IQVIA NLP Takes Mundipharma Research Toward IDMP Compliance And Beyond

IQVIA's NLP text mining solution provides data extraction for Mundipharma Research Limited's IDMP compliance project, with potential applications identified across the wider enterprise

QUICK FACTS

Situation: IDMP compliance will require Market Authorization Holders to submit and maintain a broad range of data elements about medicinal products with the EMA, 70% of which currently only exist in unstructured text, hidden in multiple document formats, styles and languages.

Solution: The IQVIA (formerly Linguamatics) natural language processing (NLP) text mining solution rapidly reveals structured facts and relationships from unstructured text by understanding meaning. It has the potential to deliver relevant, high quality IDMP data elements from documents in real time.

Success: An NLP IDMP compliance experiment resulted in good or excellent data capture accuracy from English-language Summary of Product Characteristics documents, and evidence of clear pathways to drive accuracy even higher. Mundipharma Research has now moved from experiment into production, and is embedding text mining into its broader architecture to support data-driven decisions.

Summary

Like all European pharmaceutical and life science organizations, Mundipharma Research faces a major challenge to deliver and maintain compliance with IDentification of Medical Products (IDMP), following the EMA timetable. Unlike some, it already knows it has a quick, accurate and affordable approach to delivering a number of the data elements required by IDMP, having successfully piloted the required regulatory workflows using IQVIA's' NLP-based text mining tool.

Situation

IDMP is a set of five ISO standards designed to standardize the description of marketed medicinal products. Better patient safety is the main goal, by improving the reporting of adverse events and safety signals globally. IDMP will be mandatory worldwide for all Marketing Authorization Holders in ISO countries.

The EMA will be the first health agency to mandate compliance with ISO IDMP (initially from Q1 2022, mandatory from Q1 2023), with the FDA not far behind. Pharma organizations face a rapidly shortening time frame to prepare for full transition of regulatory data and systems to comply with this complex mandate.

The challenge for such companies is to find a quick, accurate, and affordable way to search, extract, standardize, and structure the 300–2,000 data elements required per product. Not easy when you consider that 70% of this data lies in unconnected silos of unstructured text.

Sources of such unstructured text include Summary of Product Characteristics (SmPC) documents; manufacturing licenses; Chemistry, Manufacturing and Control (CMC) documents; compliance documents; and regulatory dossiers. Complicating matters further, such documentation can be in multiple languages, from internal and external sources, and have different author styles and formats.

IDMP compliance will demand time, resources and investment. Aware of these potential costs and of NLP-based text mining's potential for the IDMP scenario, the Mundipharma Research team started to look for a market leader.

They soon discovered IQVIA and its NLP text mining tool. The IQVIA platform provides NLP-based text mining, which rapidly reveals structured facts and relationships from unstructured text, thus delivering relevant, high quality results in real time.

Cliff Gibson, Senior Enterprise Architect for Mundipharma IT Services Limited, explains: "Manual extraction of such a quantity of data is slow, expensive, unwieldy and open to human error. A data-driven approach to IDMP compliance is needed, by which I mean developing the business rules for extracting text by looking at types of document and refining the rules and queries depending on the data pulled out."

"[IQVIA NLP] enables such an approach. It can use any controlled vocabularies (both internal dictionaries and external controlled vocabularies such as MedDRA) to find the appropriate term or phrase in a document, extract this and map to the right term in the controlled vocabulary."

Solution

WHY IQVIA NLP?

The IQVIA NLP platform is a world-leading agile, scalable, real time NLP-based text mining solution. At the point of introduction to Mundipharma Research, it was already used by 17 of the world's top 20 pharmaceutical and biotech companies across the pipeline of discovery and development, and into the delivery of therapeutics. Various government bodies (including the FDA) and healthcare providers were also relying on it.

IQVIA's knowledge and experience of adapting the platform's capabilities to deliver regulatory requirements placed it in a unique position to help Mundipharma Research find, extract, standardize and structure IDMP data elements from unstructured text, particularly from SmPC documents.

Working with IQVIA made sense, so Mundipharma Research commissioned an experiment to determine whether NLP could automate the identification, extraction and coding of data elements from regulatory documents in preparation for IDMP implementation.

Success

The result was the accurate and automated extraction of 16 out of 18 target data elements from unstructured text across 40 product presentations and 15 different SmPC documents, with just 12 days of effort over 8 weeks.

"[IQVIA NLP] queries were developed to extract the individual data elements using standard and customized ontologies, as well as linguistic features of SmPCs. We evaluated accuracy against a 'gold standard' data set that had been manually extracted by an independent expert," explains Cliff.

"For 16 out of the 18 data elements [NLP] delivered good or excellent accuracy. For the remaining 2 data elements, accuracy was lower due to issues with the training set. It was clear that by further refining rules according to IDMP logic or by adding non-SmPC documents, both of which are easy to do with [NLP], we could drive this accuracy even higher," he adds.

Jon Sanford, Head of Regulatory Information Management and Operations at Mundipharma Research, concludes: "We were really impressed when we saw the accuracy with which [IQVIA NLP] had been able to extract data elements from the documents. We will still need to perform some level of QC to ensure data quality, but [the platform] will help us avoid the need for manual data extraction, which we know to be slow, expensive and error-prone."

What next?

The ease with which IQVIA NLP queries can be changed make it an ideal tool to help deliver IDMP compliance. Business rules can be rapidly translated into search queries, and this flexibility is key when the IDMP framework is still evolving. In addition, the platform enables text mining across multiple languages, with extensive capabilities for multilingual concept-matching, stemming, part-of-speech tagging and more.

Mundipharma Research is now planning to productionize the workflow for all medicinal products with English-language SmPCs, in addition to exploring how to extend its use to other languages. The broad functionality that IQVIA NLP enables has also led to Mundipharma Research planning further experiments with the platform across the wider enterprise, including:

- improving MedDRA coding consistency across clinical and regulatory data
- improving efficiency and effectiveness of Quality Control and Assurance across regulatory and other documents
- capturing regulatory Response to Questions and global regulatory guideline changes
- enhancing pharmacovigilance risk-monitoring capabilities
- improving Clinical Development Program (CDP) design by providing more realistic CDP timelines and cost estimates to the Board, to support investment decisions.

Mundipharma Research, it seems, is actioning Gartner's advice to use the time gifted by the extension of the IDMP deadline to enable more strategic decisions. IQVIA is proud to be a partner.

CONTACT US

+44 (0)1223 651 910 (U.K.) | +1 617 674 3256 (U.S.)

nlp@iqvia.com

iqvia.com/nlp