

ADDRESSING SUSTAINABILITY OF THE GLOBAL PATENT SYSTEM: THE ROLE OF AI IN ENHANCING PRODUCTIVITY

CAS

A division of the
American Chemical Society



Executive summary

The sustainability of the global patent system is under pressure from the rapid growth in patent application volume and complexity. In countries seeing the fastest growth, resulting capacity gaps delay patent examination by years in some cases, put patent quality at risk, and threaten to slow the pace of innovation and investment.

Traditionally, patent offices have managed growth by hiring more examiners, but that approach increases the cost of intellectual property protection for all stakeholders and may not be sustainable as volume continues to increase and labor markets tighten. More recently, patent offices have also been exploring new approaches to increase examination efficiency, including the application of artificial intelligence (AI) and related workflow enhancements to expedite the most labor-intensive parts of patent examination, including prior art searches, and speed first office actions.

This whitepaper explores challenges and opportunities for patent offices as they seek to ensure sustainability and plan for future growth, with a focus on the application of AI-enabled workflow solutions to enhance productivity. These optimized workflow solutions have the potential to improve many patent office functions, including:

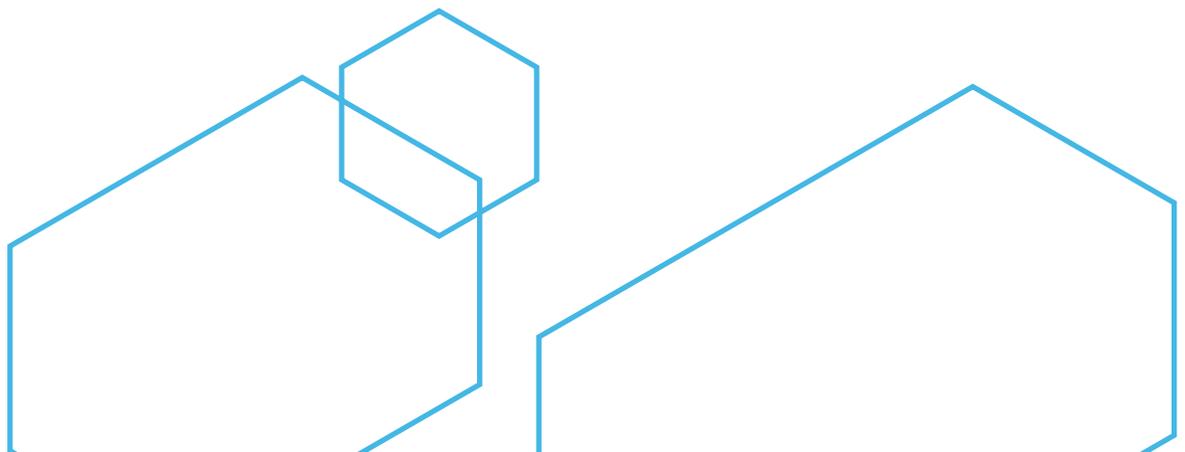
- Classification assignment and conversion
- APIs for efficient document delivery
- Online tools for reviews and analyses
- Fully formatted reporting, and more
- Outsourcing of specific tasks

Potential roadblocks in deploying AI solutions for these purposes are also addressed, such as insufficient data quality for training and operating algorithms, inefficient workflow integration for streamlining processes, and inadequate expertise to augment internal resources during projects.

The paper also includes key insights and learnings from a collaboration between CAS and the National Institute of Industrial Property (INPI) of Brazil to address their application backlog and improve examination workflow efficiency. The project deployed a custom workflow solution based on an AI-driven approach to achieve significant operational improvements, including:

- Up to 50% reduction in examination times
- 77% of all national applications processed required less examiner search time
- 29% of all national applications processed required little or no additional search
- Workloads were efficiently managed without adding staff
- Examiners were freed up to focus on other priorities
- Productivity improvements contributed to a reduction of 80% in the office's backlog

As innovation accelerates, worldwide patent offices continue to face unprecedented demands for speed and efficiency. Ensuring patent quality and minimizing patent pendency will remain challenging, requiring unique and innovative solutions. Key insights discussed in this whitepaper provide ideas for other patent offices seeking solutions to optimize workflow efficiency and ensure operational sustainability.





Challenge of managing growth

The world's patent offices share the same basic mission – to encourage innovation within their jurisdictions. To accomplish this mission, offices are pursuing long-term sustainability in all facets of their operations. Within patent prosecution, skilled examiners must carefully and accurately review applications and determine whether prior art exists that could invalidate claims. Inefficiencies in this process can quickly add up to higher operating costs, delays in first office actions, longer pendency, reduced patent quality, and in some cases, longer application backlogs.

These challenges are due, in part, to growing patent application volumes and complexity and the rapid proliferation of new technologies. The top five patent offices experienced a 5.1% compound annual growth rate in applications from 2011 to 2020,¹ despite a slowdown in 2020 from the COVID-19 pandemic.

Additionally, patent applications are becoming more complex to review, due to a growing number of claims per patent, patents cited per claim, inventors per patent, and prior art citations per patent.

In fact, scientific publications are growing at a rate of 8% a year, doubling every nine years. Patent complexity requires examiners to become familiar with expanding knowledge areas and to design increasingly sophisticated search strategies to uncover prior art.

Patent quality is a fundamental concern to IP stakeholders and can be adversely impacted by rapid application growth or insufficient examination capacity. Much attention has been given to improving application requirements and examination procedures to ensure granted patents can survive post-grant challenges to validity. However, patent quality is negatively impacted when examiners lack the time or expertise to effectively review prior art, communicate with applicants, and write thorough communications. Overburdening offices makes it more likely that examinations will be done by someone with less knowledge of the subject matter, which can reduce patent quality.

Approaches to improving capacity and throughput

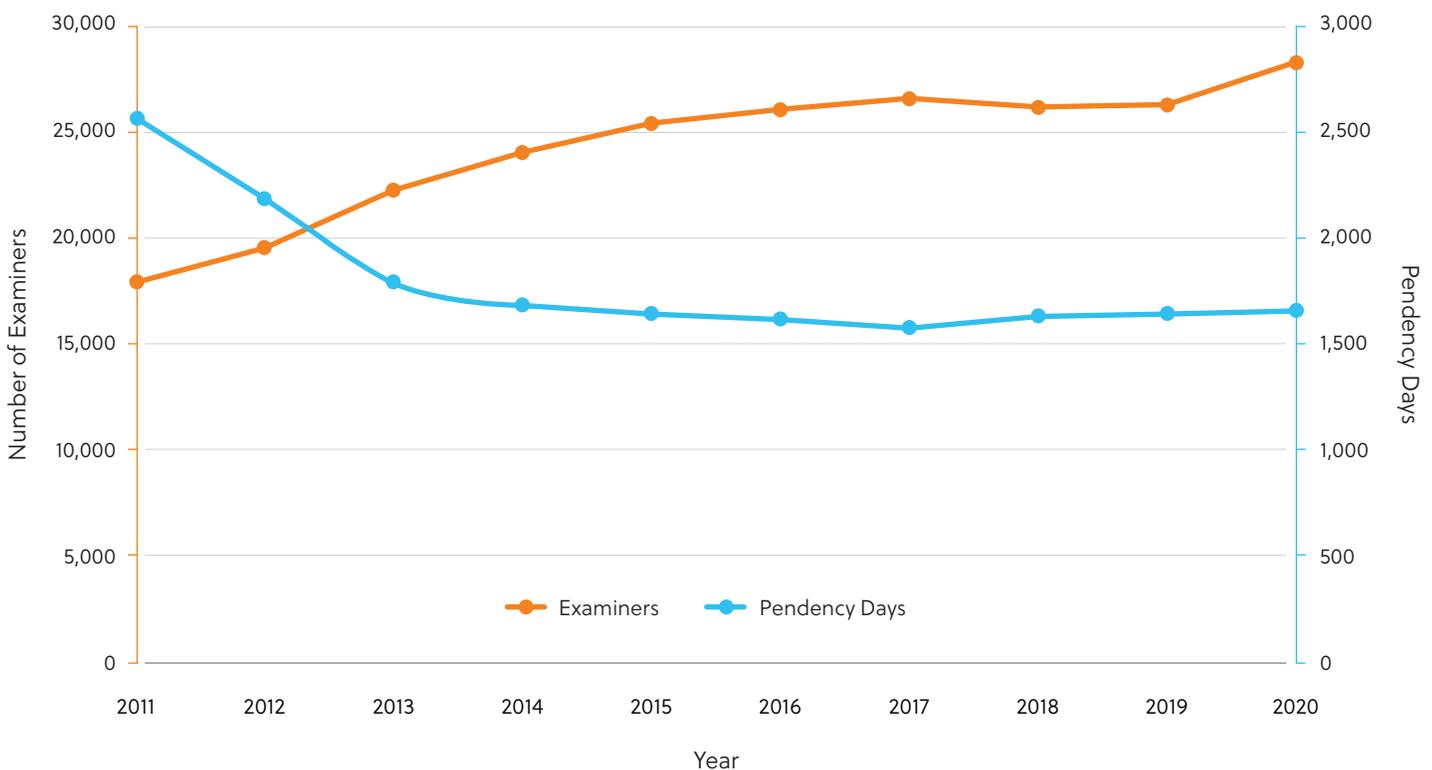
The world's patent offices are taking steps to reduce pendency by increasing the speed, quality, and efficiency of patent prosecutions. These include global IP harmonization, hiring more examiners, outsourcing prior art searches, and implementing new technologies.

Hiring more examiners has been an important answer to address application and pendency growth (**Figure 1**). However, hiring alone may be unsustainable in the long-term, owing to budgetary demands. Revenues that support hiring can fall out of line with examination demands because program costs do not necessarily decline with a drop in fee collections from the post-allowance activities that, in part, fund patent office operations.

Outsourcing prior art searches has long been used to augment patent office capacity, fill gaps in new subject matter expertise, and perform time-consuming, highly complex searches. As examiners spend a significant amount of time performing and analyzing searches, outsourcing can enable them to focus efforts on other important activities.

Perhaps the most promising solution for transforming the examination process is the adoption of new approaches that use artificial intelligence. Optimizing prior art search workflows in this way can improve examination efficiency and patent quality by saving time and ensuring the discovery of additional art that the examiner or searcher might have missed previously. AI is currently being used by over a dozen patent offices around the world to optimize prior art searches, patent classifications, machine translation, and other functions. AI augments human expertise by easily analyzing millions of data sets in seconds and returning highly relevant search results. This reduces reviews from hours or even days, to minutes. For offices hiring examiners, AI can also greatly reduce onboarding time, allowing the new examiners to achieve desired levels of competency and productivity faster.

Figure 1. Aggregate Pendency and Staffing Trends for Top 5 Patent Offices 2011-2020¹



Applying AI successfully in patent office workflows

Searching for prior art is a complicated, iterative, and time-consuming process. AI breaks it down into manageable components that algorithms can perform and analyze. A recent collaboration between CAS and the National Institute of Industrial Property (INPI) of Brazil delivered a custom solution that optimized prior art search workflows by integrating scientist-curated content, specialized technology, and unique scientific and patent expertise (see INPI Brazil case study summary). This novel approach yielded important insights into the requirements for the successful implementation of predictive solutions to optimize patent office workflows.

CAS scientists and technologists work with organizations around the world to improve workflows by harmonizing data and connecting disparate information systems. Workflow optimization ensures the availability of accurate information for critical decision-making and maintains ongoing data consistency and integrity. It begins with assessing the information landscape and then establishing effective guidelines for data management, allowing organizations to better leverage data and other assets to improve decision-making, accelerate innovation, and reduce risks.



1. Clean, structured data can improve predictive accuracy

Using high-quality structured data to train AI models and perform searches can reveal more patents with similarities and identify adjacent patents that could raise obviousness concerns. A CAS study² demonstrated that human-curated content could increase the prediction accuracy of AI algorithms by 30%, delivering the most accurate and relevant results for prior art searches.

In the INPI Brazil project, the quality of both the data used to train algorithms and the data used in performing searches was vital to achieving high levels of precision in the results delivered by AI-enabled searches.

The challenge for many organizations when it comes to data management for AI is that vast amounts of their data is unstructured, which makes it less useful for highly tuned algorithms. Publicly available scientific and patent data all carry inherent challenges, including transcription errors, mislabeled units, and overly complex patent language. The global nature of available data published in more than 60 languages also presents a challenge.

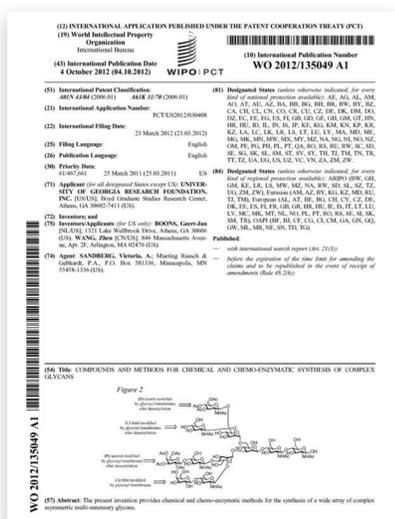
As a result, normalizing, preparing, and connecting data in a structured format is critical to success in training AI algorithms for specific models and for running searches in a live environment. Few organizations have the expertise to curate patent and non-patent publication data and extract it in ways that enable algorithms to deliver the most relevant results. When AI fails to live up to expectations, it is often because of poor data quality.

Having quality data is especially important in complex searches, such as those in fields including chemistry and biomolecules. These types of substances are often described inconsistently across publications, with their attributes embedded in tables or images, which can render the information unreadable by most search technologies.

The algorithms used to assist INPI Brazil with its chemistry application searches were trained and powered by the highly curated and structured data within the CAS Content Collection™. This collection is the most robust chemical and biomolecular substance dataset in the world. It includes over 250 million unique substances, related structures, and properties, and cross-links them to provide better hit rates for every type of patent and scientific search. The data has been curated by hundreds of scientific analysts who assign a unique CAS Registry Number® to each substance for more consistent indexing across patent and non-patent literature.

The vast amount of structured data CAS derives from millions of patent and non-patent documents is connected and made available to patent offices for prior art searches. CAS curates and connects data across the breadth of scientific disciplines to provide greater harmony between substance references for generating search results with greater precision and contextual insights. This comprehensive collection of data can be customized to support other specific search needs and applications.

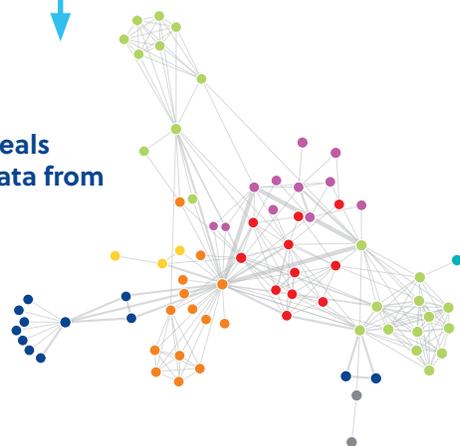
Figure 2: The power of data curation



Curation by CAS scientists yields:

- 7 concepts
- 138 substances
- 4614 chemical reactions
- 4 patent family members
- 3 cited documents

Patent knowledge graph reveals vast connections between data from different publications



WO 2012135049

Compounds and methods for chemical and chemo-enzymatic synthesis of complex glycans



2. Multiple algorithms can return similarities with the highest relevance

Greater precision and relevancy of search results can be achieved by utilizing multiple search algorithms trained using the large corpus of highly diverse and structured patent and non-patent data.

Once data is curated and structured to make it more accessible, multiple algorithms can use different models to analyze it. While some AI approaches may use one facet of similarity searching, such as semantics, for searches, more precise search results can be generated by integrating multiple algorithms trained to find semantic, syntactic, connectedness, and substance similarities. (Figure 3).

In the INPI Brazil project, more than ten algorithms were deployed, as well as a patented algorithm for ensemble learning that combined results to produce an optimal predictive model, generating relevant ranked results. In doing so, the algorithms worked together to leverage their strengths, while offsetting their limitations.

To evaluate semantics, such as title, abstract, and claims between publications, certain algorithm models were developed to focus on specific data fields, while others were trained for specific technology fields. These similarities were computed and rank-ordered by each model.

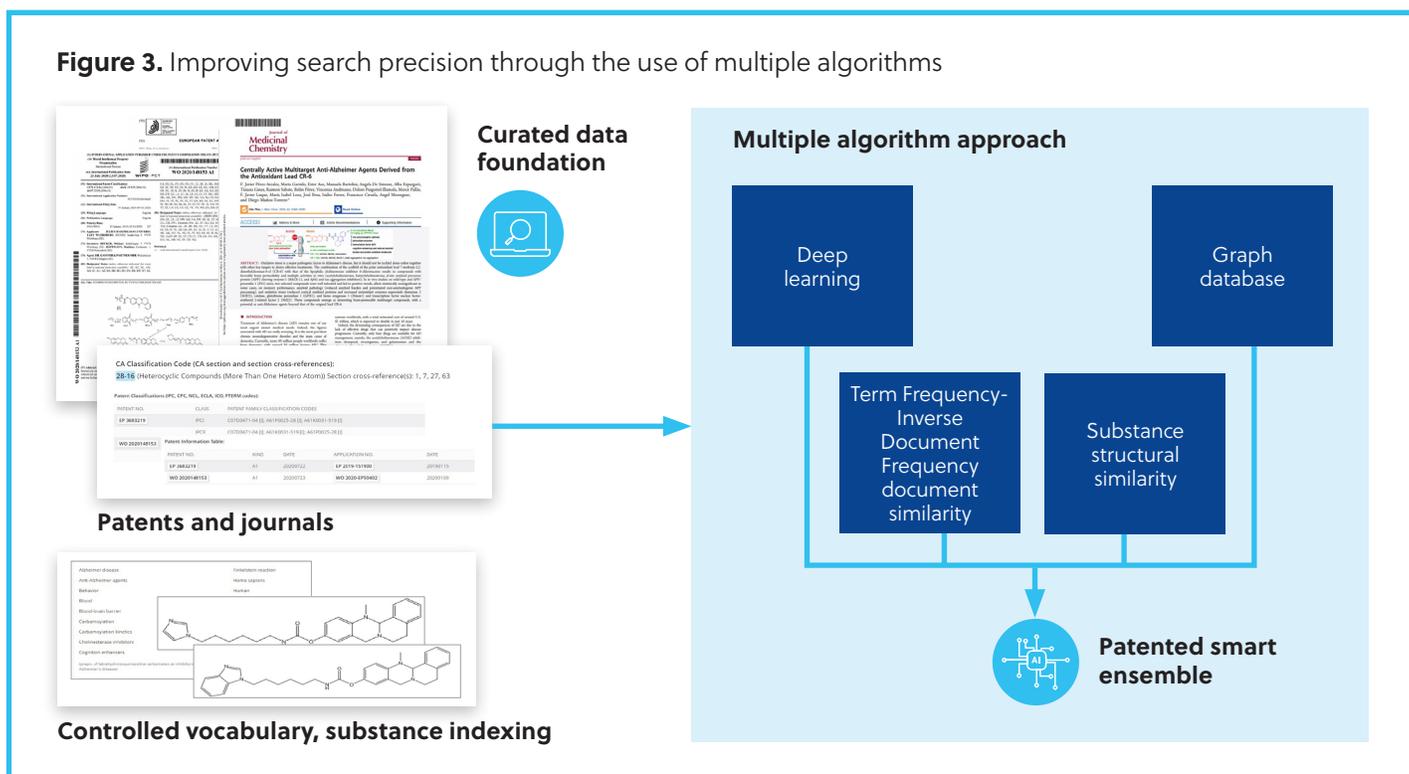
Similarly, a syntactic-driven algorithm compared the prevalence of special terms in the target document to their uniqueness across all other documents. If a special

term was prevalent in the target application and found as frequently in other patents, then patents where it was discovered were likely to have higher relevancy. Additionally, adjacent words in the text were also analyzed to reveal potential obviousness spaces around existing inventions.

Traditional knowledge graphs were also added to analyze the connectedness between the vast amounts of data. The INPI Brazil project deployed one for chemistry and one for non-chemistry to determine ontological similarity and connectedness between documents using keywords, scientific topics, roles, and nomenclature.

To compare the similarity of small-molecule chemical substances between patent documents, the project team deployed CAS indexing, which included CAS search screening and substance roles. Information about substance roles and substance rarity features was used to weight the similarity scores, thereby improving search precision.

Once the first-level algorithms returned their similarity results, the ensemble algorithm analyzed results based on search context and each algorithm's strengths and limitations. This final step arrived at a single, prioritized list of the publications that were most likely to be in conflict with the target patent.



3. Human expertise alongside technology can improve outcomes

AI projects require teams of individuals with a wide array of expertise, from data analytics and workflow integration, to technology development, and patent and scientific search patterns. Pairing technology with a team of experts who can improve workflow integration, validate algorithm models, and perform searches to augment the office's efforts, contributes to improved project outcomes.

Throughout the INPI Brazil project, CAS measured progress by hit rates on a random, holdout sample of over 37,000 historic patents that were evaluated by patent examiners in the patent offices of China, Japan, Europe, and United States, and the World Intellectual Property Organization. Hit rate, in this context, is defined as the number of examined patents where the analytics find a conflicting patent published prior to the application date.

Results were initially given to the office's patent examiners for further refinement as a list of patent application and patent citation numbers. Validation of the results was performed by INPI Brazil. However, examiners had to use multiple systems to curate the results, access source documents, track work accomplished, and produce the final report. CAS data

and workflow experts worked with them to refine their search process, reducing the total number of steps by half. CAS experts also implemented a cloud-based examiner workflow solution to consolidate search results, references, and examination activities within a single system.

The patent office relied on the expertise of the CAS IP ServicesSM search team for two key functions: validating algorithm results during development and performing highly complex searches to augment office capabilities.

Prior art search approaches can vary, and different search professionals are likely to design different approaches to a given search. Having a team of search experts available to analyze algorithm results can yield insights into how those algorithms can be further fine-tuned to improve relevancy.

Custom workflows can be fully automated or include partial outsourcing as used by INPI Brazil. The CAS search team fine-tuned the results provided by the AI engine and presented those in a formatted report, freeing up INPI Brazil examiners to focus on other important prosecution tasks to expedite first office actions.

How proven approaches can optimize progress for the future

Patent offices challenged by the growth in application volume and complexity will continue to require new approaches that enable them to meet growing stakeholder expectations within operational and resource limits. Flexible solutions that significantly improve workflow efficiencies and productivity will enable them to optimize their organizations for a dynamic and challenging future.

As no two offices have the same workflows, needs, or resources, it is vital to customize solutions for each patent office's unique and evolving challenges. The expertise CAS has gained through partnerships with patent offices, such as INPI Brazil, can be adapted to other offices, other types of patents, and other processes. These include classifications assignment and conversion, APIs for more efficient document

delivery, online tools for reviews and analysis, fully formatted reports, and fully automated workflows or workflows partially outsourced, and more. Customized workflow solutions that enhance productivity can, in this way, create a multiplier effect in improving global innovation.

As a non-profit division of the American Chemical Society, CAS builds long-term business partnerships and invests deeply in solutions that benefit the entire innovation ecosystem. Partnering with organizations around the world, CAS solutions leverage scientist-curated data, custom services, and unique technology platforms that can optimize digital workflows and other patent office initiatives for sustainability and continuous improvement.



Case Study: CAS Partnership with INPI Brazil Demonstrates AI Success in Increasing Examiner Efficiency

The challenge

In 2019, INPI Brazil faced a backlog of nearly 150,000 applications that averaged more than 10 years, significantly delaying innovation in Latin America's largest economy. That year, the office adopted measures to reduce the total backlog by 80% over the next two years, designed to have significant positive impact on innovation investment across the nation's industries. In response to these measures, Rana Gosain, senior partner at Daniel Law, of Rio de Janeiro, made the following observation:

"Countries that have modern and dynamic intellectual property systems are more successful in terms of innovation, and as a result, are economically successful. Considering that Brazil will continue to invest in infrastructure and innovation at the INPI and adopt positive and modern IP legislation, it has the potential to become one of the most coveted markets in the world."

As part of the initiative to reduce its backlog, INPI Brazil partnered with CAS³ to develop a custom solution to improve examiner productivity and reduce INPI Brazil's chemistry patent application backlog, which made up 15% of the total backlog. The project involved streamlining prior art searches using an AI-enabled workflow and CAS search services.

The solution

Prior art searches for chemistry patent applications are more complex and take longer to perform than for other categories. Searchers must review both patent and non-patent publications and include full text and structure queries.

The novel solution from CAS combined new AI-enabled similarity searches, re-engineered workflow, and CAS search support for validating algorithm models and augmenting office search capabilities. The specialized AI technology was expertly tuned with CAS's unique scientist-curated content and feedback from INPI Brazil examiners. Enhancements to the prior art search workflow were based on input from examiners and integration of the patent office's Internal Search Report formats, and all search results were centralized in a single custom, cloud-based interface.

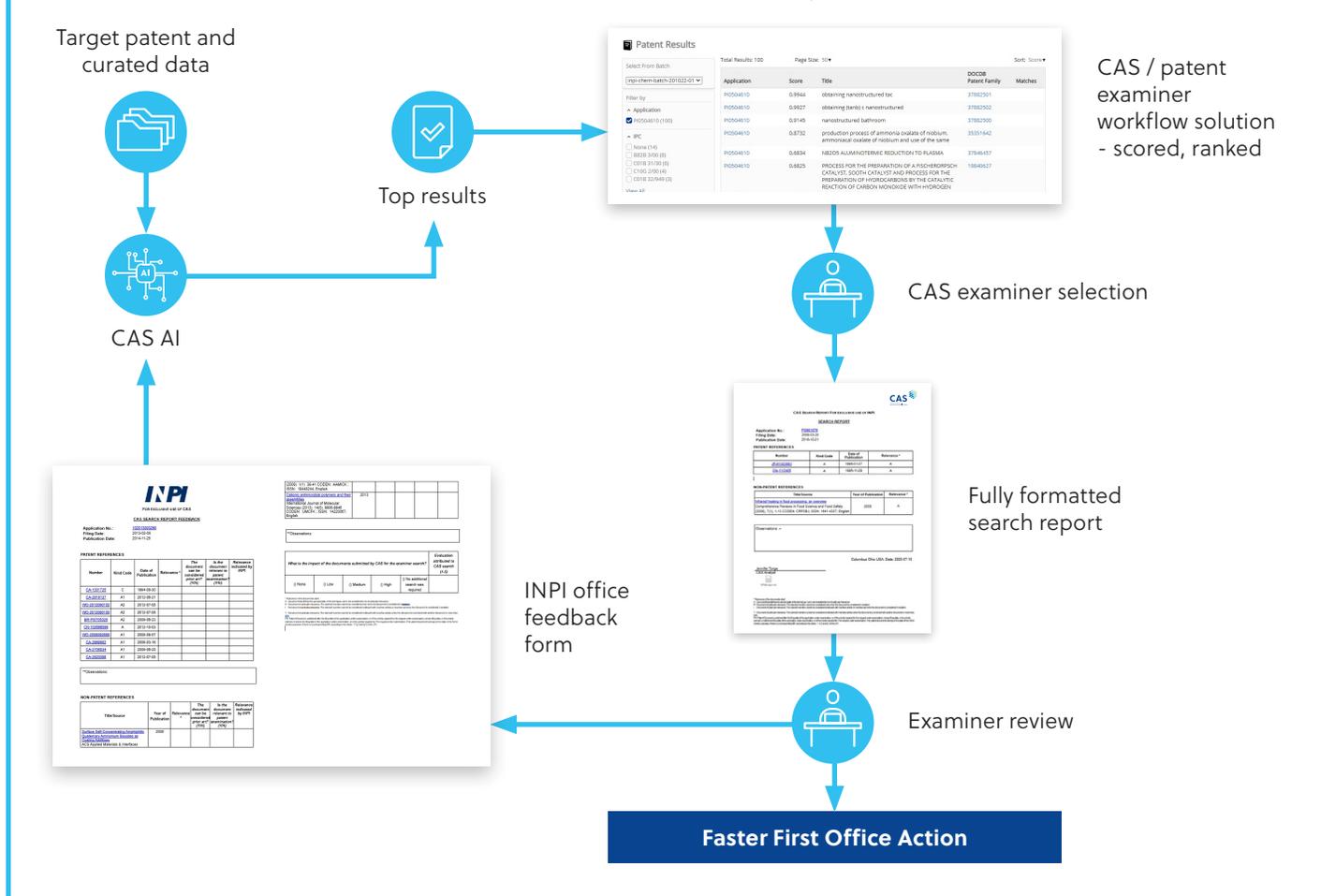
An in-house CAS IP Services team of search professionals assisted in reviewing and validating search results, providing inputs to refine and customize algorithms to the needs of the office. When an examiner picked up an application through the system's cloud-based examiner

workflow solution, they were able to access both the application and the set of supporting documents showing prior art search results, allowing them to make their decisions faster and with greater confidence. The workflow was customized, with different weights applied to patent attributes within the CAS AI algorithms, such as section codes, CAS Registry Numbers, and substances.

Several steps were eliminated within the prior art search workflow to make it more efficient, allowing examiners to spend more time performing highly complex searches and other activities (**See Figure 4**). As a result, they no longer conduct the following time-consuming search steps:

- Devising a search strategy
- Selecting which databases to search
- Compiling the necessary data to conduct searches (such as keywords in claims, structure drawings, and substance information)
- Performing the search
- Modifying the necessary data, databases, or strategy throughout the evaluation process

Figure 4. Customized workflow streamlines examinations to increase productivity



Outcomes

INPI Brazil evaluates the performance of the CAS solution based on its precision in returning relevant prior art references, impact on examiner search time, and impact on workflow. Based on feedback from the patent office, the CAS solution provided the following benefits:

- Up to 50% reduction in examination times
- 77% of all national applications processed required less examiner search time
- 29% of all national applications processed required little or no additional search
- Workloads were efficiently managed without adding staff

- Examiners were freed up to focus on other priorities
- Productivity improvements contributed to a reduction of 80% in the office’s backlog

To date, INPI Brazil has achieved significant success with its overall backlog reduction program and the transformation of its chemical application examination process. The office will continue making significant reductions in backlog as it meets aggressive future growth rates. While INPI Brazil chose to deploy the technology and workflow solution as a cloud-based application, the technology itself can also be implemented within existing patent office workflows or within CAS technology platforms, such as CAS STNext or CAS SciFinderⁿ.

References

1. WIPO patent application data for patent offices in China, the European Union, Japan, the Republic of Korea, and the United States.
2. Impact of Data Quality on Machine Learning Results, 2021, CAS.
3. INPI and CAS Sign Technical Cooperation Agreement Formalizing Collaboration to Tackle Brazil’s Patent Application Backlog, 2021, CAS.





Learn more about how CAS can develop AI-enabled solutions to streamline your critical workflows. www.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

cas.org

