Data Extraction and NLP: Semantic Searching on Research documents

95%
Accuracy
In extracting data from running text

Ease of use
Unified Graphical Interface

Case-Detection
Was significantly Improved

Easy to set-up
Integrated Platform

The Client

World's fifth-largest independent biotech company.

Industry

Pharmaceutical

Overview

Drug manufacturing and research companies work with many research documents. The cost of repeating a similar experiment or getting to discuss with other researchers improves the chances of success of experiments. There is a need for mining specific information blocks and accessing these reams of PDF documents.

Business Challenge

This leading pharma company wanted Quadratyx to extract structured information from documents amenable for rapid searching and advanced analytics. A major challenge was that there was no uniformity in the structure, content or style across the clinical trial protocol documents prepared in different parts of the world.

Our Approach:

Quadratyx designed a solution encompassing:
→ Robust, sophisticated text pattern recognition & mining.
→ Extracted values for all key parameters provided.
→ Generated a single consolidated file from thousands of free-text documents.

Outcomes

✓ Digitized key information extracted from free-text documents.
✓ Ease in searching required content by end users.
✓ Ease in building advance analytics solutions.