

**Use case**

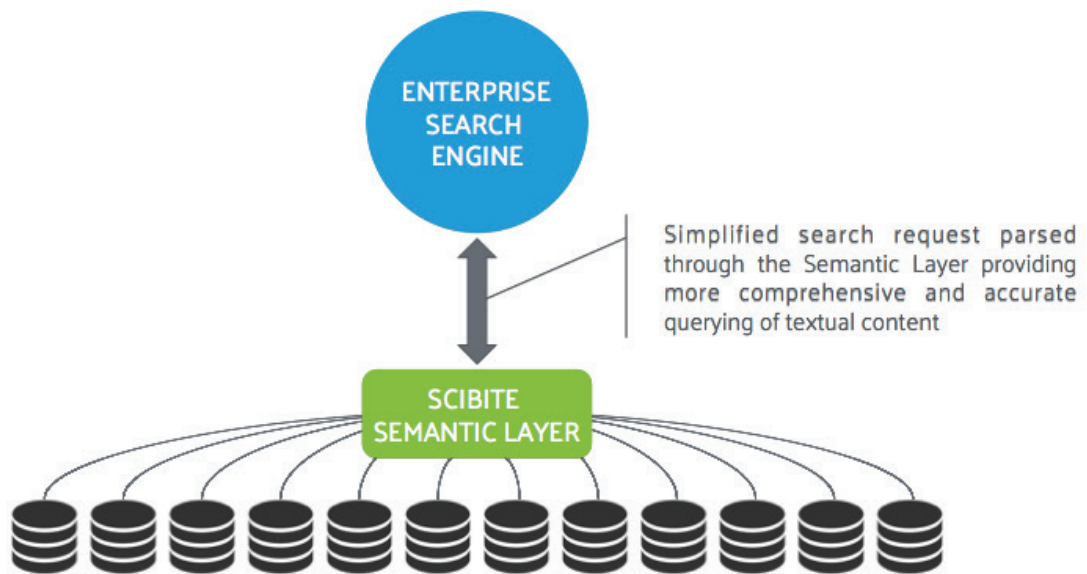
# **Transforming enterprise search**



## Transforming enterprise search

Sinequa, a leader in real-time big data search and analytics for Fortune Global 2000 companies have partnered with SciBite, a Cambridge, UK based Informatics Company, to provide access to drug discovery intelligence data for pharma and biotech companies including, amongst others, Astra Zeneca.

The partnership brings together Sinequa's flexible enterprise search program, which scales hundreds of millions of documents, and SciBite's semantic layer providing deep ontologies and lightning-fast text indexing engine to turn unstructured, diverse text into rich, structured resources. The combination now opens unparalleled access to drug discovery intelligence and vast amounts of knowledge, previously hidden in scattered document repositories.



**Figure 1:** SciBite identifies scientific entities during the data ingestion phase



SciBite CSO Lee Harland said, 'Our partnership with Sinequa allows us to apply semantic enrichment over a company's entire document collection as well as public data such as PubMed, patents, grant applications, and so on. This allows healthcare professionals to become more informed, without overloading them with information.'

## A semantic layer

A coupling of two modules forms the Core of SciBite's semantic enrichment capabilities.

**CORE PRODUCTS**

**TERMite**

Our entity recognition engine analyses and identifies relevant data, scanning 1 million words per second, relating to:

**VOCabs**

We have a highly-curated collection of 20 million + scientific entities (terms) and their synonyms, covering over 80 topics

Figure 2: The SciBite/Sinequa integration utilises TERMite and selected VOCabs

Incoming documents are analysed by TERMite, the entity recognition and extraction engine. The VOCabs provide an extensive reference library for entity recognition and are vastly enriched over any publically available alternative. The semantic layer is composed of a simple API specifically designed to integrate and enrich complimentary technologies such as the Sinequa platform. Input of any format and location can quickly be transformed into machine-readable data and fed back into the Sinequa platform.

## What does semantic enrichment bring?

SciBite semantic enrichment provides critical scientific search, retrieve and explore functionality to end-users. Examples include.

- **Improved search experience.** Identify occurrences of entities via different synonyms. So a search for 'Lipitor' finds documents mentioning any related synonym e.g. 'Atorvastatin'.
- **Concept-type searches.** Find any documents that mention an 'X' (where X means drugs, indications, parameters etc...).
- **Summary perspectives.** Ask 'What are all the targets discussed in documents?'
- **Ontology Queries.** Find any documents that mention a 'type' of entity such as a gene.

## Transforming enterprise search

The results of this innovative partnership are helping to transform the search experience for end-users by opening up content once locked up in unstructured text. The Sinequa platform collates information scattered across a customer's infrastructure regardless of format or location into a powerful, easy to navigate interface. The SciBite Semantic Layer provides the scientific context powering advanced search and analyse of the results beyond simple keyword matching.

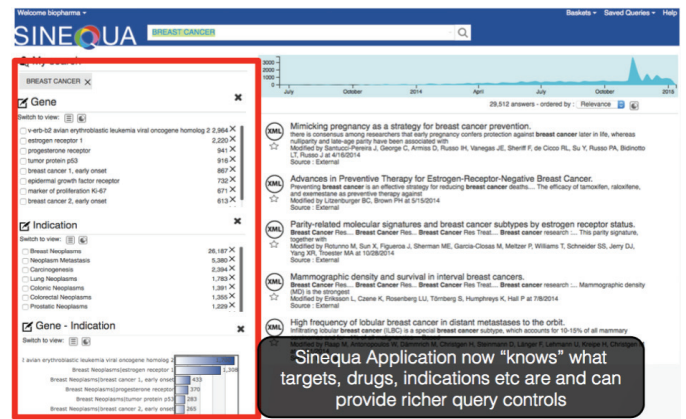


Figure 3: Sinequa application now 'knows' what targets, drugs, indications etc. are, and can provide richer query controls

Users can confidently explore relationships between drugs, indications, targets, safety signals, biomarkers and much more. Sinequa's flexible API caters for the development of further bespoke solutions such as key-opinion-leader mining, drug repurposing and target identification.

Contact SciBite today to learn more about the value of adding a Semantic Layer into your technologies and unlocking the value of unstructured scientific information.



SciBite's data-first, semantic analytics software is for those who want to innovate and get more from their data. At SciBite we believe data fuels discovery and we are leading the way with our pioneering infrastructure that combines the latest in machine learning with an ontology-led approach to unlock the value of scientific content. Supporting the world's leading scientific organisations with use-cases from discovery through to development, SciBite's suite of fast, flexible, deployable API technologies empower our customers, making it a critical component in scientific, data-led strategies. Contact us to find out how we can help you get more from your data.

To learn how SciBite can unlock the value of your data, speak to one of our experts today or email us at [contact@scibite.com](mailto:contact@scibite.com)

**Head Office:**

SciBite Limited  
BioData Innovation Centre  
Wellcome Genome Campus  
Hinxton, Cambridge CB10 1DR  
United Kingdom

 [www.scibite.com](http://www.scibite.com)  
 [contact@scibite.com](mailto:contact@scibite.com)  
 LinkedIn: SciBite  
 Twitter: @SciBite  
 +44 (0)1223 786 129

