



**CAS CUSTOM SERVICES<sup>SM</sup>**

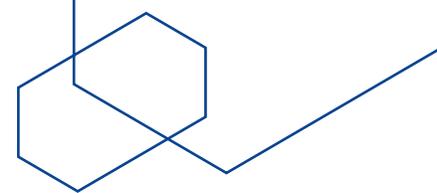
# **BETWEEN GROWTH AND EFFICIENCY, THERE IS INNOVATION**

## **How CAS's Unique Approach to AI Drives Patent Office Success**

Patent offices that are exploring AI solutions to improve prior art searches are discovering that project success relies on the use of highly structured data, multi-functional algorithms, and a team of cross-trained experts providing customized solutions.

CAS partners with over 30 worldwide patent offices and has developed AI-enabled solutions that improve application timeliness, patent quality, and productivity. Our approach has five unique advantages that drive its success.

# Five advantages of CAS AI solutions



## 1. Leveraging a century of data science leadership

As a leader in data science for over a century, CAS uniquely curates and structures data to improve search algorithm modeling and results. As technology advances, we refine this data and introduce new ways to access and retrieve it:

- In 2017, we introduced a multifactor relevance-ranking algorithm for search results in CAS SciFinder® to increase researcher productivity and accelerate scientific discovery.
- In 2020, we added an AI algorithm for predictive retrosynthesis to CAS SciFinder to identify the most efficient synthetic paths for target compounds.
- In 2021, we launched AI-powered prior art analysis capabilities in CAS STNext® and CAS SciFinder® to improve productivity and search insights.
- In 2021, we completed a custom AI-enabled workflow solution for INPI Brazil to cut examination times in half and help resolve its 10-year application backlog.

These and other recent CAS innovations put more data power into the hands of scientific and patent professionals to improve decision-making and increase productivity.

## 2. Structured data leads to more relevant search results

The more data machine learning algorithms can access, the more relevant, reliable, and trustworthy their results can be. However, for complex arts like chemistry and life sciences, publicly available data sets are of limited use because they often contain anomalies, such as transcription errors, or have keywords and structures embedded in unsearchable tables and images. CAS overcomes these challenges using proprietary, curated data for testing and production:

- The CAS Content Collection™, the world's largest collection of chemistry and life sciences data which is extracted, indexed, and linked to simplify access and retrieval of relevant information.
- A random sample of patents held out of algorithm training that we use as a control set for measuring the accuracy of search results and hit rates. These patents have been evaluated by examiners in patent offices in China, Japan, Europe, United States, and the World Intellectual Property Organization, and have been validated for relevance by our own team of IP search professionals.

## 3. Using multiple algorithm streams improves outcomes

The right combination of search algorithms and quality data will deliver results with the highest relevance. While some AI search solutions may rely on single algorithms based on semantics to perform similarity searches, more relevant results are generated by integrating multiple algorithms to perform searches. For instance, CAS runs machine learning against full-text patent documents, graph databases against our curated, indexed data, and similarity algorithms against the massive CAS Content Collection.

Each algorithm has its strengths. Machine learning is effective for searching text and indexed terms, but less effective for the composition of matter patents where important data is often contained inside structures. Likewise, a graph database can find similarities and connections that machine learning cannot. We also add an ensemble algorithm that analyzes the results from the search algorithms and arrives at a single, prioritized list of publications that are most likely to have conflicts with the target application.

#### 4. Mixed-mode subject matter experts increase project effectiveness

AI projects involve a complex interplay between data, technology, and workflows. The most effective solutions are customized to an organization's unique goals and processes and include team members who have cross-functional knowledge.

A typical AI-enabled prior art project brings together data scientists to identify and curate data, computational scientists to model and refine algorithms, patent searchers to validate AI results, software programmers to develop user interfaces, and project managers to keep everything aligned. AI projects in arts like chemistry and life sciences are more successful when these experts understand how the underlying science and data can affect outcomes.

CAS deploys mixed-mode teams where people have experience in multiple disciplines. Our data scientists who train algorithms are also chemists, our computational chemists creating machine learning models also understand data science, and our IP search professionals validating results have advanced science degrees. In this way, each member of the project team better anticipates challenges in these specialized areas and understands how to collaborate more effectively.

#### 5. Transparency of results saves time

It is not enough for AI-enabled searches to return relevant results. Organizations want to know how results are generated and why certain reference documents are returned.

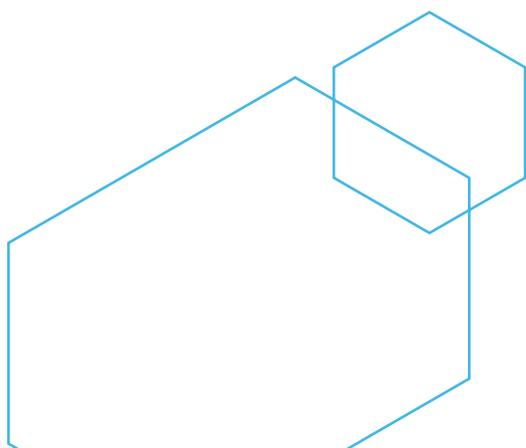
Patent office searchers and examiners need traceability to document decisions for prosecution and internal quality reviews. They need to know which terms, structures, and other attributes appear between documents and how relevancy rankings are derived.

CAS machine learning algorithms return documents in order of highest to lowest relevance. They also link each result to supporting references. Similar terms between the target application and other documents are highlighted, giving searchers a clear understanding of the logic behind the results. Everything is traceable.

#### Expertise you can trust

No two patent offices have the same needs, workflows, or resources. CAS Custom Services combines advanced data analytics, specialized technology, and deep patent expertise to create innovations aligned to each organization's business goals and initiatives.

**Learn more at [cas.org](https://cas.org)**





CAS is a leader in scientific information solutions, partnering with innovators around the world to accelerate scientific breakthroughs. CAS employs over 1,400 experts who curate, connect, and analyze scientific knowledge to reveal unseen connections. For over 100 years, scientists, patent professionals, and business leaders have relied on CAS solutions and expertise to provide the hindsight, insight, and foresight they need so they can build upon the learnings of the past to discover a better future. CAS is a division of the American Chemical Society.

**Connect with us at [cas.org](https://cas.org)**