

CASE STUDY

Professional Services

A Top Pharmaceutical Company

Customized Data Sets for Improved Discovery



SUMMARY

A review of how Elsevier R&D Professional Services helped a pharmaceutical company find reported mutations in cancer cells.

Knowing how a mutation will alter the behavior of a tumor is invaluable.

Identifying Driver Mutations

One hypothesis in cancer research is that somatic mutations caused by genetic instability drive the development of cancer. The mutations can be passenger mutations, which do not affect the growth, survival and penetration of the cancer; or driver mutations, which alter the behavior of tumors.

The majority of cancer research focuses on finding these driver mutations and identifying their roles. Knowing how a mutation will alter the behavior of a tumor is invaluable. However, the sheer number of mutations reported in the literature makes finding information about them a monumental task.

Customized information retrieval

One of our customers is very involved in cancer research, so these driver mutations are of great interest to them. They wanted to find all of the reported mutations in cancer cells. The goal was to have a complete and searchable list of these mutations.

Since the customer could not find such a list or database, they approached the Elsevier R&D Solutions Professional Services team to see if one could be created. This group of search specialists knew both the background of the problem and the strategies that would generate the list. They also knew that it could be done quickly.

What was Elsevier's strategy?

The Professional Services team identified two important aspects to the task: finding the data in the huge volume of available literature on somatic mutations; and assembling a searchable database. For the first part, the team took Elsevier's Natural Language Processing (NLP) Engine, which is ideal for such specialized and focused search strategies. For the latter, the team decided to attach the database as a custom cartridge for Pathway Studio, Elsevier's biological research solution.

The Professional Services team created a custom search that allowed the NLP Engine to search full-text literature and find statements indicating that the piece contained information about protein mutations. The team focused their attention on the different ways that scientists specify mutations in the text and linguistic structure of the sentences describing the link between mutation and disease.

What was the outcome?

It took just one month of testing and optimizing the search strategies to complete the task. More than 100,000 mutations were found in PubMed® abstracts (>20 million) and in all of the Elsevier and open-source full-text articles (>3 million). The extracted mutations were assembled into a custom-made data set. Using Pathway Studio, the customer could now search the list and find mutation they needed — and the search strategies will keep the their data set up-to-date.



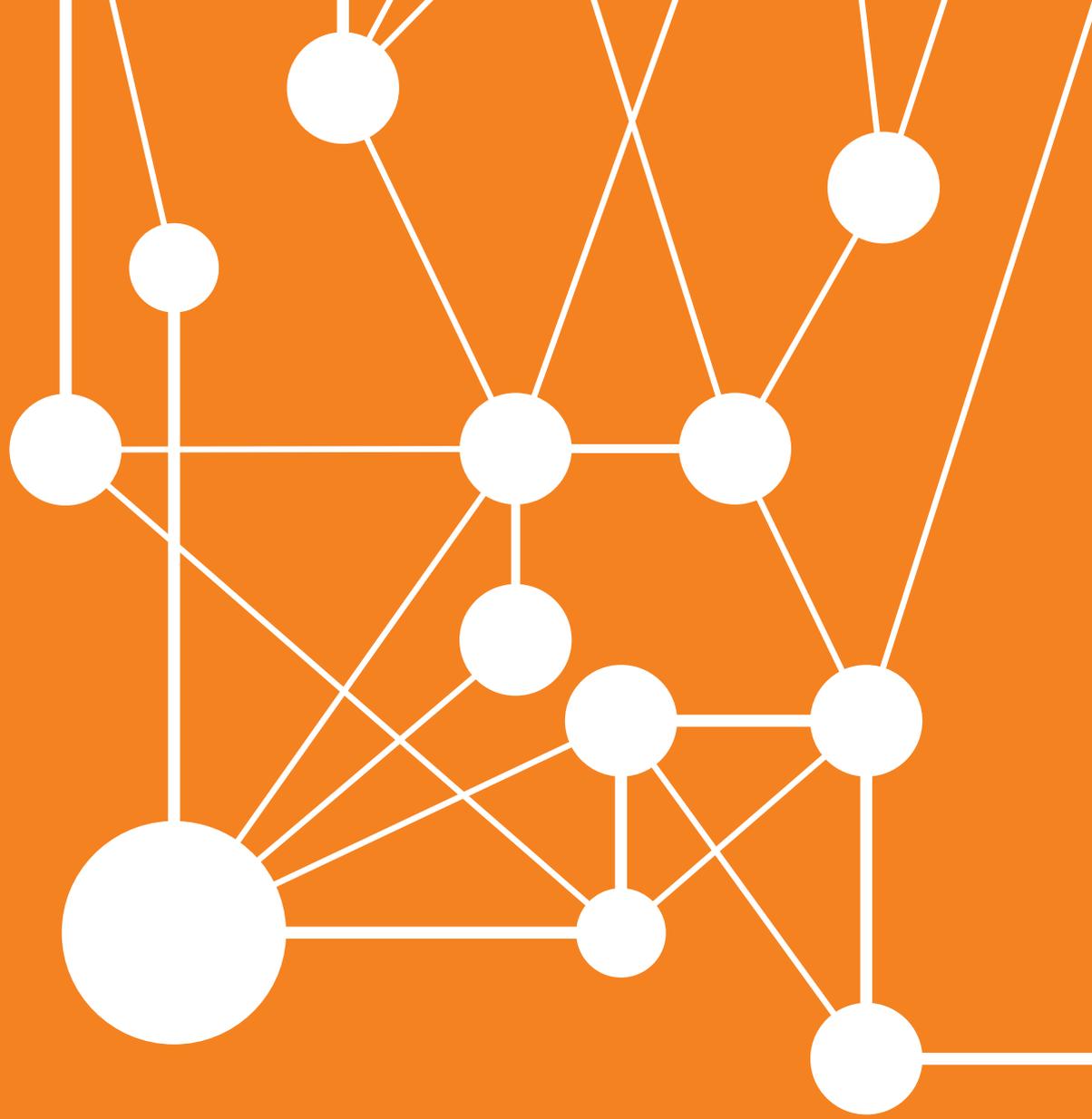
How long does it take to create a customized set of facts?

The members of the Professional Services team are experts in research informatics. They provide tools and expertise to meet specific needs, including tailoring search strategies to customer needs — and they can return results fast.

In creating a customized database, the task of the Professional Services team is to compile a comprehensive set of queries in the form of linguistic patterns to find all relevant facts that are expressed in different syntactic and semantic forms in scientific texts. The unmatched speed of Elsevier’s NLP Engine allows the rapid development of queries and linguistic patterns. The linguistic research can take from couple days to a month, depending on the task.

Even for the complex task of finding all the published mutations related to cancer, the team can develop, test, and optimize the strategy dozens of times within the space of a month.





Visit elsevier.com/products/solutions/professional-services
or contact your nearest Elsevier office.

ASIA AND AUSTRALIA

Tel: + 65 6349 0222

Email: sginfo@elsevier.com

JAPAN

Tel: + 81 3 5561 5034

Email: jpinfo@elsevier.com

KOREA AND TAIWAN

Tel: +82 2 6714 3000

Email: krinfo.corp@elsevier.com

EUROPE, MIDDLE EAST AND AFRICA

Tel: +31 20 485 3767

Email: nlinfo@elsevier.com

NORTH AMERICA, CENTRAL AMERICA AND CANADA

Tel: +1 888 615 4500

Email: usinfo@elsevier.com

SOUTH AMERICA

Tel: +55 21 3970 9300

Email: brinfo@elsevier.com