



## CUSTOMER SUCCESS STORY: LEXISNEXIS UNIVENTIO

### ENABLING UNPRECEDENTED INSIGHT IN INTELLECTUAL PROPERTY



#### TABLE OF CONTENTS

<b>Introduction</b>	<b>1</b>
<b>Tackling the Challenges</b>	<b>2</b>
<b>Why Omniscien Technologies</b>	<b>3</b>
<b>Continued Innovation</b>	<b>3</b>
<b>Benefits to LexisNexis Univentio</b>	<b>3</b>

LexisNexis Univentio is a leading global provider of patent information. LexisNexis Univentio combines patent data from more than 100 authorities (patent offices) around the world into a single, unified, tagged and structured format and supplies this data to LexisNexis and Elsevier as well as to other patent research and analytics platforms and enterprises worldwide. LexisNexis Univentio provides one of the world's largest collections of patent data including machine translated contents and extracted images.

For corporate divisions like the research & development or the legal department it is crucial to protect their company's intellectual property as well as to ensure that other parties' rights are not infringed. Therefore, conducting research on patents is a necessary task of these departments. Besides, business intelligence often combine intellectual property data with other key business information in order to support strategic decisions. The importance of this research has even increased in the light of recent high-profile court cases in the high-tech domain. Intellectual property data and in particular patents, however, are very complex documents. Therefore, it requires extensive experience and knowledge in this domain, as well as rich tools and access to updated information in order to effectively do a patent research. Since patents are global and published by the thousands per week, deriving quality information from text by machine (text data mining) with powerful search tools ranging from Boolean to semantic turned out to be the only feasible option. For cost and time reasons but also for consistency, for example on terminology, applying Machine Translation in this case proved to be the best option.

As part of the process LexisNexis Univentio had to tackle a number of complex challenges which are discussed in the following section.

## HIGHLIGHTS

- **Omniscien Technologies was the only supplier capable of providing the range of language pairs and the capacity needed for effective translation and at an economically feasible price**
- **The level of optimization and workflow control offered by Omniscien Technologies was unprecedented and allowed LexisNexis Univentio to improve the output drastically.**
- **The range of language pairs that Omniscien Technologies provided and in particular the Asian ones, enabled LexisNexis Univentio to take lead over the competition**

## TACKLING THE CHALLENGES

### 1. Volume

One of the problems that LexisNexis Univentio was facing, specifically with large, non-English patent offices, such as CIPO (China), KIPO (Korea), JAPIO (Japan) but also to a certain extent with WIPO (UN) and EPO (EU), was the sheer volume of documents in the so called front files (updates) and back files (historical data). In order to process (and reprocess) millions of patent documents within a reasonable timeframe, a powerful and scalable system was required.

### 2. Complexity of Content

Patent content is complex and spans a wide range of areas from the latest developments in Bio- and Nanotechnology to Oil & Gas as well as Electronics and many more. Furthermore, patents typically contain a lot of references, formulas, images and tags which increase the complexity of this data by an order of magnitude. Therefore the translation of these kinds of data requires a translation system that must be able to perform extensive rules instructions in pre- and post-processing in order to handle all this 'non-text' information as well as it must be able to ensure that those 'non-text' data will be re-inserted into the document at the "correct" place after the translation.

### 3. Normalization

Search algorithms like Boolean and others rely very much on normalization of terminology in order to get positive and relevant search results. Therefore, besides the required translation abilities, normalizing functions drastically improves the output and also the search results.

### 4. Variety of Languages and Domains

As patent data are global, the range of language support needed is extensive. Besides, it covers a broad variety of different domains, from biotechnology to mechanics and engineering, using different and sometimes even conflicting terminologies among those domains.

*"We have always pioneered advanced solutions to enhance our offerings, and our deployment of the new Language Studio™ NMT capabilities extends our existing partnership with Omnicien Technologies. It provides us with the ability once again to establish our leadership position in the industry and provide a unique, high-quality offering to our clients."*

**Eric van Stegeren,**

Senior Director Technical Solutions  
and Managing Director of LexisNexis  
IP Solutions

*"Language Studio™ from Omnicien Technologies has served us well over the years, and the addition of NMT allows us the flexibility to select the best technology for the job dynamically, be it SMT, NMT or a combination of both. This represents a unique opportunity to achieve far better quality translations and solutions as part of our products."*

**Laura Rossi,**

Manager Language Technology Solutions  
at LexisNexis IP Solutions

## WHY OMNISCIENT TECHNOLOGIES

---

After a number of attempts with other suppliers and after doing an extensive market research and testing, LexisNexis Univentio chose Omnicien Technologies in 2008 as the supplier for their Machine Translation System and has continued to use Omnicien Technologies' system until today.

LexisNexis Univentio opted for Omnicien Technologies' Language Studio™ platform as an in-house system running on a very high number of chores for the reasons that Language Studio™ could provide LexisNexis Univentio with the required range of capabilities and control and last but not least, because of the output quality.

## CONTINUED INNOVATION

---

Both LexisNexis Univentio and Omnicien Technologies share a passion for continued innovation and market leadership. In an ongoing quest to provide best quality translations and to increase accessibility for a global customer base, LexisNexis Univentio and Omnicien Technologies have deepened their cooperation in 2017, with LexisNexis Univentio becoming the first global patent information provider applying machine learning technology based on Neural Machine Translation (NMT). Through the application of this new technology supplied by Language Studio™, LexisNexis Univentio has once again taken the lead in the industry in terms of quality. This collaboration also enabled LexisNexis Univentio to provide very high quality patent contents in Asian languages in addition to the English language, thus capitalizing on an increasing demand from the larger Asian markets.

## BENEFITS TO LEXISNEXIS UNIVENTIO

---

LexisNexis Univentio could benefit from this collaboration in a number of ways, most prominently by gaining an extensive lead over the competition. Together with Omnicien Technologies, LexisNexis Univentio has created a base platform that enables fast adding of language pairs and additional control for any requirements. The ability to process data faster, to shorten the time to market and to provide better quality played an important role for this competitive advantage.

## ABOUT OMNISCIENT TECHNOLOGIES

---

Omniscien Technologies is a leading global supplier of high-performance and secure high-quality Language Processing, Machine Translation (MT) and Machine Learning technologies and services for content intensive applications. Our wide range of solutions serves clientele from various industries including the Localization Industry, Online Research Services, Publishing, eCommerce, Media, Online Travel, Technology, Enterprise and Government.

Omniscien Technologies has gained a reputation for cutting edge solutions with its Language Studio™ platform. Depending upon the customer's unique requirements, Language Studio™ can be deployed in a variety of ways to integrate with their in-house data processing and translation management systems, and it offers unparalleled levels of customization and control as well as feature rich pre- and post-processing, enabling customers with even the most complex data to achieve both high quality and high volume output to satisfy every use case. Omniscien Technologies has by far the most comprehensive and feature rich system in the market today.

Covering 550 language pairs and with a number of industry specific solutions, Omniscien Technologies remains the partner of choice for customers with complex, high-volume bespoke data processing and machine translation needs.