UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
18/611,498	03/20/2024	Daniel Hunter	56710-0002001	9701

26191 7590 06/21/2024
FISH & RICHARDSON P.C. (TC)
PO BOX 1022
MINNEAPOLIS, MN 55440-1022

EXAMINER	
LE, MICHAEL	

ART UNIT	PAPER NUMBER
2163	

NOTIFICATION DATE	DELIVERY MODE
06/21/2024	ELECTRONIC

Please find below and/or attached an Office communication concerning this application or proceeding.

The time period for reply, if any, is set in the attached communication.

Notice of the Office communication was sent electronically on above-indicated "Notification Date" to the following e-mail address(es):

PATDOCTC@fr.com

Office Action Summary	Application No. 18/611,498	Applicant(s) Hunter et al.	
	Examiner MICHAEL LE	Art Unit 2163	AIA (FITF) Status Yes

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTHS FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

1)☑ Responsive to communication(s) filed on 3/20/2024.

☐ A declaration(s)/affidavit(s) under **37 CFR 1.130(b)** was/were filed on _____.

2a)☐ This action is **FINAL.** 2b) ☑ This action is non-final.

3)☐ An election was made by the applicant in response to a restriction requirement set forth during the interview on _____; the restriction requirement and election have been incorporated into this action.

4)☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims*

5) ☑ Claim(s) 1-20 is/are pending in the application.

5a) Of the above claim(s) _____ is/are withdrawn from consideration.

6) ☐ Claim(s) _____ is/are allowed.

7) ☑ Claim(s) 1-20 is/are rejected.

8) ☐ Claim(s) _____ is/are objected to.

9) ☐ Claim(s) _____ are subject to restriction and/or election requirement

* If any claims have been determined allowable, you may be eligible to benefit from the **Patent Prosecution Highway** program at a participating intellectual property office for the corresponding application. For more information, please see http://www.uspto.gov/patents/init_events/pph/index.jsp or send an inquiry to **PPHfeedback@uspto.gov.**

Application Papers

10)☐ The specification is objected to by the Examiner.

11)☑ The drawing(s) filed on 3/20/2024 is/are: a)☑ accepted or b)☐ objected to by the Examiner.

Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).

Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).

Priority under 35 U.S.C. § 119

12)☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).

Certified copies:

a)☐ All b)☐ Some** c)☐ None of the:

1.☐ Certified copies of the priority documents have been received.

2.☐ Certified copies of the priority documents have been received in Application No. _____.

3.☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

** See the attached detailed Office action for a list of the certified copies not received.

Attachment(s)

1) ☑ Notice of References Cited (PTO-892)

2) ☑ Information Disclosure Statement(s) (PTO/SB/08a and/or PTO/SB/08b) Paper No(s)/Mail Date 3/28/24.

3) ☐ Interview Summary (PTO-413) Paper No(s)/Mail Date _____.

4) ☐ Other: _____.

DETAILED ACTION

Summary and Status of Claims

1. The present application, filed on or after March 16, 2013, is being examined under the first inventor to file provisions of the AIA.

2. This Office Action is in response to Application No. 18/611,498 filed 3/20/2024.

3. Claims 1-20 are pending.

4. Claims 4-10 are rejected under 35 U.S.C. 112(b).

5. Claims 1-20 are rejected under 35 U.S.C. 101.

6. Claims 1-20 are rejected under 35 U.S.C. 103 as being unpatentable over Gupta (US Patent Pub 2022/0092097) in view of Madisetti et al. (US Patent 12,001,462).

Information Disclosure Statement

7. The information disclosure statement filed 3/28/2024 has been fully considered, initialed, and signed by the Examiner. A copy is attached to this Office action.

Claim Rejections - 35 USC § 112

8. The following is a quotation of 35 U.S.C. 112(b):

> (b) CONCLUSION.—The specification shall conclude with one or more claims particularly pointing out and distinctly claiming the subject matter which the inventor or a joint inventor regards as the invention.

9. Claims 4-10 are rejected under 35 U.S.C. 112(b) as being indefinite for failing to particularly point out and distinctly claim the subject matter which the inventor or a joint inventor regards as the invention.

10. Claim 4 recites "... to generate a topic aware outline of the document." It is unclear whether this is an additional outline or the "document outline" recited in base claim 1 is further limited here to be a "topic aware outline." Clarification is required.

11. The remaining claims are rejected because they depend on a rejected claim.

Claim Rejections - 35 USC § 101

12. 35 U.S.C. 101 reads as follows:

> Whoever invents or discovers any new and useful process, machine, manufacture, or composition of matter, or any new and useful improvement thereof, may obtain a patent therefor, subject to the conditions and requirements of this title.

13. **Claims 15-17 are rejected under 35 U.S.C. 101** because the claimed invention is directed to non-statutory subject matter. The claim(s) does/do not fall within at least one of the four categories of patent eligible subject matter because the specification at pg. 24, lines 16-20 describes possible computer readable media as an open ended list of examples. Since the list of examples is open ended, interpretation of the medium can include non-statutory subject matter, such as transitory media. Additionally, the specification at pg. 22, line 28 to pg. 23, line 6 describes a computer storage medium can be signals. For at least these reasons, claims 15-17 are directed to non-statutory subject matter.

14. **Claims 1-20 are also rejected under 35 U.S.C. 101** because the claimed invention is directed to a judicial exception (i.e., a law of nature, a natural phenomenon, or an abstract idea) without significantly more. Determining whether claims are statutory under 35 U.S.C. 101 involves a two-step analysis. Step 1 requires a determination of whether the claims are directed to the statutory categories of invention. Step 2 requires a determination of whether the claims are directed to a judicial exception without significantly more. Step 2 is divided into two prongs, with the first prong having a part 1 and part 2. *See* MPEP 2106; *See* 2019 Revised Patent Subject Matter Eligibility Guidance (2019 PEG).

15. Pursuant to Step 1, claims 1-14 are directed to the statutory category of a process. Claims 15-17 would have been directed to a manufacture, if amendment to overcome the nonstatutory subject matter rejection above. Claims 18-20 are directed to a machine.

16. Claim 1

17. Pursuant to Step 2A, part 1, claims are analyzed to determine whether they are directed to an abstract idea. Under the 2019 PEG, claims are deemed to be directed to an abstract idea if they fall within one of the enumerated categories of (a) mathematical concepts, (b) certain methods of organizing human activity, and (c) mental processes. Here, claim 1 are directed to an abstract idea categorized under mental processes. Courts consider a claims to be directed to a mental process if it "can be performed in the human mind, or by a human using a pen and paper." MPEP 2016(a)(2)(III). Claim 1 recites a mental process because the recited steps recite the action of observation. See MPEP 2106(a)(2)(III). For example, claim 1 recites limitations of receiving a query related to a document from a user, submitting the document to a first LLM along with a first prompt to generate an outline, receiving the document outline, submitting the document to a second LLM with a second prompt to generate metadata of the document and receiving the metadata, submitting the query, metadata, and outline to a third LLM with a third prompt to generate a natural language response to the query and receiving the response, correlating a sentence from the response to a sentence from the document and providing the response and correlation to the user. These limitations merely recite steps that can be performed by a person with the aid of a computer as a tool to perform the abstract idea, which is to summarize a document into an outline and metadata and response to the query about it. Therefore, claim 1 is directed to an abstract idea categorized as a mental process.

18. Pursuant to Step 2A, part 2, claims are analyzed to determine whether the recited abstract idea is integrated into a practical application. In this case, as explained above, claim 1 merely recite a mental process. These limitations describe generating an outline and metadata of a document in order to

generate a natural language response to the query about the document with an indication to a correlation of the response to a part of the document. These steps do not provide meaningful limits on the abstract idea because summarizing documents in order to answer questions about them is a mental process often performed by people. The additional components of using LLMs to perform the various steps are recited at a high level of generality, which do not add meaningful limits on the recited abstract idea to integrate it into a practical application by providing an improvement to the functioning of a computer or technology, implementing the abstract idea with a particular machine or manufacture that is integral to the claim, effecting a transformation or reduction of a particular article to a different state or thing, nor applying the abstract idea in some meaningful way beyond linking its use to computer technology. *See* 2019 PEG. At best, they merely tie the abstract idea to the field of technology and act as a tool for the person to perform an existing process.

19. Since claim 1 is directed to an abstract idea categorized as a mental process and does not integrate the judicial exception into a practical application, claim 1 is directed to a judicial exception.

20. Pursuant to Step 2B, claims are analyzed to determine whether they recite significantly more than the abstract idea. In other words, it is determined whether the claims provide an inventive concept. In this case, claim 1 does not recite limitations that amount to significantly more than the abstract idea. The limitations are steps involving processes that can be practically performed by a human with the aid of pen and paper or with the aid of a computer as a tool to execute the abstract idea. For example, a person can summarize a document into an outline and metadata and response to a query about the document while highlighting a relevant portion of a document with the response. These limitations do not improve the functioning of a computer, improve the technology, apply the abstract idea to a particular machine, effect a transformation, nor provide meaningful limitations beyond linking the abstract idea to computer technology. For at least these reasons, claim 1 is nonstatutory because they are directed to a judicial exception without significantly more.

21. Claims 2-4

22. Pursuant to step 2A, part 1, claims 2-4 depend on claim 1 and therefore recites the same abstract idea. Pursuant to step 2A, part 2, claims 2-4 recites the additional limitations of the metadata comprises a title, date, or information indicating one or more parties to the document, wherein the second LLM generate the metadata based on a predetermined portion of the document, and the first prompt is to generate a topic aware outline of the document. These limitations do not provide meaningful limits on the abstract idea because they merely describe extra solution activity or features, such as type of metadata, where a person should look to generate metadata, and that the outline is a topic aware outline. Therefore, these additional limitations do not integrate the abstract idea into a practical application. Pursuant to step 2B, the additional limitations do not amount to significantly more than the abstract idea because the limitations are not recited in a manner that provides improvements to the functioning of a computer or any other technology or technical field.

23. Claim 5

24. Pursuant to step 2A, part 1, claim 5 depends on claim 4 and therefore recites the same abstract idea. Pursuant to step 2A, part 2, claim 5 recites the additional limitations of submitting the document to a fourth LLM with a 4th prompt to generate a numerical outline of the document, receiving it, and submitting it with the query, metadata, and topic outline to the 3rd LLM. These limitations describe limitations that can be performed by a person with the aid of a computer as a tool, which in this case, is an LLM. Here, the 4th LLM is recited at a high level of generality and does not provide meaningful limits on the abstract idea. Therefore, these additional limitations do not integrate the abstract idea into a practical application. Pursuant to step 2B, the additional limitations do not amount to significantly more than the abstract idea because the limitations are not recited in a manner that provides improvements to the functioning of a computer or any other technology or technical field.

25. Claim 6

26. Pursuant to step 2A, part 1, claim 6 depends on claim 5 and therefore recites the same abstract idea. Pursuant to step 2A, part 2, claim 6 recites the additional limitations of submitting the query to a 5th LLM with a 5th prompt to transform the query into outline requests for each of the topic aware outline and the numerical outline, receiving the outline requests, and including them in the 1st and 4th prompts respectively. These limitations merely describe using a computer as a tool to convert a query into a format that may be used for each of the first and fourth LLMs respectively. However, the additional elements are recited at a high level of generality that does not provide meaningful limits on the abstract idea. Therefore, these additional limitations do not integrate the abstract idea into a practical application. Pursuant to step 2B, the additional limitations do not amount to significantly more than the abstract idea because the limitations are not recited in a manner that provides improvements to the functioning of a computer or any other technology or technical field.

27. Claim 7

28. Pursuant to step 2A, part 1, claim 7 depends on claim 6 and therefore recites the same abstract idea. Pursuant to step 2A, part 2, claim 7 recites the additional limitations of wherein correlating comprises submitting relevant document sentences and the selected sentence from the response to a 6th LLM with a prompt to rank each sentence based on relevancy. These limitations merely recite steps that can be performed by a person with the aid of a computer as a tool, such as comparing some sentences from the document to a response sentence and ranking them by how similar they are. Therefore, these additional limitations do not integrate the abstract idea into a practical application. Pursuant to step 2B, the additional limitations do not amount to significantly more than the abstract idea because the limitations are not recited in a manner that provides improvements to the functioning of a computer or any other technology or technical field.

29. Claim 8

30. Pursuant to step 2A, part 1, claim 8 depends on claim 7 and therefore recites the same abstract idea. Pursuant to step 2A, part 2, claim 8 recites the additional limitations of performing the relevancy determination based on embeddings of the sentences. These limitations merely recite steps that can be performed by a person with the aid of a computer as a tool, which can be used to create an embedding of a sentence and perform vector distances between them to determine relevancy scores. Therefore, these additional limitations do not integrate the abstract idea into a practical application. Pursuant to step 2B, the additional limitations do not amount to significantly more than the abstract idea because the limitations are not recited in a manner that provides improvements to the functioning of a computer or any other technology or technical field.

31. Claim 9

32. Pursuant to step 2A, part 1, claim 9 depends on claim 8 and therefore recites the same abstract idea. Pursuant to step 2A, part 2, claim 9 recites the additional limitations of determining relevancy of the sentences based on vector distance of the embeddings. As stated above, these limitations merely recite steps that can be performed by a person with the aid of a computer as a tool, which can be used to create an embedding of a sentence and perform vector distances between them to determine relevancy scores. Therefore, these additional limitations do not integrate the abstract idea into a practical application. Pursuant to step 2B, the additional limitations do not amount to significantly more than the abstract idea because the limitations are not recited in a manner that provides improvements to the functioning of a computer or any other technology or technical field.

33. Claim 10

34. Pursuant to step 2A, part 1, claim 10 depends on claim 7 and therefore recites the same abstract idea. Pursuant to step 2A, part 2, claim 10 recites the additional limitations of the LLMs are selected based on one or more of latency, maximum context window size, accuracy of results, quality of results, or resource usage. These limitations merely describe selecting computer components based on

performance or specific attributes. Such decisions can be made by a person when reviewing possible choices. Therefore, these additional limitations do not integrate the abstract idea into a practical application. Pursuant to step 2B, the additional limitations do not amount to significantly more than the abstract idea because the limitations are not recited in a manner that provides improvements to the functioning of a computer or any other technology or technical field.

35. Claim 11

36. Pursuant to step 2A, part 1, claim 11 depends on claim 1 and therefore recites the same abstract idea. Pursuant to step 2A, part 2, claim 11 recites the additional limitations of using a 7th LLM to classify the query into a first query classification of two possible classifications, determining a 7th prompt based on the first query classification to prompt the 3rd LLM based on the metadata and the document outline and provide the 7th prompt to the third LLM. These limitations do not provide meaningful limits on the abstract idea because a person can classify a query based on what type of query it is and accordingly, mentally determine what type of query is being asked to best get a response. Therefore, these additional limitations do not integrate the abstract idea into a practical application. Pursuant to step 2B, the additional limitations do not amount to significantly more than the abstract idea because the limitations are not recited in a manner that provides improvements to the functioning of a computer or any other technology or technical field.

37. Claims 12 and 13

38. Pursuant to step 2A, part 1, claims 12 and 13 depend on claim 7 and therefore recites the same abstract idea. Pursuant to step 2A, part 2, claims 12 and 13 recites the additional limitations of wherein two or more of the LLMs are the same or different, which does not provide meaningful limits on the abstract idea. Therefore, these additional limitations do not integrate the abstract idea into a practical application. Pursuant to step 2B, the additional limitations do not amount to significantly more than the

abstract idea because the limitations are not recited in a manner that provides improvements to the functioning of a computer or any other technology or technical field.

39. Claim 14

40. Pursuant to step 2A, part 1, claim 14 depends on claim 1 and therefore recites the same abstract idea. Pursuant to step 2A, part 2, claim 14 recites the additional limitations of the document is a plurality of documents. This limitation does not provide meaningful limits on the abstract idea. Therefore, these additional limitations do not integrate the abstract idea into a practical application. Pursuant to step 2B, the additional limitations do not amount to significantly more than the abstract idea because the limitations are not recited in a manner that provides improvements to the functioning of a computer or any other technology or technical field.

41. **Claims 15-17** recite essentially the same subject matter as claims 1-3, respectively, in the form of a computer readable storage media. Therefore, they are rejected for the same reasons.

42. **Claims 18-20** recite essentially the same subject matter as claims 1-3, respectively, in the form of a system. Therefore, they are rejected for the same reasons.

43. **Claims 1-20** are therefore not drawn to eligible subject matter as they are directed to an abstract idea without significantly more.

44. To expedite a complete examination of the instant application, the claims rejected under 35 U.S.C. 101 (nonstatutory) above are further rejected as set forth below in anticipation of applicant amending these claims to overcome the rejection.

Note on Prior Art Rejections

45. In the event the determination of the status of the application as subject to AIA 35 U.S.C. 102 and 103 (or as subject to pre-AIA 35 U.S.C. 102 and 103) is incorrect, any correction of the statutory

basis for the rejection will not be considered a new ground of rejection if the prior art relied upon, and the rationale supporting the rejection, would be the same under either status.

Claim Rejections - 35 USC § 103

46. The following is a quotation of 35 U.S.C. 103 which forms the basis for all obviousness rejections set forth in this Office action:

> A patent for a claimed invention may not be obtained, notwithstanding that the claimed invention is not identically disclosed as set forth in section 102 of this title, if the differences between the claimed invention and the prior art are such that the claimed invention as a whole would have been obvious before the effective filing date of the claimed invention to a person having ordinary skill in the art to which the claimed invention pertains. Patentability shall not be negated by the manner in which the invention was made.

47. The factual inquiries set forth in *Graham v. John Deere Co.*, 383 U.S. 1, 148 USPQ 459 (1966), that are applied for establishing a background for determining obviousness under 35 U.S.C. 103 are summarized as follows:

1. Determining the scope and contents of the prior art.

2. Ascertaining the differences between the prior art and the claims at issue.

3. Resolving the level of ordinary skill in the pertinent art.

4. Considering objective evidence present in the application indicating obviousness or nonobviousness.

48. This application currently names joint inventors. In considering patentability of the claims the examiner presumes that the subject matter of the various claims was commonly owned as of the effective filing date of the claimed invention(s) absent any evidence to the contrary. Applicant is advised of the obligation under 37 CFR 1.56 to point out the inventor and effective filing dates of each claim that was not commonly owned as of the effective filing date of the later invention in order for the examiner to consider the applicability of 35 U.S.C. 102(b)(2)(C) for any potential 35 U.S.C. 102(a)(2) prior art against the later invention.

49. Claims 1-20 are rejected under 35 U.S.C. 103 as being unpatentable over Gupta (US Patent Pub 2022/0092097) in view of Madisetti et al. (US Patent 12,001,462) (Madisetti).

50. In regards to **claim 1**, Gupta discloses a method comprising:

a. receiving, from a user, a query related to a document (Gupta at para. 0034)[1];

b. submitting the document to the system to generate an outline of the document (Gupta at paras. 0024-25)[2];

c. receiving a document outline of the document from the system (Gupta at para. 0024-25);

d. submitting the document to generate metadata of the document (Gupta at paras. 0032-34)[3]

e. receiving document metadata of the document from the system (Gupta at paras. 0032-34);

f. submitting at least a portion of the query, the document metadata, and the document outline to the system to generate a natural language response to the query based at least in part on the document metadata and the document outline (Gupta at para. 0034)[4];

g. receiving the natural language response from the system (Gupta at para. 0034);

i. providing the natural language response to the user. Gupta at paras. 0034, 0047.

51. Gupta does not expressly disclose submitting the document to a first LLM with a first prompt to generate an outline of the document and using a second LLM with a second prompt to generate metadata of the document and a third LLM with a third prompt to generate a natural language response to the query, correlating a selected sentence from the natural language response to a sentence from the

[1] A user submits a question (i.e., query) related to a document
[2] Documents are submitted to the submit to generate outlines for them.
[3] The system extracts entities of the document, such as dates and important clauses (i.e., metadata).
[4] The extracted entity information and outline are used by the system to generate an answer to a question in the given context.

document and providing an indication that the selected sentence from the natural language response is correlated to at least one sentence from the document. Gupta does disclose using an artificial intelligence with natural language processing capabilities to generate the outline of the document and answer questions about the document in natural language. Gupta at paras. 0004, 0034. What is not specifically disclosed is using multiple LLMs to perform these steps, correlating a selected sentence from the natural language response to a sentence from the document and providing an indication that the selected sentence from the natural language response is correlated to at least one sentence from the document.

52. Madisetti discloses a system and method that utilizes a plurality of LLMs for processing documents, such as knowledge documents, and providing answers to queries pertaining to the knowledge documents. In response to a user prompt (i.e., query), the system performs similarity analysis using prompt embeddings (i.e., correlating) to determine the best context aware answers to return to the user along with citations from the knowledge documents (i.e., providing an indication ...). Madsetti at col. 7, lines 30-67; col. 8, lines 1-7.

53. Gupta and Madisetti are analogous art because they are both directed to the same field of endeavor of artificial intelligence processing of documents.

54. At the time before the effective filing date of the instant application, it would have been obvious to one of ordinary skill in the art to modify Gupta by adding the features of submitting the document to a first LLM with a first prompt to generate an outline of the document and using a second LLM with a second prompt to generate metadata of the document and a third LLM with a third prompt to generate a natural language response to the query, correlating a selected sentence from the natural language response to a sentence from the document and providing an indication that the selected sentence from the natural language response is correlated to at least one sentence from the document, as disclosed by Madisetti.

55. The motivation for doing so would have been to provide more accuracy with responses. Madisetti at col. 7, lines 30-54.

56. In regards to **claim 2**, Gupta in view of Madisetti discloses the method of claim 1, wherein the document metadata comprises a document title, a document date, or information indicating one or more parties party to the document. Gupta at para. 0032.[5]

57. In regards to **claim 3**, Gupta in view of Madisetti discloses the method of claim 1, wherein the second large language model generates the document metadata based on a predetermined portion of the document. Gupta at para. 0034.[6]

58. In regards to **claim 4**, Gupta in view of Madisetti discloses the method of claim 1, wherein the first prompt prompts the first large language model to generate a topic aware outline of the document. Gupta at para. 0025.[7]

59. In regards to **claim 5**, Gupta in view of Madisetti discloses the method of claim 4, but does not expressly disclose further comprising: submitting the document to a fourth large language model along with a fourth prompt prompting the fourth LLM to generate a numerical outline of the document, receiving the numerical outline of the document, and submitting the numerical outline along with the query, the document metadata, and the topic aware outline to the third LLM. In particular, Gupta in view of Madisetti does not expressly disclose a generating a "numerical outline" of the document.

60. As set forth in the rejection of claim 1 above, Gupta discloses using artificial intelligence with natural language processing to generate outlines of submitted documents in order to provide answers

[5] Extracted entities include important dates (i.e., a document date).
[6] The system searches a document outline with the hierarchical structure for entities based on a predefined configuration to search a section (i.e., predetermined portion of the document).
[7] The outline is based on topics detected in the document (i.e., topic aware document).

to an input question. In doing so, Gupta in view of Madisetti discloses feeding LLMs a document outline, the query, and any additional metadata context in order to generate a response for the user's query.

61. A numerical outline seems to be an outline of a document's numerical values. Although it is unclear based on Applicant's specification. Given this interpretation, Gupta discloses extracting entities, which can include important dates. Gupta at para. 0032. Dates are often represented with numbers. Since Gupta discloses extracting dates of a document and producing an outline of a document based on extracted information, at the time before the effective filing date of the instant application, it would have been obvious to one of ordinary skill in the art to modify Gupta in view of Madisetti by adding the features of submitting the document to a fourth large language model along with a fourth prompt prompting the fourth LLM to generate a numerical outline of the document, receiving the numerical outline of the document, and submitting the numerical outline along with the query, the document metadata, and the topic aware outline to the third LLM. The motivation for doing so would have been to increase the accuracy of a response. Madisetti at col. 7, lines 30-45.

62. In regards to **claim 6**, Gupta in view of Madisetti discloses the method of claim 5, further comprising:

a. submitting the query to a fifth LLM along with a fifth prompt prompting the fifth LLM to transform the query to a first outline request for generate the topic aware outline of the document and into a second outline request for generating the numerical outline of the document (Madisetti at col. 7, lines 30-54);

b. receiving the first outline request and the second outline request from the fifth LLM (Madisetti at col. 7, lines 30-54);

c. including the first outline request in the first prompt (Madisetti at col. 7, lines 30-6; col. 8, lines 1-6); and

d. including the second outline request in the fourth prompt. (Madisetti at col. 7, lines 30-6; col. 8, lines 1-6)[8]

63. In regards to **claim 7**, Gupta in view of Madisetti discloses the method of claim 6, wherein correlating the selected sentence from the natural language response to a sentence from the document comprises submitting relevant document sentences and the selected sentence from the natural language response to a sixth LLM along with a sixth prompt prompting the sixth LLM to rank each relevant document sentence based on a relevance of the relevant document sentence to the selected sentence from the natural language response. Madisetti at col. 7, lines 55-67; col. 8, lines 1-7.[9]

64. In regards to **claim 8**, Gupta in view of Madisetti discloses the method of claim 7, wherein relevance of the relevant document sentence to the selected sentence from the natural language response is determined based on embeddings of the relevant document sentence and the selected sentence from the natural language response. Madisetti at col. 7, lines 55-67.[10]

65. In regards to **claim 9**, Gupta in view of Madisetti discloses the method of claim 8, wherein determining relevance based on embeddings of the relevant document sentence and of the selected sentence of the natural language response comprises determining relevance based on vector distance between embeddings of the relevant document sentence and of the selected sentence. Madisetti at col. 7, lines 55-67.[11]

66. In regards to **claim 10**, Gupta in view of Madisetti discloses the method of claim 7, wherein the first LLM, the second LLM, the third LLM, the fourth LLM, the fifth LLM, and the sixth LLM are selected

---

[8] Madisetti discloses deriving prompts based on a user's input prompt. This is interpreted as transforming the query into various prompts for appropriate LLMs, such as for creating a first outline and for creating a second outline.

[9] Relevant parts of a document are ranked by the LLM based on the query prompt, which is processed to create derived prompts. The relevant citations in the knowledge document are ranked to determine the best result to the user.

[10] Prompts (i.e., sentences) are converted to embeddings and are processed for similarity with reference to the user's prompt (i.e., selected sentence...).

[11] The similarity between embeddings is interpreted as using vector distance since the embeddings are vectors.

based on one or more of latency, maximum context window size, accuracy of results, quality of results or resource usage. Madisetti at col. 3, lines 30-37.[12]

67. In regards to **claim 11**, Gupta in view of Madisetti discloses the method of claim 1, further comprising:

a. using a seventh LLM to classify the query into a first query classification of at least two possible query classifications (Madisetti at col. 7, lines 30-54);

b. determining a seventh prompt, based on the first query classification, for prompting the third LLM to generate a natural language response to the query based at least in part on the document metadata and the document outline (Madisetti at col. 7, lines 30-54); and

c. providing the seventh prompt to the third LLM. (Madisetti at col. 7, lines 30-54).[13]

68. In regards to **claim 12**, Gupta in view of Madisetti discloses the method of claim 7, wherein two or more of the first LLM, the second LLM, the third LLM, the fourth LLM, the fifth LLM, and the sixth LLM are a same LLM. Madisetti at col. 6, lines 53-67; col. 1, lines 1-12.[14]

69. In regards to **claim 13**, Gupta in view of Madisetti discloses the method of claim 7, wherein two or more of the first LLM, the second LLM, the third LLM, the fourth LLM, the fifth LLM, and the sixth LLM are different LLMs. Madisetti at col. 6, lines 53-67; col. 1, lines 1-12.[15]

70. In regards to **claim 14**, Gupta in view of Madisetti discloses the method of claim 1, wherein the document is a plurality of documents. Gupta at para. 0024.[16]

---

[12] LLMs are chosen based on levels of accuracy.

[13] Madisetti discloses deriving prompts from a user's query. The prompts are derived based on different categories (i.e., at least two possible query classifications). Each LLM receives a different type of prompt and thus requires a different derived prompt. The system determines which prompt to derive based on the user's query (i.e., classify the query into a first query classification; determining a seventh prompt based on the first query classification).

[14] Various LLMs in the hLLM family can be used. Each can be trained on different training sets or not. This makes two or more the same or different.

[15] Various LLMs in the hLLM family can be used. Each can be trained on different training sets or not. This makes two or more the same or different.

[16] A plurality of documents can be uploaded to the system.

71. **Claims 15-18** are essentially the same as claims 1-3, respectively, in the form of a computer readable storage media. Gupta at para. 0030. Therefore, they are rejected for the same reasons.

72. In regards to **claim 18**, Gupta discloses a system comprising:

a. one or more computers and one or more storage devices on which are stored instructions that are operable, when executed by the one or more computers, to cause the one or more computers to perform operations (Gupta at para. 0030) comprising:

i. receiving, from a user, a query related to a document (Gupta at para. 0034)[17];

ii. submitting the document to the system to generate an outline of the document (Gupta at paras. 0024-25)[18];

iii. receiving a document outline of the document from the system (Gupta at para. 0024-25);

iv. submitting the document to generate metadata of the document (Gupta at paras. 0032-34)[19]

v. receiving document metadata of the document from the system (Gupta at paras. 0032-34);

vi. submitting at least a portion of the query, the document metadata, and the document outline to the system to generate a natural language response to the query based at least in part on the document metadata and the document outline (Gupta at para. 0034)[20];

vii. receiving the natural language response from the system (Gupta at para. 0034);

---

[17] A user submits a question (i.e., query) related to a document
[18] Documents are submitted to the submit to generate outlines for them.
[19] The system extracts entities of the document, such as dates and important clauses (i.e., metadata).
[20] The extracted entity information and outline are used by the system to generate an answer to a question in the given context.

viii. providing the natural language response to the user. Gupta at paras. 0034, 0047.

73. Gupta does not expressly disclose submitting the document to a first LLM with a first prompt to generate an outline of the document and using a second LLM with a second prompt to generate metadata of the document and a third LLM with a third prompt to generate a natural language response to the query, correlating a selected sentence from the natural language response to a sentence from the document and providing an indication that the selected sentence from the natural language response is correlated to at least one sentence from the document. Gupta does disclose using an artificial intelligence with natural language processing capabilities to generate the outline of the document and answer questions about the document in natural language. Gupta at paras. 0004, 0034. What is not specifically disclosed is using multiple LLMs to perform these steps, correlating a selected sentence from the natural language response to a sentence from the document and providing an indication that the selected sentence from the natural language response is correlated to at least one sentence from the document.

74. Madisetti discloses a system and method that utilizes a plurality of LLMs for processing documents, such as knowledge documents, and providing answers to queries pertaining to the knowledge documents. In response to a user prompt (i.e., query), the system performs similarity analysis using prompt embeddings (i.e., correlating) to determine the best context aware answers to return to the user along with citations from the knowledge documents (i.e., providing an indication ...). Madsetti at col. 7, lines 30-67; col. 8, lines 1-7.

75. Gupta and Madisetti are analogous art because they are both directed to the same field of endeavor of artificial intelligence processing of documents.

76. At the time before the effective filing date of the instant application, it would have been obvious to one of ordinary skill in the art to modify Gupta by adding the features of submitting the document to

a first LLM with a first prompt to generate an outline of the document and using a second LLM with a second prompt to generate metadata of the document and a third LLM with a third prompt to generate a natural language response to the query, correlating a selected sentence from the natural language response to a sentence from the document and providing an indication that the selected sentence from the natural language response is correlated to at least one sentence from the document, as disclosed by Madisetti.

77. The motivation for doing so would have been to provide more accuracy with responses. Madisetti at col. 7, lines 30-54.

78. **Claims 19 and 20** are essentially the same as claims 2 and 3, respectively, in the form of a system. Therefore, they are rejected for the same reasons.

Additional Prior Art

79. Additional relevant prior art are listed on the attached PTO-892 form. Some examples are:

80. Ozeran (US Patent Pub 2020/0335188) discloses a system and method for interrogating clinical documents.

81. Saleh et al (US Patent 10,885,436) discloses a system and method for training text summarization neural networks.

82. Raghavan et al. (US Patent Pub 2021/0056510) discloses a system and method for automation of corporate workflow processes via machine learning.

83. Zhang et al. (US Patent Pub 2021/0166014) discloses a system and method for generating a document summary using a neural network.

84. Tsai et al. (US Patent 11,182,545) discloses a system and method for machine learning on mixed data documents.

Conclusion

85. Any inquiry concerning this communication or earlier communications from the examiner should be directed to Examiner Michael Le whose telephone number is 571-272-7970 and fax number is 571-273-7970. The examiner can normally be reached Mon-Fri 9:30 AM – 6 PM.

86. Examiner interviews are available via telephone, in-person, and video conferencing using a USPTO supplied web-based collaboration tool. To schedule an interview, applicant is encouraged to use the USPTO Automated Interview Request (AIR) at http://www.uspto.gov/interviewpractice.

87. If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Tony Mahmoudi can be reached on 571-272-4078. The fax phone number for the organization where this application or proceeding is assigned is 571-273-8300.

88. Information regarding the status of published or unpublished applications may be obtained from Patent Center. Unpublished application information in Patent Center is available to registered users. To file and manage patent submissions in Patent Center, visit: https://patentcenter.uspto.gov. Visit https://www.uspto.gov/patents/apply/patent-center for more information about Patent Center and https://www.uspto.gov/patents/docx for information about filing in DOCX format. For additional questions, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free). If you would like assistance from a USPTO Customer Service Representative, call 800-786-9199 (IN USA OR CANADA) or 571-272-1000.

/MICHAEL LE/
Examiner, Art Unit 2163

/ALEX GOFMAN/
Primary Examiner, Art Unit 2163