New artificial intelligence functionality in PE2E Search

Summary

The United States Patent and Trademark Office (USPTO or Office) recently added the new artificial intelligence (AI)-based "Similarity Search" feature to the Patents End-to-End (PE2E) search suite as a tool to further assist patent examiners in conducting a search of the prior art. When a search is performed using Similarity Search, it will be reflected in the search history recorded in the file wrapper.

Background

In the examination of a patent application, patent examiners conduct a thorough search of the prior art. See 37 CFR 1.104(a)(l). The Office provides examiners with access to a wide variety of search tools to perform an effective search of the prior art, including domestic and foreign patent documents and non-patent literature. One such search tool is the Patent Linguistic Utility Service (PLUS), a query by example search system for U.S. patents and U.S. pre-grant publications (PGPubs) from 1971 to the present. The PLUS search system receives as input a keyword list generated from scanned portions of the specification of an application being searched (e.g., title, abstract, background, and brief summary) and uses word frequency search technology to produce a list of U.S. patents and U.S. PGPubs that most closely match the application.

In an effort to modernize the Patents Automated Information Systems, the USPTO launched PE2E, a single web-based system that provides examiners with a unified and robust set of tools to use in the examination process. PE2E Search is a system within PE2E that presents a modern interface design and introduces new tools and features, such as AI search capabilities. PE2E Search contains both U.S. and foreign published patent documents. It is replacing legacy search tools, including Examiner's Automated Search Tool (EAST), Web-based Examiner Search Tool (WEST), and PLUS, which are set to be decommissioned on September 30, 2022.

In October 2021, the USPTO released the AI-based "More Like This Document" feature in PE2E Search to examiners. The Office continues to add new search features with AI capabilities to PE2E Search. The Office has also been proactively notifying the public of any new AI-based features in PE2E Search that are being made available to examiners. See, for example, the notice on the [New PE2E Search Tool Using AI Search Features (uspto.gov)](uspto.gov).

AI Similarity Search feature

The USPTO has introduced a new AI-based tool in PE2E Search: the Similarity Search feature. The Similarity Search feature receives examiner-selected application information, including the

specification, as input and uses trained AI models to output a list of domestic and foreign patent documents that are similar to the patent application being searched. The Similarity Search feature in PE2E Search is designed to be an enhanced replacement of PLUS and provides examiners with optional new search capabilities to access prior art alongside traditional document retrieval approaches. An examiner can refine the AI-based Similarity Search queries using Cooperative Patent Classification (CPC) symbols and text snippet selection capabilities. This provides the examiner with the flexibility to emphasize certain CPC symbols, paragraphs, sentences, or words to focus on specific concepts. Moreover, this new feature provides significantly faster results for examiners. AI-based Similarity Search queries are completed in seconds, whereas PLUS searches historically can take hours to generate.

AI models supporting PE2E Search are trained using publicly available patent data, including the disclosure text, patent classifications, document citations, and human-rated similarity. Potential model biases due to applicant, inventor, and assignee information are mitigated by excluding this information from the training data. Models are updated regularly to maintain the highest level of performance possible. To ensure data security and maintain patent application confidentiality as required by 35 U.S.C. 122(a), the USPTO has implemented National Institute of Standards and Technology (NIST) security controls[1] and runs search queries using this feature within USPTO system boundaries using a Federal Risk and Authorization Management Program authorized cloud. This is consistent with the National AI Initiative strategic pillar "Advancing Trustworthy AI" as well as Executive Order 13960, "Promoting the Use of Trustworthy Artificial Intelligence in the Federal Government," and achieves the goal of granting more robust and reliable patents by helping examiners find potentially relevant documents for consideration.

Search recordation

To provide a complete, accurate, and uniform record of what the examiner has searched and considered for each application, the USPTO has established procedures for recording Similarity Search data in the application file. When an examiner uses the Similarity Search feature of the PE2E search tool as part of a search for an application, the examiner's PE2E search history will indicate that AI Similarity Search capabilities were used so the public receives clear notice that aspects of the examiner's search were performed using AI. When an examiner selects a Similarity Search query to be included in the search notes of the application file wrapper, all documents retrieved by that query, along with the query itself, are listed in the search notes. This enhances public awareness of the search strategy and ensures that the model outputs are transparent and explainable. Measures are taken to ensure application confidentiality. The search results produced using emphasized text and/or CPC symbols for a given application will be listed in the PE2E search history in addition to the application number and any applicable emphasized text snippets and CPC symbols, except where the inclusion of the text or the application number would violate the confidentiality provision of 35 U.S.C. 122(a).

---

[1] At the time of this notice, the environment that the model runs on complies with NIST 800-53 controls for a Federal Information Processing Standards Publication 199 categorized system.

Contact information

You can submit inquiries and feedback concerning this notice, including comments and concerns regarding search recordation with respect to the emphasized text, to ai_for_pe2e@uspto.gov, or by mail addressed to Commissioner for Patents, Box Comments-Patents, Post Office Box 1450, Alexandria, VA 22313-1450, and marked to the attention of Jonathan Horner.