

# CAS creates first English version of world's oldest chemical journal with Iconic language technology

## PROJECT FACTS

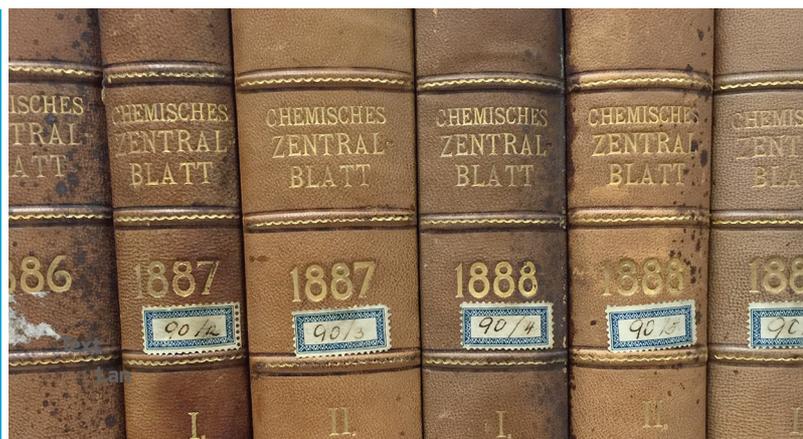
**Language:** German into English

**Volume:** 800,000 documents, 3 million abstracts

**Content:** Historical chemical information

**Formats:** Physical books, PDF, XML, Images

**Technologies:** Custom MT, OCR, Information extraction, Machine learning



## OVERVIEW

*Chemisches Zentralblatt* is the world's oldest journal of German chemical abstracts from 1830 to 1969. Only available in German, the compendium was challenging to search. In 2016, CAS, a division of the American Chemical Society, introduced a new solution, ChemZent™, developed in partnership with Iconic, unlocking 3 million abstracts of valuable, historical scientific information in English.

## THE CHALLENGE

It was difficult to explore and identify foundational chemical information in the original German hardback version of *Chemisches Zentralblatt*. Organisations were also missing access to the complete digital repository of the earliest chemistry history. To solve this, CAS turned to Iconic, tasking us with translation of the mammoth compendium from German into English for indexing and search.

## HOW ICONIC HELPED

Iconic developed innovative machine learning technology to extend its existing Machine Translation (MT) and Natural Language Processing (NLP) solutions for the large-scale digitisation, translation and indexing of 140 years' worth of German journals and patents. Customising proprietary MT technology for historical chemical information with statistical and neural approaches, Iconic also developed NLP technologies for digitisation of 800,000 PDFs, and extraction and fielding of 3 million abstracts by author and title. Iconic's expertise together with CAS industry-leading scientific information analysis made the launch of ChemZent™ possible within one year of idea inception. This foundational chemistry is now searchable in English with summaries in English as well.

## RESULTS



### 1 billion words translated and indexed

English-language searchable version of world's oldest chemistry abstracts journal, *Chemisches Zentralblatt*, available for first time in history.



### CAS releases ChemZent™

CAS makes history by providing unparalleled discoverability of chemistry with the launch of ChemZent™ in 2016, with first-year sales projections exceeded.



### Fueling future scientific innovation

The global scientific community can now search and find valuable foundational chemical information expected to fuel future scientific discoveries.



## ABOUT THE CLIENT

CAS is dedicated to the American Chemical Society vision of improving people's lives through the transforming power of chemistry. The CAS team of highly trained scientists find, collect, and organize all publicly disclosed substance information, creating the world's most valuable collection of content that is vital to continued scientific innovation worldwide. [www.cas.org](http://www.cas.org)

## ABOUT LANGUAGE WEAVER

RWS Language Weaver® is a leading machine translation software and solutions provider who specialise in custom solutions tailored with subject matter expertise for specific industry sectors including legal, life sciences and financial services.



"Access to the historic collection of *Chemisches Zentralblatt* via ChemZent in SciFinder has opened this scientific resource to chemists the world over now searchable in English with summaries in English as well. The ability to do an exhaustive search that includes this important foundational chemistry makes information relevant to research more easily accessible."

### Dr. Matthew J. Toussant

Senior Vice President of Product and Content Development

