

CAS CUSTOM SERVICES<sup>SM</sup>

# UNRAVELING THE POTENTIAL OF INTERNAL R&D DATA

Solution success story

## The challenge: The complexity and missed connections in scientific data

A large health-tech organization had decades of key research trapped in “dark data” – unstructured, inaccessible documents. Key research findings were housed in several disconnected, legacy systems with no systematic way to search the documents.

As a result, relevant individuals struggled to locate, access, and apply this valuable information to their work. Scientists had to know what they were looking for and manually sift through the files until they found it. As veteran scientists retired, proprietary knowledge was not being utilized by newly hired researchers who were unaware of the

available insights. Many experiments were unnecessarily repeated across R&D groups, which cost considerable time and resources. Without the ability to access and utilize the knowledge gained by other researchers across the organization, innovation was hindered, costs increased, and schedules were pushed out.

The lack of organization, limited search capabilities, and an insufficient data model meant that valuable information remained hidden in their systems. This rendered it useless for future research, discovery, and decision-making.

## The solution: Access to the full breadth of the organization’s knowledge

The organization sought to better utilize decades of stored research through improved searching and retrieval of data, helping to drive more efficient innovation. Utilizing the same process used to curate the world’s published science that powers CAS SciFinder®, the leading scientific information solution, CAS built a custom knowledge management system to curate and connect internal R&D data for the company. This enabled the organization to search the full text of their documents, connect similar concepts and substances, and search by concepts tailored to their R&D efforts, such as indicating if a substance is a natural product.

Moving beyond standard keyword search to enable scientific context, information extracted from documents is associated with key concepts and custom criteria based on the client’s innovation needs. Now, a researcher can immediately find where a concept is referenced in the document. Refined lexicons, ontologies, and taxonomies allow for matching on synonyms or even matching on classes of substances. This sophisticated feature is far beyond the capabilities of any out-of-the-box solution.

Knowledge graphs were constructed based on the available data, enabling the organization to link information on substances, scientific information, and concepts, and highlighting connections between that data across the organization. By utilizing knowledge graphs, researchers can find others in the organization working on similar areas of science, creating collaboration opportunities. Data in internal documents is further enhanced when connected with custom content curated from the CAS Content Collection™. Researchers can now identify and fill data gaps as a result.

With access to the full breadth of the organization's knowledge, teams now leverage their valuable R&D data to accelerate the innovation of new health-tech solutions. Before implementing the customized search and retrieval tool, one search for a specific data point within the unstructured files took a senior scientist approximately eight hours; now it takes minutes. This not only frees up resources for more specialized tasks, but also creates cost savings by preventing unnecessary experimentation and repetition.

Find out how CAS Custom Services can help you transform scientific data into actionable, evidence-based insights that maximize investment and fuel success at [cas.org](https://cas.org).

**"One search took a senior scientist  
~8 hours; now it takes minutes with  
scientifically aware search from CAS."**

Company estimate