

CAS CUSTOM SERVICES<sup>SM</sup>

# IMPROVED DATA INTEGRITY ENHANCES R&D EFFICIENCY

Solution success story



## The challenge: The complexity of an ever-expanding data set

A large diversified chemical company was investing in R&D efforts to be first-to-market with new innovations. Every day, more than 800 new substances were being added to their databases, including chemical names, synonyms, structures, additional identifiers, and associated regulatory data. Different CROs, suppliers, scientists, and analysts all contributed data in a variety of formats and quality levels. The information was housed in several disconnected systems, including the compound registry, inventory system, and knowledge management platform. These systems were vital to product research, development, and strategic decisions. The increasing level of inconsistent and inaccurate data made being first-to-market more challenging.

Reliable information was not available to inform product management decisions or submit accurate external regulatory documents. In efforts to harmonize the information, the organization attempted to manually confirm CAS Registry Numbers®, proper compound names, and other chemical information. Time limitations and human error limited the effectiveness of this approach to clean up the existing data and ensure the quality of new data inputs. As individuals across the organization invested considerable time and resources to identify, correct, and manage the impacts of inconsistent and inaccurate data, the organization needed to replace manual processes to enable teams to work more efficiently and effectively.

## The solution: A unified representation for substances

To ensure the company could better leverage their data as an asset, CAS scientists and technologists worked with them to assess their information landscape and forge a path forward. CAS identified the primary area of inconsistency was substance representations, coupled with an overall lack of data cleanliness. CAS guided the teams to select a unified representation for substances and establish guidelines for management of the data.

CAS experts then went to work cleaning and harmonizing substances across the existing systems. Using CAS REGISTRY® as the authoritative source for chemical identity, CAS chemists leveraged substance comparison technology in combination with their own expertise to ensure the company's data was of the highest quality and aligned with their expectations. The internal systems were updated to contain the same information for all substances, including the company's internal names and identifiers, CAS Registry Numbers®, and all known relevant names as indexed by CAS. Since substances are added to each of the systems frequently, CAS integrated an API connection that enables seamless look-up of the CAS Registry Number and associated information for new substances being added to the database.

In collaboration with key stakeholders, CAS could quickly identify, repair, and harmonize the incorrect data throughout disparate databases across the organization with minimal staff intervention. Researchers and product owners can now complete their work with timeliness and confidence, no longer doubting the quality of internal data or spending time loading and validating it. The organization can now leverage its data as an asset to accelerate innovation, drive strategic decisions, and reduce risk.

## After partnering with CAS for its new data solution, the company estimated the following:

- Total cost savings of over \$800k annually
- 3,300 hours in annual saved time of end users
- \$300 thousand annual cost reduction in ingesting and correcting data

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).