Broader Search, Narrower Discovery: The Familiarity Bottleneck in AI-Augmented Knowledge Work*

Yuan Gao — Boston University

Dokyun Lee — Boston University

Daniel Sokol — University of Southern California

Mike Teodorescu — University of Washington

Preliminary and incomplete. Please do not distribute or cite.

Abstract

We study the introduction of the first AI search tool for patent examiners at the US Patent and Trademark Office (USPTO). This presents a very structured type of workplace, as examination procedures are codified in the Manual of Patent Examination Procedure (MPEP), productivity is measured by a simple count-based production system, and the workers are unionized federal employees insulated from AI displacement. This poses an interesting study of the utility of AI systems in a workplace where incentives are fixed and jobs are secure. Exploiting variation in patent claims amenability to text-based AI search in a difference-in-differences setup, together with novel data on examiner search behavior extracted from unstructured search records using vision-language models (VLMs), we do not find that AI changes the quantity of work: it reallocates attention in the search space to areas examiners may be less familiar with, but familiarity gates selection. Examiners rely on the majority of their citations on pre-existing knowledge. In heavily regulated workplaces, AI operates less as a technology driving productivity than as a knowledge allocation technology.

Data Note

Our newly curated dataset on examiners' prior art search behavior provides a comprehensive view of an important but long understudied part of the patent examination workflow: how examiners search for and explore prior art. It enables microlevel analyses of how examiners conduct iterative searches, identify relevant prior art, explore and shift across technology classes, draw on internal and external sources, and interact with different search technologies. The presentation will focus on the data collection process, research opportunities enabled by the new data, and preliminary findings from this paper. We plan to open-source the dataset after the project has progressed through peer review.

*We thank colleagues and seminar participants for helpful comments and discussions. All errors are our own.

Yuan Gao: Questrom School of Business, Boston University. Dokyun Lee: Questrom School of Business and Computing & Data Sciences, Boston University. Daniel Sokol: Gould School of Law and Marshall School of Business, University of Southern California. Mike Teodorescu: Information School, iStartup Lab, University of Washington. Correspondence should be addressed to Yuan Gao at yuangg@bu.edu.

1 Introduction

Which work can AI do, support, or change? A growing literature has been looking at task-level automation (Brynjolfsson and Mitchell, 2017; Brynjolfsson et al., 2018; Eloundou et al., 2023), which shows that up to four fifths of workers may have at least some tasks automated by AI. OTher literature shows that over a third of occupations use AI to *augment* tasks rather than automate tasks (Handa et al., 2025). Augmentation itself takes various forms, which can be classified by task complexity and locus of decision (Teodorescu et al., 2021). However, the literature has not yet looked at what changes occur in a workplace where the work product is highly regulated and jobs are very secure, such as the workplace of patent examiners. Given the recent introduction of AI Search for examiners at the USPTO, and our paper's first analysis of the data related to the AI search, we are equipped to answer the question in this paper.

Patent examination is a highly specialized, technically sophisticated, and economically important knowledge work in the United States. The USPTO employs more than 9,000 patent examiners, and the 2024 cohort of new examiners alone received more than 590,000 hours of formal training during their first year (USPTO, 2024, 2025b). The economic importance of patent examination is reflected in the broader role of patents in the U.S. economy. In 2024, utility-patent-intensive industries contributed approximately $5.6 trillion, or 23%, of U.S. private-sector GDP (USPTO, 2026).

Throughout the examination process, prior art search is a fundamental task, accounting for around half of examination time (USPTO, 2017). Examiners need to conduct in-depth searches across a vast knowledge base spanning domestic and foreign patents and non-patent literature to identify prior art relevant to the focal application. Based on this prior art, examiners then write office actions assessing whether the application satisfies the requirements of novelty under Section 102 and non-obviousness under Section 103. Incomplete or ineffective prior art search can allow low-quality patents to slip through examination, weakening innovation incentives, limiting competition, and increasing opportunistic litigation (Galasso and Schankerman, 2015; Frakes and Wasserman, 2015; Farrell and Shapiro, 2008). Beyond its role in evaluating the focal application, prior art search may also affect subsequent innovation, as examiner-cited prior art can serve as a channel of knowledge flow to later innovators (Alcacer and Gittelman, 2006).

To improve prior art search and the overall examination efficiency, the USPTO introduced a new AI search tool in late 2021 that uses an examiner-selected prior art as a seed to retrieve semantically similar documents. The tool therefore augments a complex task without moving the locus of decision: the examiner decides whether to use the citations provided. Despite substantial legal and practitioner discussion,[1] systematic evidence on AI search effectiveness and examiner search behavior remains limited by two data constraints: long lags between filing, examination, and disclosure, and the difficulty of accessing and structuring detailed search records. To address these limitations, we use a vision language model to extract and structure examiner search logs from tens of millions of PDF pages containing images, tables, and mixed layouts. We then link the resulting records to recently available examination outcomes for the earliest applications exposed to the new AI search tool. Our analysis suggests that AI search changes the composition, but not the amount, of observable search effort. Although AI broadens access to knowledge,

[1] https://www.gtlaw-techventureviews.com/2025/08/patent-examiner-ai-searches-required-ai-based-searching/

examiners focus more on familiar content.

Our findings speak to a growing new literature on AI's impact on productivity and knowledge work. On the one hand, this literature shows that AI speeds up software development (Cui et al., 2026), improves consultants' performance and speed (Dell'Acqua et al., 2026), and disproportionately benefits novices (Brynjolfsson et al., 2025). Writing tasks are particularly vulnerable to AI disruption (Noy and Zhang, 2023). While one line of research finds that AI increases knowledge-worker productivity, a separate stream shows that these gains depend on the setting. We contribute to this latter stream by showing that AI may not change how much work is done, but rather redirect how existing work is performed. Further, the gains from AI-augmented knowledge work may depend on workers' pre-existing knowledge. Put simply, what workers already know conditions what they can learn with AI.

2 Data and Institutional Background

Our sample consists of utility patent applications filed between 2019 and 2022 that reached a first office action on the merits (FAOM). This window allows sufficient time for examination outcomes to become observable given the substantial examination backlog, while reducing potential confounding from the subsequent diffusion of generative AI. Applications may ultimately be granted, abandoned, or remain pending, but we focus on FAOM decisions to keep the empirical setting clean. Although FAOM decisions do not determine the final disposition of an application, prior work suggests that they provide a relatively exogenous measure of examiners' initial beliefs (Farre-Mensa et al., 2020). Later stages of examination may introduce additional confounding factors, including applicant responses, rebuttals, requests for continued examination, and other forms of strategic behavior. Moreover, FAOM decisions strongly predict final examination outcomes (Farre-Mensa et al., 2020).

We also exclude patent applications with parent applications or related family members in the sample, as prior art searches for related applications may influence the focal search process in unobserved ways. The final sample contains 268,496 original utility patent applications examined by 8,641 examiners across 619 art units.

2.1 New Data

To further understand how examiners search for and explore prior art, and how they interact with the newly introduced AI search tool during the search process, we construct a new dataset on examiner prior art search behavior from unstructured USPTO search records. The USPTO requires examiners to document prior art searches in each application file (MPEP §719.05; §904). These records are primarily stored under two document codes available through the USPTO Open Data Portal: SRNT, which records search strategies and results, and SRFW, which records classifications, databases, and other search information.

These records contain detailed information on examiner search behavior, with an example provided in Appendix Fig. 1. Specifically, we observe query content, including keywords, CPC classes, and direct lookups of documents or inventors; the number of results returned; search sequence and timing; the databases searched, such as US patent databases or worldwide patent databases; and the search tools used.

These tools include both USPTO internal search systems, in which AI search is integrated and examiner use is explicitly recorded, and external tools such as Google Search.

A feature unique to the AI search records is that they preserve both the input to and the document-level output from each search. The AI search tool we study uses text embedding retrieval to identify documents most similar to a seed document specified by the examiner. We therefore observe both the seed document chosen by the examiner and the 20 documents ranked by the algorithm as the most relevant results. This level of visibility is not available for Boolean searches in the internal search system records. For these searches, we observe the query entered by the examiner and the number of documents returned, but not which documents within the returned set the examiner actually viewed. A specific document becomes observable only if the examiner subsequently uses that document as an explicit search query.

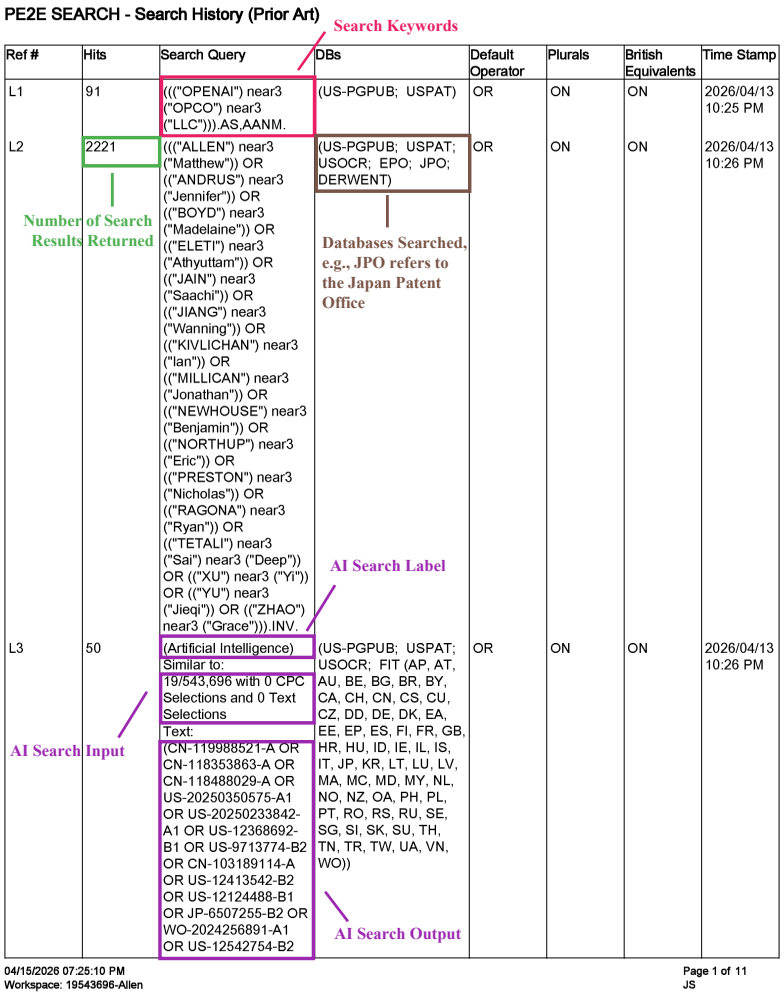


Figure 1: Example of PE2E Search Record Data. When examiners use USPTO internal search systems, the systems automatically record the search query, the number of results returned, the databases searched, and the timestamp.

Collecting these search records at scale is challenging because they are not stored in a standardized, machine-readable format. Searches conducted through USPTO internal systems automatically generate

image-based records that are relatively standardized and manageable using conventional OCR. Examiners also use external websites and commercial databases, including IP.com, IEEE Xplore, STN for chemical searches, and other domain-specific resources, such as specialized plant and horticultural databases like the UPOV Plant Variety Database. Use of these external resources is also required to be disclosed, but is often self-reported in highly unstructured formats, including screenshots, exported tables, and scanned notes.

To process these records at scale, we use GLM-OCR, a compact vision-language model specialized for document understanding, to extract and structure information from image-based PDFs (Duan et al., 2026). GLM-OCR is well suited to our setting because it performs strongly on complex document parsing, including text recognition, table reconstruction, and structured information extraction, while remaining computationally efficient. Despite its relatively small 0.9B parameter size, it achieves state-of-the-art performance across multiple benchmarks for complex document parsing.

Our final dataset consists of 500,724 utility patent applications filed between January 1, 2019, and December 31, 2022, together with the corresponding search activity conducted by assigned examiners from January 2019 through March 2026. Our observation period spans the transition from the Examiner's Automated Search Tool (EAST), introduced in 1999, to the web-based Patents End to End (PE2E) Search platform, whose rollout began in 2020. PE2E introduced two AI search tools during this transition: More Like This Document, the tool studied in this paper, in October 2021, and Similarity Search in September 2022.

We extract 22,008,059 query records from examiners' use of the USPTO's internal search system, including 21,361,795 prior art search queries and 646,264 interference search queries. Across both categories, applications contain an average of 44.0 recorded queries, with a median of 33, a 90th percentile of 84, and a maximum of 4,815. We also observe 68,974 applications, or 13.8% of the sample, in which examiners used external websites or other resources during prior art search. For this paper, our analysis focuses on search activity conducted within the USPTO's internal search systems.

We run GLM-OCR through the vLLM inference engine on a workstation equipped with two NVIDIA RTX 6000 Ada Generation GPUs, at a computational cost of more than 3,000 GPU hours. A more detailed data schema is provided in Appendix Table A1. We plan to conduct human validation of the model outputs and to open-source the resulting dataset after the project has progressed through peer review.

Moreover, we aim to deepen public understanding of the two newly introduced AI search tools and inform broader discussion of the implications and regulation of AI in patent systems. Public disclosures from the USPTO indicate that these models are trained on patent text, classifications, citations, human similarity ratings, and are updated regularly (USPTO, 2022a,b, 2025a). However, the USPTO provides limited information on the underlying training data, model architecture, and model updating process.

2.2 USPTO Institutional Records

To better understand the decision process in developing and adopting the AI search tool at the institutional level, one of the coauthors submitted two Freedom of Information Act requests (FOIA). The first request was centered on all training materials, slide decks, and related internal documents that are examiner-

facing. On September 22 2026, the agency released 333 pages of records pertinent to this request, which we summarize below.

The records comprise 8 slide decks and several QRFs (quick reference guides), as well as an FAQ. They are intended for examiner training use. The records describe both the tools used by the USPTO that are bundled in "AI search" - More Like This Document (MLTD) and the newer Similarity Search (the two tools are one year apart, MLTD began OCtober 2021 whereas Similarity Search October 2022). The materials, which we will provide snapshots of at the NBER i3 event, do not include anything about how the tools were designed, or what the architecture and training data might be. We do know that users see an ordered list showing each result's rank, CPC codes, and text passages that the model found similar to the queried document. Interestingly, the actual numeric similarity score is not available to the user. We do also know the contractor who implemented the tool is Google, but we do not have the technical specifications.

The information we have is limited to what is described, however we have filed another FOIA request requesting much more in depth technical details, including specifications and design documents, and more. We expect that second FOIA request to clear by December, in time for NBER i3. We will update this working paper with the additional information assuming we receive it before NBER i3.

3 AI Search and Patent Examination Outcomes

In our data, we can directly observe whether examiners use AI search in a focal application, but such adoption may be highly endogenous. Our baseline identification strategy, instead, exploits pre-treatment variation in claim format that creates differential exposure to text-based AI search. As shown in Figure 2, applications relying more on formulas, chemical structures, or other non-textual content are less amenable to text-based search.

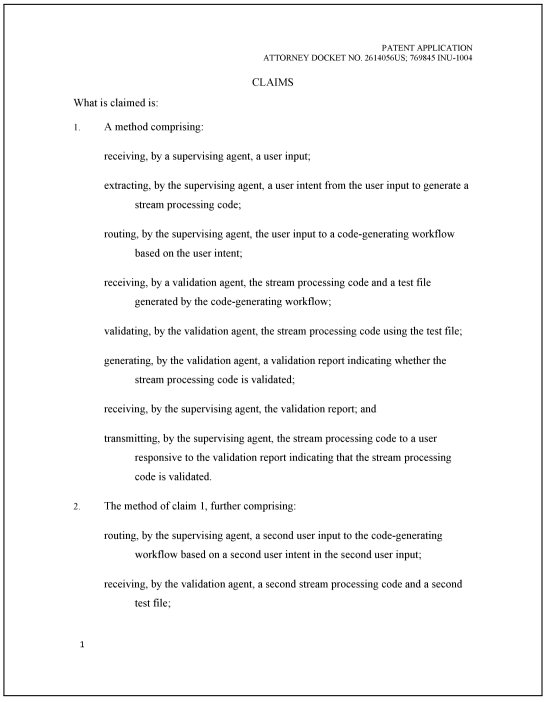
This variation is more plausibly exogenous because it is predetermined at the time of filing, before prior art search begins. Applications are also randomly assigned across examiners with different propensities to use AI search, and applicants cannot access the AI tool to strategically adjust their claims. That said, one potential concern is that sophisticated applicants or those represented by counsel may anticipate how claim format matters for prior art search using the new AI search tool. We plan to address this concern in future analyses by examining relevant subsamples, such as first-time versus experienced applicants and pro se versus represented applicants.

Appendix Figure 3 shows the distribution of exposure score overall and within selected CPC classes. About 90% of applications have fully textual claims, consistent with MPEP §2173.05(s), which generally requires utility patent claims to be self-contained and permits references to external material, such as figures and tables, only when there is *"no practical way to define the invention in words."* The remaining applications span a wide range of exposure scores. Although average exposure is lower in technologies such as organic chemistry than in computing and communications, substantial variation remains within technologies. For example, 26.7% of organic chemistry applications are fully textual, while roughly 10% of applications in computing and communications contain some non-textual claim content. Thus, the exposure measure captures meaningful within-technology variation rather than simply differences across

Table 1: Summary Statistics

	Mean	SD	Percentile 10th	Percentile 50th	Percentile 90th	*N*
Panel A. Examination Outcomes						
Non-final rejection at FAOM	0.801	0.399				234,640
Prior-art rejection share (§§102/103)	0.591	0.417	0.000	0.750	1.000	268,496
Non-prior-art rejection share (§§101/112)	0.246	0.410	0.000	0.000	1.000	268,496
Panel B. Prior-Art Evidence Supporting FAOM Rejections						
# prior art cited	3.622	2.579	1.000	3.000	7.000	201,804
Examiner-sourced share	0.770	0.148	0.500	0.800	0.933	201,898
Self-reuse share	0.016	0.092	0.000	0.000	0.000	201,804
Panel C. Examiner's Effort						
# search queries	80.31	70.54	20.00	60.00	164.0	268,348
Search time (minutes)	215.4	185.9	41.00	166.0	449.0	268,402
Filing to FAOM lag (months)	16.41	7.341	7.000	16.00	26.00	217,035
Examiner-initiated interview after first action	0.166	0.372				217,382
Panel D. AI Search Usage (Post-AI search launch Applications)						
Uses AI search	0.196	0.397				110,111
Subsequent search follows AI search output	0.124	0.330				110,111
Cites AI search output	0.088	0.283				110,111

Notes: Summary statistics for the variables used in the main identification. Columns report the mean, standard deviation, 10th, 50th, and 90th percentiles, and the number of nonmissing observations; percentiles are omitted for binary variables. Query count, search time, and filing to FAOM lag are winsorized at the 1st and 99th percentiles.



PATENT APPLICATION
ATTORNEY DOCKET NO. 2614056US; 769845 INU-1004

CLAIMS

What is claimed is:

1. A method comprising:

receiving, by a supervising agent, a user input;

extracting, by the supervising agent, a user intent from the user input to generate a stream processing code;

routing, by the supervising agent, the user input to a code-generating workflow based on the user intent;

receiving, by a validation agent, the stream processing code and a test file generated by the code-generating workflow;

validating, by the validation agent, the stream processing code using the test file;

generating, by the validation agent, a validation report indicating whether the stream processing code is validated;

receiving, by the supervising agent, the validation report; and

transmitting, by the supervising agent, the stream processing code to a user responsive to the validation report indicating that the stream processing code is validated.

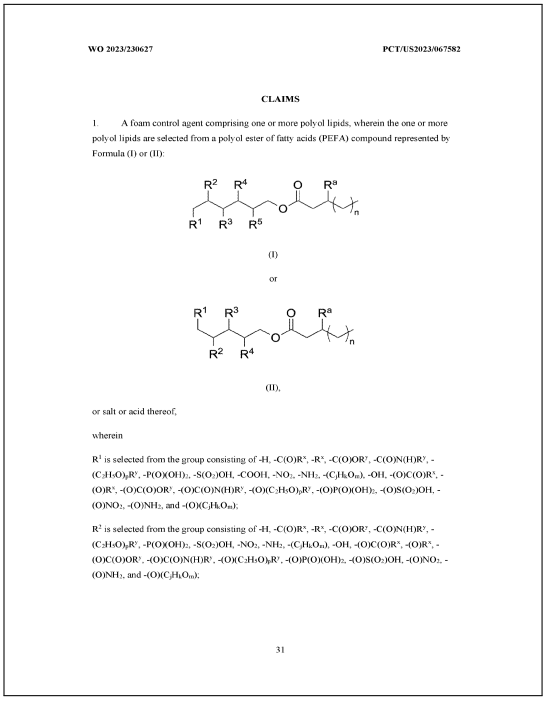
2. The method of claim 1, further comprising:

routing, by the supervising agent, a second user input to the code-generating workflow based on a second user intent in the second user input;

receiving, by the validation agent, a second stream processing code and a second test file;

1

(a) High Exposure



WO 2023/230627 PCT/US2023/067582

CLAIMS

1. A foam control agent comprising one or more polyol lipids, wherein the one or more polyol lipids are selected from a polyol ester of fatty acids (PEFA) compound represented by Formula (I) or (II):

(I)

or

(II),

or salt or acid thereof,

wherein

R^1 is selected from the group consisting of -H, -C(O)R^A, -R^A, -C(O)OR^7, -C(O)N(H)R^7, -$(C_2H_4O)_jR^7$, -P(O)(OH)$_2$, -S(O$_2$)OH, -COOH, -NO$_2$, -NH$_2$, -$(C_jH_kO_m)_n$, -OH, -(O)C(O)R^A, -(O)R^A, -(O)C(O)OR^A, -(O)C(O)N(H)R^A, -(O)$(C_2H_4O)_jR^7$, -(O)P(O)(OH)$_2$, -(O)S(O$_2$)OH, -(O)NO$_2$, -(O)NH$_2$, and -(O)$(C_jH_kO_m)_n$;

R^2 is selected from the group consisting of -H, -C(O)R^A, -R^A, -C(O)OR^7, -C(O)N(H)R^7, -$(C_2H_4O)_jR^7$, -P(O)(OH)$_2$, -S(O$_2$)OH, -NO$_2$, -NH$_2$, -$(C_jH_kO_m)_n$, -OH, -(O)C(O)R^A, -(O)R^A, -(O)C(O)OR^7, -(O)C(O)N(H)R^7, -(O)$(C_2H_4O)_jR^7$, -(O)P(O)(OH)$_2$, -(O)S(O$_2$)OH, -(O)NO$_2$, -(O)NH$_2$, and -(O)$(C_jH_kO_m)_n$;

31

(b) Low Exposure

Figure 2: Illustrative examples of application-level exposure for text-based AI search. Panel a shows a high-exposure case, while Panel b shows a low-exposure case with more visually structured content.

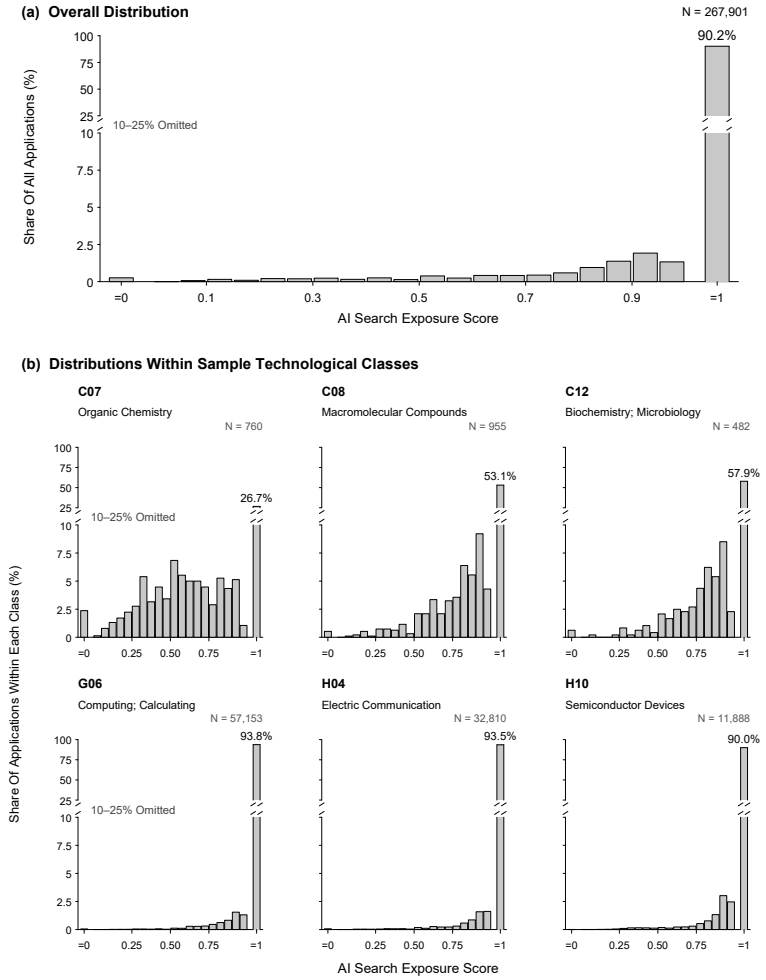


Figure 3: The AI search exposure score captures how readily an application's claims can be represented in text and searched by text-based AI. It equals one minus the share of claims containing nontextual features, including chemical or Markush notation, images, sequence information, biological deposit references, or numbered tables. A score of one indicates fully textual claims. Panel (a) shows the score distribution across 267,901 applications, while Panel (b) shows distributions for six selected CPC classes. C07, C08, and C12 are expected to be less amenable to text-based AI search, whereas G06, H04, and H10 are expected to be more amenable.

technological fields. Further, Table 2 shows that AI search exposure significantly predicts whether examiners use AI search, follow up on AI search-returned prior art, and cite AI search-returned prior art in the office action, validating the design of the AI search exposure score.

Formally, using repeated cross-sectional data at the patent application level, we estimate the following exposure-based difference-in-differences regression:

$$Y_i = \beta \left(Post_s \times Exposure_i \right) + \gamma Exposure_i + \mu_e + \delta_{a \times s} + \eta_{a \times f} + \varepsilon_i. \tag{1}$$

Here, Y_i is the examination outcome for application i. $Post_s$ equals 1 for searches after the October 2021 AI search tool rollout. $Exposure_i$ is the share of claims representable in text, with higher values indicating greater exposure to text-based AI search. The coefficient β captures the intent-to-treat effect of differential AI search exposure after the rollout. The specification includes examiner fixed effects, μ_e, art

Table 2: AI Search Exposure Predicts AI Search Use

	(1) Examiner used AI search	(2) Subsequent search follows AI search output	(3) Office action cited AI returned prior art
Post × Exposure	0.0185*	0.0154*	0.0217***
	(0.0103)	(0.0091)	(0.0074)
Observations	264,791	264,791	264,791
Adjusted R^2	0.404	0.295	0.217
Examiner FE	Yes	Yes	Yes
Art Unit × Search Month FE	Yes	Yes	Yes
Art Unit × Filing Year FE	Yes	Yes	Yes

Notes: Each column regresses an AI uptake indicator on Post × Exposure using the same specification as Equation (1). The regressions assess the exposure-based intent-to-treat design by testing whether greater exposure predicts higher AI uptake after the October 2021 launch. The three indicators equal one if the examiner used AI search, incorporated an AI-returned document into a subsequent search, or cited an AI-returned document in the office action, respectively. Standard errors in parentheses are clustered by examiner. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

unit × search-month fixed effects, $\delta_{a\times s}$, and art unit × filing-year fixed effects, $\eta_{a\times f}$, accounting for time-invariant examiner differences, technology-specific time shocks, and filing-cohort composition. Standard errors are clustered at the examiner level.

Table 3 reports our main results. We find that applications with greater exposure to AI search are significantly more likely to be rejected than to receive a direct allowance. However, these additional rejections are not supported by more evidence. Instead, the evidence base narrows: higher exposure applications rely more on prior art submitted by applicants rather than identified by examiners, and examiner-cited references are more likely to have been used previously rather than newly discovered. Although we do not assess examination quality, these patterns suggest less learning and discovery across cases, as examiners rely more heavily on previously accumulated knowledge when examining new applications.

These changes are not readily explained by changes in examiner productivity or effort. We find no change in observable effort, including the number of search queries, search time, and filing-to-FAOM lag, a proxy for examination efficiency. Examiner-initiated interviews, where examiners directly propose engaging with applicants to discuss claims and potential amendments, also remain unchanged. Because these interviews require additional examiner engagement beyond routine procedural work, they provide an additional measure of examination effort (deGrazia et al., 2021).

4 AI Search and Prior-Art Discovery

4.1 Descriptive Patterns

We next leverage our newly curated search log data to examine why AI search narrows the knowledge base examiners draw on. Our descriptive analysis focuses on applications in which examiners used AI search and traces the technological spaces surfaced by AI, measured at the CPC subgroup level, through subsequent searches and Office Action citations.

We focus on the technological spaces of AI-surfaced documents rather than the documents themselves mainly for four reasons. First, document-level search is only partially observable in search logs. For AI search, we observe both queries and returned documents, whereas conventional (keyword) search records

Table 3: AI Search Exposure and Examination Outcomes

Panel A. Examination Outcomes	(1) P (Non-final rejection)	(2) Prior-art rejection share (§§102/103)	(3) Non-prior-art rejection share (§§101/112)
Post × Exposure	0.0414***	0.0433***	−0.0061
	(0.0152)	(0.0145)	(0.0156)
Exposure	0.0270**	0.1209***	−0.1643***
	(0.0106)	(0.0104)	(0.0109)
Observations	230,872	264,791	264,791
Adjusted R^2	0.190	0.274	0.336
Panel B. Prior-Art Evidence Supporting FAOM Rejections	(4) # prior art cited	(5) Examiner-sourced share	(6) Self-reuse share
Post × Exposure	−0.1408	−0.0181***	0.0103**
	(0.0913)	(0.0060)	(0.0044)
Exposure	0.4836***	0.0359***	−0.0044
	(0.0599)	(0.0042)	(0.0032)
Observations	198,159	198,258	198,159
Adjusted R^2	0.485	0.386	0.071

Panel C. Examiner's Effort	(7) # search queries	(8) Search time (minutes)	(9) Filing to FAOM lag (months)	(10) Examiner-initiated interview
Post × Exposure	0.0691	5.175	0.0644	0.0123
	(1.940)	(5.047)	(0.1439)	(0.0143)
Exposure	6.570***	22.18***	0.0571	−0.0087
	(1.446)	(3.787)	(0.0899)	(0.0094)
Observations	264,645	264,700	213,361	213,705
Adjusted R^2	0.501	0.493	0.826	0.226
Examiner FE	Yes	Yes	Yes	Yes
Art Unit × Search Month FE	Yes	Yes	Yes	Yes
Art Unit × Filing Year FE	Yes	Yes	Yes	Yes

Notes: $Exposure_i$ is the share of claims representable in text; Post equals one for applications first searched after the October 2021 AI-tool launch. Panels A–C report FAOM outcomes, prior-art evidence, and examiner effort. Prior-art rejection share is the fraction of claims rejected under §102 or §103 in the first office action. Non-prior-art rejection share is the fraction of claims rejected under §101 or §112. Self-reuse denotes prior art cited by the examiner within the previous 24 months. All specifications include three fixed effects, with standard errors clustered by examiner. Sample size varies by outcome. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

capture only inputs; documents inspected from returned results are generally unobserved unless explicitly queried. Second, this level better reflects examiner search practice, where CPC-based searches are common and direct document lookup typically occurs only when a specific document is already known. Third, it captures exploration beyond adoption of the exact AI-suggested document, including cases in which AI redirects examiners toward previously unexplored technological areas. Fourth, technological space is a legally relevant construct: prior art supporting novelty rejections may lie outside the focal application's immediate area, while non-obviousness rejections may combine prior art across technological areas. Technological space therefore provides a more behaviorally and legally relevant measure of examiner knowledge discovery.

One explanation for the narrower prior art base is that AI surfaces little new or useful knowledge, leading examiners to rely on what they already know. Our descriptive evidence suggests otherwise. About 79.7% of AI-surfaced technological spaces are new to the examiner, indicating that AI expands the knowledge available for exploration. The narrowing instead appears to arise from how examiners select from

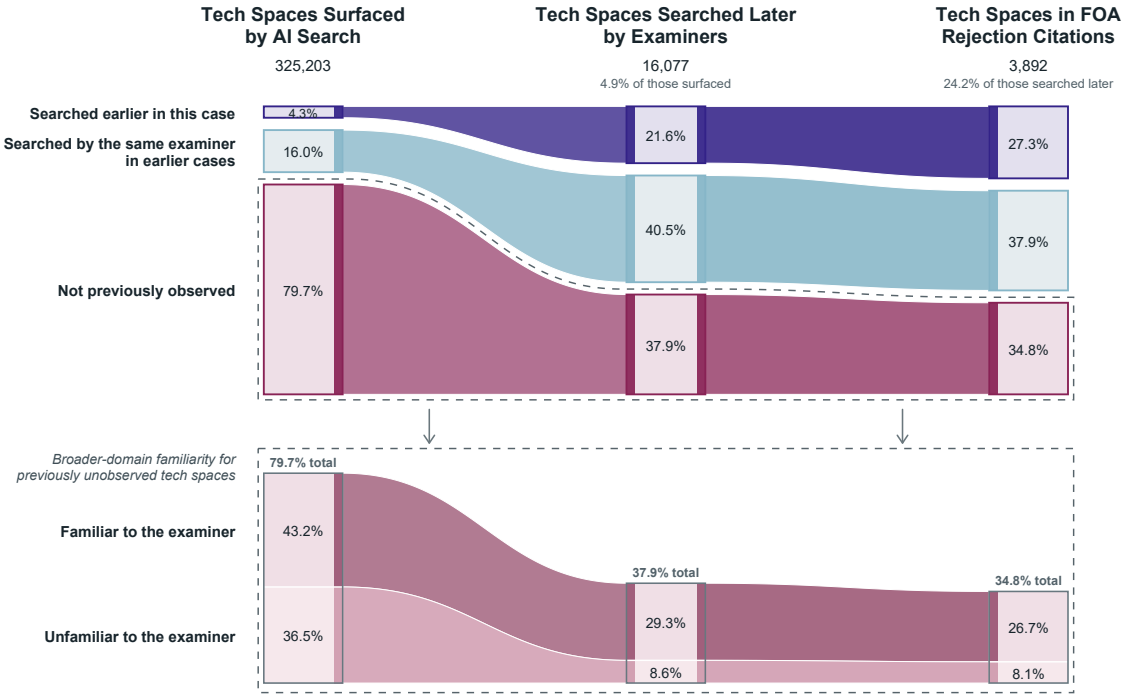


Figure 4: From AI-surfaced tech spaces to FAOM citations. The figure traces AI-surfaced CPC subgroups through examiner search and FAOM citations. The upper Sankey classifies subgroups by examiner familiarity and normalizes each stage to 100%. The lower Sankey further classifies previously unobserved subgroups by broader CPC subclass familiarity.

this broader set.

We identify two channels through which this selection occurs. First, through a *recall channel*, AI resurfaces knowledge examiners have encountered before and makes it readily available. Although previously searched technological spaces account for only 20.3% of those surfaced by AI, they account for 62.1% of subsequent search follow-up and 62.0% of final citation uptake. AI therefore serves partly as a recall technology, bringing previously accumulated knowledge back into the focal search. Second, through a *bounded discovery channel*, AI leads examiners into genuinely new technological spaces. These spaces account for 37.9% of subsequent search follow-up and 34.8% of final citation uptake, but remain concentrated within broader CPC subclasses already familiar to the examiner. The narrower evidence base therefore does not reflect a lack of new knowledge provision. Rather, AI expands the knowledge opportunity set, but familiarity gates its realization, directing examiners toward known knowledge and bounding the scope of new discovery.

4.2 Identification Strategy and Results

To examine these descriptive patterns more systematically, we turn to a second identification strategy at the search-query level. The key challenge is that search is sequential and path-dependent: what an examiner searches next depends on what has already been observed (Stigler, 1961; Weitzman, 1979). We therefore focus on the examiner's first use of AI search within an application and compare changes in search behavior before and after this event.

To construct a comparable control group, we use a two-step matching procedure. First, we exactly

match AI-using applications to non-AI applications within the same Art Unit, CPC subclass, filing year, and search year, to obtain applications with similar technological and examination environments. Second, within matched application pairs, we match on the pre-treatment search state. Simply assigning the same query position to the control application would be insufficient because searches at the same position may occur at very different stages of the search process. We therefore construct comparable pre- and post-periods based on three pretreatment characteristics (Li et al., 2001): search depth, measured by the number of queries already conducted; search scope, measured by the number of CPC subgroups already searched; and search trajectory, measured by the number of hits returned by the three most recent queries. We then assign the matched search state in the non-AI application as a pseudo-treatment point and compare the before-to-after change in search behavior around the actual AI use with the corresponding change around the pseudo-treatment point.

Our search-behavior outcomes focus on tech-space entry, defined as instances in which examiners actively shift the technological space of their search. We examine two margins: the incidence of tech-space entry and, conditional on entry, the novelty and distance of the technological space entered.

We consider the following difference-in-differences specification:

$$\bar{Y}_{apt} \;=\; \alpha_a \;+\; \lambda_p \,\mathrm{Post}_t \;+\; \beta\left(\mathrm{AI}_a \times \mathrm{Post}_t\right) \;+\; \varepsilon_{apt} \tag{2}$$

Here, a indexes applications in matched pair p, and $t \in \{\text{pre}, \text{post}\}$ indexes the analysis period. The pre period includes queries at $\tau \leq -3$, and the post period includes queries at $\tau \geq 1$, where τ is time relative to the pseudo onset. We exclude the AI search transition window $\tau \in \{-2, -1, 0\}$. Because AI search requires a seed document as input, the one or two preceding queries typically involve locating or opening that document before it is submitted to the AI search tool. The outcome $\bar{Y}_{apt}$ is the mean query level entry indicator over at risk queries. AI_a equals one for the adopting application. Application fixed effects and pair-by-period fixed effects compare the change in each treated application with that of its matched control. Standard errors are clustered at the examiner level.

Our results in Table **??** echo the earlier descriptive patterns. Compared with Boolean search, AI search leads examiners to enter more new technological spaces in the subsequent search process. These entries are not simply revisits to areas already searched within the focal application, but extend into spaces not previously explored in the current search. Panel B separates these entries into two channels. The *recall channel* captures spaces that are new to the focal search but previously familiar to the examiner. The *discovery channel* captures spaces that are new to both the focal search and the examiner, representing genuinely new discovery. Panel C further examines the distance of these entries from the examiner's recent search trajectory. AI-induced entries are more likely to reach more distant technological areas, including new CPC subgroups and main groups. However, they do not expand the broader boundary of search beyond familiar main groups and subclasses. Together, these results support a *bounded discovery channel*: AI induces exploration into new and more distant technological spaces, but largely within broader technological domains already familiar to the examiner.

Table 4: AI Search and Tech Space Entry: Matched Difference-in-Differences

Outcome	Estimate (AI × Post)	Baseline (Treated Pre)	Obs.	Adj. R^2
Panel A. Incidence of Entry				
P(Entry)	0.0141***	0.113	18,724	0.566
	(0.0043)			
Panel B. Novelty of Entry				
P(Entry, revisit)	−0.0018	0.011	18,724	0.236
	(0.0015)			
P(Entry, new to search, known to examiner)	0.0062**	0.053	18,724	0.434
	(0.0030)			
P(Entry, new to search, new to examiner)	0.0097***	0.050	18,724	0.396
	(0.0028)			
Panel C. Distance of Entry				
Relative to the last tech space				
P(new subgroup, same main group)	0.0101***	0.048	18,724	0.461
	(0.0027)			
P(new main group, same subclass)	0.0042*	0.035	18,724	0.337
	(0.0023)			
P(new subclass, same class)	−0.0001	0.008	18,724	0.185
	(0.0010)			
P(new class)	−0.0000	0.023	18,724	0.242
	(0.0020)			
Relative to all previously searched tech spaces				
P(main group never searched)	0.0038	0.053	18,724	0.411
	(0.0027)			
P(subclass never searched)	−0.0005	0.024	18,724	0.222
	(0.0018)			
P(class never searched)	−0.0007	0.018	18,724	0.131
	(0.0017)			

Notes: The unit of observation is application × period, with pre defined as $\tau \leq -3$ and post as $\tau \geq 1$. We exclude $\tau \in \{-2, -1, 0\}$ as the AI search transition window. Outcomes measure the share of at risk queries with at least one event of the stated type. Tech space refers to a CPC subgroup, where subgroup ⊂ main group ⊂ subclass ⊂ class. New to search indicates a tech space not previously observed in the focal search, while new to examiner indicates one not previously typed, pulled, or cited by the examiner. Regressions include application and pair by period fixed effects with equal pair weights. Pre period trends are jointly flat (Wald $p = 0.15$ to 0.17). Standard errors in parentheses are clustered by treated examiner. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

5 Discussion

We studied the first AI search tool deployed by the USPTO to aid patent examiners in their patent evaluations. We built a dataset and evaluated internal USPTO documents related to the deployment of AI search for patent examiners. The research fits in the broader emerging literature on the AI to augment or automate workplace tasks.

Our paper is the first to examine AI use in a highly structured setting where workers face little risk of displacement because their jobs are unionized and their work is regulated by federal law. We find that the AI search tool expands the opportunity set but does not commit the worker to a path and does not replace the worker. The count-based performance system for patent examiners provides little incentive to search far beyond their technology area, and our preliminary results are consistent with this constraint. Familiarity gates exploration, as examiners are more likely to rely on references and topics they have encountered and used before.

A potential explanation for what we observe is that While AI lowers the cost of search, the cost of evaluating citations rests with the examiner. Patent examiners still need to reason whether an AI surfaced citation outside of their main field of expertise is relevant or not. Second, incentives may not be aligned to encourage more use of the AI search tool (counts are the rewarded metric, not breadth of search and citations), and the examination job is not designed around the use of an AI. The findings are preliminary, and we expect to be able to present more USPTO institutional context by the December NBER i3 gathering.

References

Alcacer, J. and Gittelman, M. (2006). Patent citations as a measure of knowledge flows: The influence of examiner citations. *The Review of Economics and Statistics*, 88(4):774–779.

Brynjolfsson, E., Li, D., and Raymond, L. (2025). Generative ai at work. *The Quarterly Journal of Economics*, 140(2):889–942.

Brynjolfsson, E. and Mitchell, T. (2017). What can machine learning do? Workforce implications. *Science*, 358(6370):1530–1534.

Brynjolfsson, E., Mitchell, T., and Rock, D. (2018). What can machines learn and what does it mean for occupations and the economy? *AEA Papers and Proceedings*, 108:43–47.

Cui, K. Z., Demirer, M., Jaffe, S., Musolff, L., Peng, S., and Salz, T. (2026). The effects of generative ai on high-skilled work: Evidence from three field experiments with software developers. *Management Science*.

deGrazia, C., Pairolero, N. A., and Teodorescu, M. H. (2021). Examination incentives, learning, and patent office outcomes: The use of examiner's amendments at the uspto. *Research Policy*, 50(10):104360.

Dell'Acqua, F., McFowland III, E., Mollick, E., Lifshitz, H., Kellogg, K. C., Rajendran, S., Krayer, L., Candelon, F., and Lakhani, K. R. (2026). Navigating the jagged technological frontier: Field experimental evidence of the effects of artificial intelligence on knowledge worker productivity and quality. *Organization Science*, 37(2):403–423.

Duan, S., Xue, Y., Wang, W., Su, Z., Liu, H., Yang, S., and Tang, J. (2026). GLM-OCR technical report. arXiv preprint arXiv:2603.10910.

Eloundou, T., Manning, S., Mishkin, P., and Rock, D. (2023). GPTs are GPTs: An early look at the labor market impact potential of large language models. *arXiv preprint arXiv:2303.10130*.

Farre-Mensa, J., Hegde, D., and Ljungqvist, A. (2020). What is a patent worth? evidence from the u.s. patent "lottery". *The Journal of Finance*, 75(2):639–682.

Farrell, J. and Shapiro, C. (2008). How strong are weak patents? *American Economic Review*, 98(4):1347–1369.

Frakes, M. D. and Wasserman, M. F. (2015). Does the US patent and trademark office grant too many bad patents?: Evidence from a quasi-experiment. *Stanford Law Review*, 67:613–676.

Galasso, A. and Schankerman, M. (2015). Patents and cumulative innovation: Causal evidence from the courts. *The Quarterly Journal of Economics*, 130(1):317–369.

Handa, K., Tamkin, A., McCain, M., Huang, S., Durmus, E., Heck, S., Mueller, J., Hong, J., Ritchie, S., Belonax, T., Troy, K. K., Amodei, D., Kaplan, J., Clark, J., and Ganguli, D. (2025). Which economic tasks are performed with AI? Evidence from millions of Claude conversations. *arXiv preprint arXiv:2503.04761*.

Li, Y. P., Propert, K. J., and Rosenbaum, P. R. (2001). Balanced risk set matching. *Journal of the American Statistical Association*, 96(455):870–882.

Noy, S. and Zhang, W. (2023). Experimental evidence on the productivity effects of generative artificial intelligence. *Science*, 381(6654):187–192.

Stigler, G. J. (1961). The economics of information. *Journal of Political Economy*, 69(3):213–225.

Teodorescu, M. H. M., Morse, L., Awwad, Y., and Kane, G. C. (2021). Failures of fairness in automation require a deeper understanding of human–ML augmentation. *MIS Quarterly*, 45(3):1483–1500.

USPTO (2017). Examination time and the production system. USPTO presentation. Accessed 2026-05-11.

USPTO (2022a). New artificial intelligence functionality in PE2E search. Official Gazette of the United States Patent and Trademark Office. 1504 Off. Gaz. Pat. Office 359, published November 15, 2022.

USPTO (2022b). New PE2E search tool using AI search features. Official Gazette of the United States Patent and Trademark Office. Notice dated December 20, 2021; 1494 Off. Gaz. Pat. Office 251, published January

11, 2022.

USPTO (2024). Quality U.S. patents drive our economy and solve world problems. United States Patent and Trademark Office blog, December 9. https://www.uspto.gov/blog/quality-us-patents-drive-our-economy-and-solve-world-problems. Accessed September 7, 2026.

USPTO (2025a). Advance notice of change to the MPEP to provide guidance on the use of PE2E search's similarity search feature. Memorandum to the Patent Examining Corps. October 24, 2025.

USPTO (2025b). FY 2025 agency financial report. Technical report, United States Patent and Trademark Office. https://www.uspto.gov/sites/default/files/documents/USPTOFY2025AFR.pdf. Accessed September 7, 2026.

USPTO (2026). Intellectual property and the U.S. economy in 2024. Report and accompanying webpage. https://www.uspto.gov/ip-policy/economic-research/intellectual-property-and-us-economy. Accessed September 7, 2026.

Weitzman, M. L. (1979). Optimal search for the best alternative. *Econometrica*, 47(3):641–654.

Appendix

Table A1: Search-Record Data Schema

Variable	Type	Description
Panel A. Application Metadata		
Application ID	string	Links search records to patent outcomes (rejections, citations, grants, and appeals).
Filing date	date	Date the application was filed.
Filing year	numeric	Year the application was filed.
Examiner ID	string	Identifier for the examiner.
Art unit	string	Examiner's art unit.
Technology class	string	CPC class.
Panel B. Search Records (one row per query)		
Query text	string	Recorded search input, including document number lookups and CPC class inputs.
Query time	date-time	Date of the query.
Search type	string	Prior-art or interference search.
Databases searched	string	Database selected in the query.
Result count	numeric	Number of results returned by the system for a given search query
Search operator	string	Default operator (AND or OR).
Query order	numeric	Printed query sequence number.
Panel C. AI Search		
Uses AI search	binary	Any AI-search use detected for the application.
AI search count	numeric	Detected AI-search occurrences per application.
AI tool	string	More Like This Document (MLT) or Similarity Search (SimSearch).
Input ID	string	MLT anchor document ID or SimSearch input application ID.
CPC selection count	numeric	Number of CPC selections in SimSearch.
Text selection count	numeric	Number of text selections in SimSearch.
Returned documents	list	Recovered document IDs in recorded order.
Document rank*	numeric	Position in the AI search output (1 = first), indicating similarity-based relevance rank.
Returned document count	numeric	Number of returned documents in the AI search output
Panel D. Search Volume and Timing (application level)		
Total query count*	numeric	Prior-art and interference queries combined.
Prior-art query count*	numeric	Recorded prior-art queries.
Interference query count*	numeric	Recorded interference queries.
First search time	date-time	First timestamp identified in the application history.
Active search days	numeric	Days with recorded search timestamps.
Search span	numeric	Days between first and last retained timing points.
Search time (minutes)	numeric	Sum of within-day timestamp gaps.

Notes: Search records cover 500,724 patent applications. Variables come from linked application, query, and AI-output tables. Binary variables take values 0/1. Asterisks (*) mark derived measures computed from the query records rather than stored columns. AI-search usage and counts are detected from recorded search histories, not click tracking; selection counts do not include the selected passages or code identities. Query counts include EAST and PE2E. Search time estimates activity, not observed labor. Timing data end in December 2025; query records extend into 2026.