, the Secretary of State for Florida. Nora Austin Hall speaks to Jeb Bush. ODemocracy@aol.com sends an email to Jeb Bush. Kevin Shelley speaks to the Attorney General to pursue criminal and civil actions against Diebold.

4. When do they speak to each other? Create a timeline for this section.

- November 16, 2002: An email is sent by David Biddulph to Governor Jeb Bush of Florida, Clmoore7@aol.com, JohnsonB@HillsboroughCounty.org, Representative Tom Feeney, and Hillsborough County Supervisor of Elections Buddy Johnson, urging them to make election records open to public inspection and audit.

- August 19, 2003: David Biddulph sends an email to Governor Jeb Bush and others, urging them to consider making election records open to public inspection and audit. He also suggests his invention, the "Perfect Voting System," as a solution to address the questions being raised about paperless voting systems.

- December 2, 2003: An editorial by Paul Krugman is published in the New York Times, discussing the issue of paperless voting machines, specifically those produced by Diebold Inc.

- December 4, 2003: ODemocracy@aol.com sends an email to Jeb Bush containing the editorial by Paul Krugman.

- May 3, 2004: David Biddulph sends an email to Jeb Bush and others, asking if they are using the same voting machines and if they have a paper trail. The

email mentions a decision made by California Secretary of State Kevin Shelley to ban touch-screen voting machines in the state due to concerns over security.

- May 3, 2004: Jeb Bush sends an email to Glenda Hood, discussing the concern raised by David Biddulph about the voting machines being used in Florida. The email includes an article about Kevin Shelley's decision to ban touch-screen voting machines in California.

5. What do they say? Summarize any back and forth conversations.

David Biddulph advocates for open election records and audit, and suggests his invention, the "Perfect Voting System," as a solution to address the questions being raised about paperless voting systems. Jeb Bush forwards emails from David Biddulph and ODemocracy@aol.com to Glenda Hood, discussing concerns about the vulnerability of electronic voting machines to election manipulation. Nora Austin Hall sends an email to Jeb Bush discussing an editorial about the vulnerability of electronic voting machines to election manipulation. Kevin Shelley makes a decision to ban touch-screen voting machines in California due to concerns over security and the vulnerability of the machines to computer hackers. He also asks the Attorney General to pursue criminal and civil actions against Diebold. Diebold is accused of being vulnerable to computer hackers who might change votes.

6. Where are they when communicating?

The location of the individuals when communicating is not explicitly mentioned in the documents. However, it can be assumed that the communication is primarily through email, as the documents are emails sent between the individuals. It is possible that some of the communication may have taken place in person or over the phone, but there is no information available to confirm this.



5. Using GPT for Document Review

In a recent paper for Law.com, we asked the question: “[Will ChatGPT Replace Ediscovery Review Teams?](#)” Our research to date strongly suggests that GPT can match or exceed human standards for document review.

We again start our testing with the emails from the TREC Jeb Bush collection. The program administrators reviewed the document set, identifying 34 topics for search testing. We used them to test GPT’s review capabilities, matching the results to the “gold standard” set by TREC’s human reviewers.

Here was a topic we used to test GPT’s review capabilities:

“Slot Machines – All documents concerning the definition, legality, and licensing of “slot machines” in Florida.”

We submitted several test documents to GPT, asking whether they were relevant (or not relevant) to that topic. We got these kinds of responses from GPT:

1. Yes. This email discusses a proposed initiative petition to amend the Florida Constitution to provide authorization for county voters to approve or disapprove slot machines within existing pari-mutuel facilities. It also mentions the Attorney General's questions concerning the ballot summary and single-subject requirement.
2. No. This email does not discuss slot machines or their legality in Florida. It is about a dispute between a private business owner and the Miccosukee tribe.

From there we tested across the 34 topics, tracking results against human review judgments. In many cases, GPT correctly took issue with the judgments from human reviewers. We showed several examples where GPT found documents relevant, although humans marked them as not relevant, and vice versa.

For example, we reviewed documents where GPT took a different view of relevance on this topic:

Space – All documents concerning the space industry, the space program, space travel (whether manned or unmanned, public or private), and the study or exploration of space in Florida.

GPT found 14 relevant documents about Florida’s space program that were marked “not relevant” by humans.

False Positives (14 documents)	
Docid	AI review
035779	Yes. This email discusses the Space Summit and the IT Incubator issue, both of which are related to the space industry in Florida.
067746	Yes. This email is relevant to the topic as it discusses the selection of a director for the Spaceport Florida Authority, which is a public organization related to the space industry in Florida.
090743	Yes. This email discusses the ribbon cutting ceremony for the Delta IV Launch Pad in Florida, which is related to the space industry. It also mentions the Air Force, Boeing, and the economic benefits of the launch pad.

And five documents on the space program marked relevant by human reviewers which GPT concluded were not relevant.

False Negatives (5 documents)	
Docid	AI review
175439	No. This email does not discuss the space industry, the space program, space travel, or the study or exploration of space in Florida.
209943	No. This email does not discuss any topics related to space.
284400	No. This email does not discuss the space industry, the space program, space travel, or the study or exploration of space in Florida. It is about economic development and job growth in Florida.

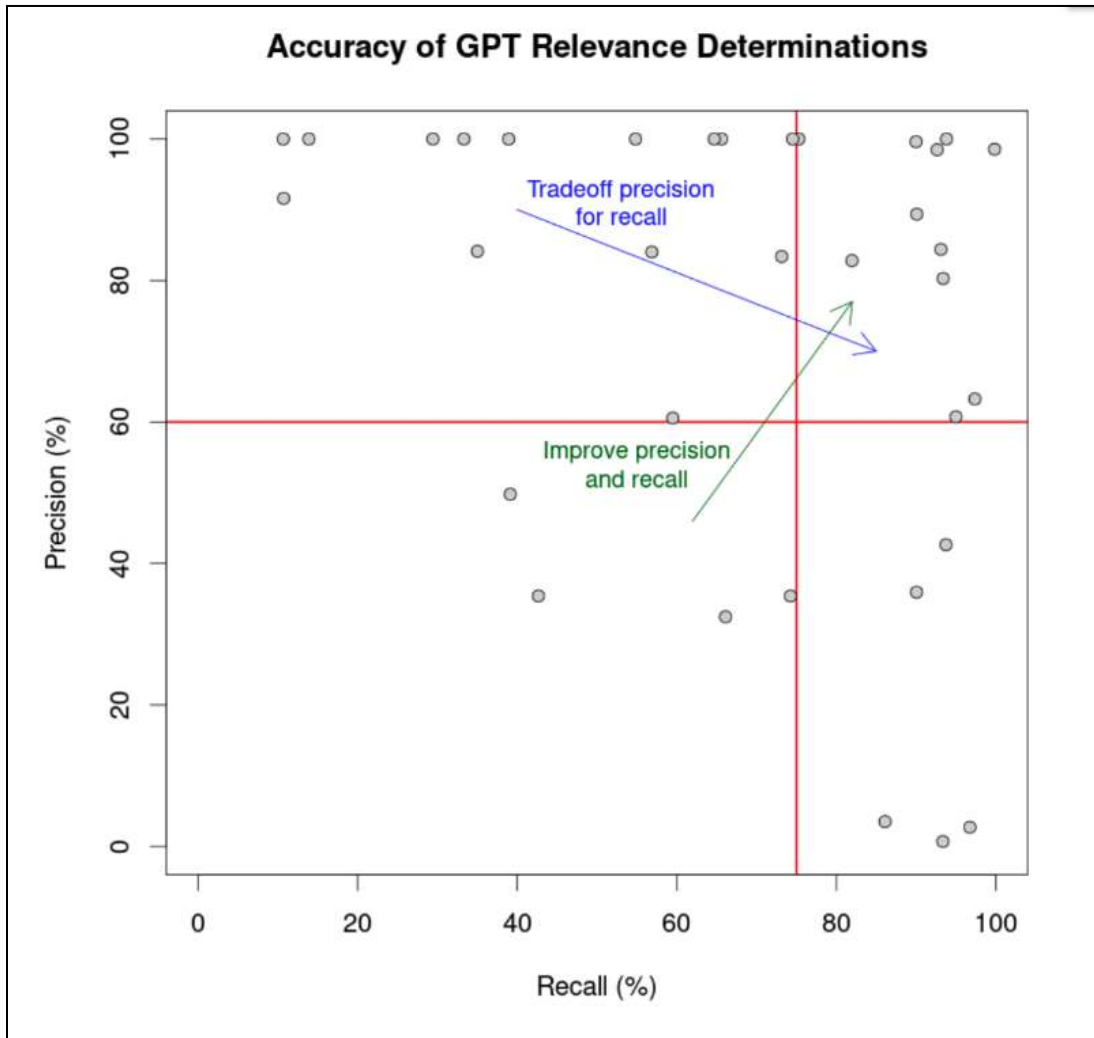
We showed one example using the word Orbit that was marked relevant by the human reviewers.

Text (8365 characters)

The UCF ORBIT mission is to provide world-class resources, content, and knowledge based information through partnerships with business and IT companies in the Florida High Tech Corridor (www.FloridaHighTech.com). In addition, UCF ORBIT is committed to achieving recognition for offering the highest quality content and IT business solution programs available anywhere in the world. A key component of ORBIT's mission is to support UCF's goal as America's leading partnership university and promote economic development in the Florida High Tech Corridor.

A quick read of the document text indicates that this document is not about space exploration. The phrase “ORBIT mission” might mislead a hasty human reviewer, but this is in fact the name of an IT program. GPT was right to mark it non-relevant.

Ultimately, in our first round of testing with GPT 3.5 (the 4.0 version had not yet been released), we found that GPT did as well or better than humans for just over half the topics and worse on the others—even taking the official reviews at face value.



We believe we will be able to improve GPT's scores using 4.0 and offering more detailed prompts along with some other techniques we are now using. Integrating review into GPT is ongoing.

Conclusion

These experiments only scratch the surface of GPT's potential to aid in investigations and discovery efforts. Using GPT to create Boolean searches, to find, summarize and analyze documents, and to create investigation reports are game changers for investigations and ediscovery on their own. If we can also use GPT to replace large human teams for document review, the savings in review costs could run into the billions of dollars across the review industry. Review



times would go from months to weeks or days, moving discovery and case understanding forward as well.

As the legal industry continues to evolve, it's crucial for legal professionals to stay ahead of the curve and embrace new technologies like GPT. The potential for improving efficiency, reducing costs, and achieving better outcomes is just too enormous to pass up. We have no doubt that GPT and other AI tools will continue to revolutionize the way we approach investigations and ediscovery in the future, and we look forward to helping make this new AI-powered future a reality for all discovery professionals.

About the Authors

John Tredennick (JT@Merlin.Tech) is the CEO and founder of Merlin Search Technologies, a cloud technology company that has developed a revolutionary new machine learning search algorithm called Sherlock® to help people find information in large document sets—without having to master keyword search.

Tredennick began his career as a trial lawyer and litigation partner at a national law firm. In 2000, he founded and served as CEO of Catalyst, an international e-discovery search technology company that was sold to a large public company in 2019. Over the past four decades he has written or edited eight books and countless articles on legal technology topics, spoken on five continents and served as Chair of the ABA's Law Practice Management Section.

Dr. William Webber (wwebber@Merlin.Tech) is the Chief Data Scientist of Merlin Search Technologies. He completed his PhD in Measurement in Information Retrieval Evaluation at the University of Melbourne under Professors Alistair Moffat and Justin Zobel, and his post-doctoral research at the E-Discovery Lab of the University of Maryland under Professor Doug Oard.

With over 30 peer-reviewed scientific publications in the areas of information retrieval, statistical evaluation, and machine learning, he is a world expert in AI and statistical measurement for information retrieval and ediscovery. He has almost a decade of industry experience as a consulting data scientist to ediscovery software vendors, service providers, and law firms.



About Merlin Search Technologies

Merlin is a pioneering cloud technology firm, specializing in developing and hosting AI-driven software for search investigations and review. With over twenty years of experience, our team has built and hosted discovery platforms for many of the largest corporations and law firms in the world.

We've built a next-generation search platform, integrating machine learning algorithms with keywords to make it quicker, easier and less costly to find information in large document sets. And introduced Cloud Utility Pricing to save on hosting.

Our mission is to use AI and cloud technologies to make search, investigations and discovery efficient and cost effective.