

Using Natural Language Processing At Novo Nordisk To Generate Actionable Insights From Real World Data

Novo Nordisk wanted to extract more value from the real world data (RWD) being collected about the use of its products, so that product managers and teams could better understand what patients and healthcare professionals were querying and reporting, and take appropriate action.

QUICK FACTS

Situation: Novo Nordisk wanted to identify healthcare market trends and detect patterns in clinical trial protocol deviations, patient sentiment, compliance, routines, behaviors, and treatment satisfaction and outcomes, from disparate RWD sources such as call center feeds and information from medical liaisons and healthcare providers. Manual scanning and extraction processes are labor-intensive and inefficient, and Novo Nordisk wanted a better solution—preferably automated, scalable and cloud-based.

Solution: Novo Nordisk already had experience using the Linguamatics NLP platform to text mine disparate data sources in separate in-house projects. To apply the NLP platform to RWD, first the Novo Nordisk team iteratively refined the platform algorithms to extract well-structured medical information from unstructured RWD sources. These data were then made available to product teams and business users via self-serve, Tableau-based dashboards for rapid discernment of actionable insights. In the second phase, Novo Nordisk then migrated the platform workflow and Tableau access methods to its AWS cloud-based big data platform for scalability, improved performance and data access.

Success: The Linguamatics NLP platform data extraction workflow replaced the need for manual scanning of RWD sources. With the new system, Novo Nordisk has reduced manual work by FTEs, reduced vendor spend, automated the process of generating insights and significantly broadened access to these insights across a global team.

Large amounts of RWD on Novo Nordisk product usage are available in multiple formats from medical information requests, field medical affairs and medical scientific liaison reports, customer call center logs and interactions with healthcare professionals. Mining these disparate structured and unstructured sources using traditional manual scanning and extraction techniques is time consuming and inefficient, and Novo Nordisk wanted to speed up, automate and scale the process, and provide staff with enhanced access to the extracted information for better derivation of actionable insights.

Novo Nordisk used the Linguamatics NLP platform to analyze and search a set of RWD sources and extract well-structured data, which was then amenable for access by product managers and teams via modern visualization and analysis tools. The data science team worked closely with medical subject matter experts to tune the NLP platform queries to detect and extract information with the desired level of precision and recall. The resulting information was then presented to business users via Tableau-based, self-serve dashboards, for exploration and insight generation.

The next step for Novo Nordisk was to provide better access to its data silos and data lakes via seamlessly integrated tools, and migrate the NLP platform and Tableau to its Amazon Web Services (AWS) cloud-based global big data and analytics platform.

Situation

Product managers and teams in pharmaceutical companies need to understand what patients and healthcare professionals are reporting about their drugs as they are used in the market, in order to discern trends and patterns, and respond appropriately. There are huge amounts of potentially useful RWD to help gain this understanding contained in electronic medical records, voice-of-the-customer call logs, field notes, and patient or healthcare professional information requests. Much RWD is contained in unstructured text, and the integration of different data sources (structured and unstructured) provides the best view for product managers, researchers and commercial product teams.

Novo Nordisk had used traditional manual scanning and extraction techniques to review and process various medical RWD sources, and make the information available to product teams and business users. However the extraction process was labor-intensive and inefficient; the resulting database was hard to search and not interactive; and scalability was becoming an issue, as data volumes and requests for access increased.

Novo Nordisk has good experience and previous success in using the Linguamatics NLP text mining platform on individual projects (see box). It had also developed an AWS cloud-based big data platform to provide storage and access to its data lakes and silos. Novo Nordisk wanted to take advantage of the performance, scalability and advanced analytics of the AWS platform to improve access to the extracted RWD, for faster and more powerful analysis and insight generation.

Solution

Novo Nordisk ran an initial project to use the NLP platform to text mine medical RWD on its products, so that the derived information could then be accessed by product teams via a medical patient dashboard.

The RWD sources to be mined were medical information requests (20,000 per year), field medical affairs notes (3,300 per year) and customer call center reports (130,000 per year). These were linguistically processed with NLP platform text analytics and indexed, using MeSH, NCI, MedDRA and in-house developed ontologies. The NLP platform queries were developed relating to Safety, Efficacy, PK/PD, Randomized Controlled Trials, Patient Populations, Dosing and Devices, and these were refined in close collaboration with pharmacist subject matter experts. Once the algorithms and post-processing were optimized, they were embodied in an integrated workflow that could be triggered as needed on a regular basis.

"We use Linguamatics [NLP] on both external and internal unstructured data. [The platform] is a very big time saver, and also introduces new capabilities for us—things we weren't able to do before because we didn't have the manual resources. [The platform] is very powerful; it offers you a lot of features, a lot of functionality. Frankly, there aren't other tools that are comparable—and I have seen quite a few tools, especially for unstructured data. [Linguamatics NLP] is definitely our tool of choice."

— Thierry Breyette, Associate Director, Information Analytics, Novo Nordisk

The extracted real world evidence was made available to business users via self-serve, Tableau-based interactive dashboards that can answer questions like, "What types of efficacy questions are being asked regarding Product X?", with interactive drill-downs

to explore who is asking the questions, where from and with what frequency. These dashboards provide rich and detailed information that can be used to explore, for example, commercial effectiveness, market development potential, and the need for extra product documentation or usage guidelines.

While the NLP platform text mining workflow and Tableau data access worked well, Novo Nordisk saw an opportunity to further accelerate the workflow, offer more frequent updates, deal with an increased number of disparate data sources and volume of legacy data, and address a global demand for service.

For the second part of the project, they migrated the NLP platform/Tableau solution to the cloud, as part of the Novo Nordisk AWS-based global big data and analytics platform, OASIS. The team used Python and Zeppelin Notebooks to create a data pipeline that extracts

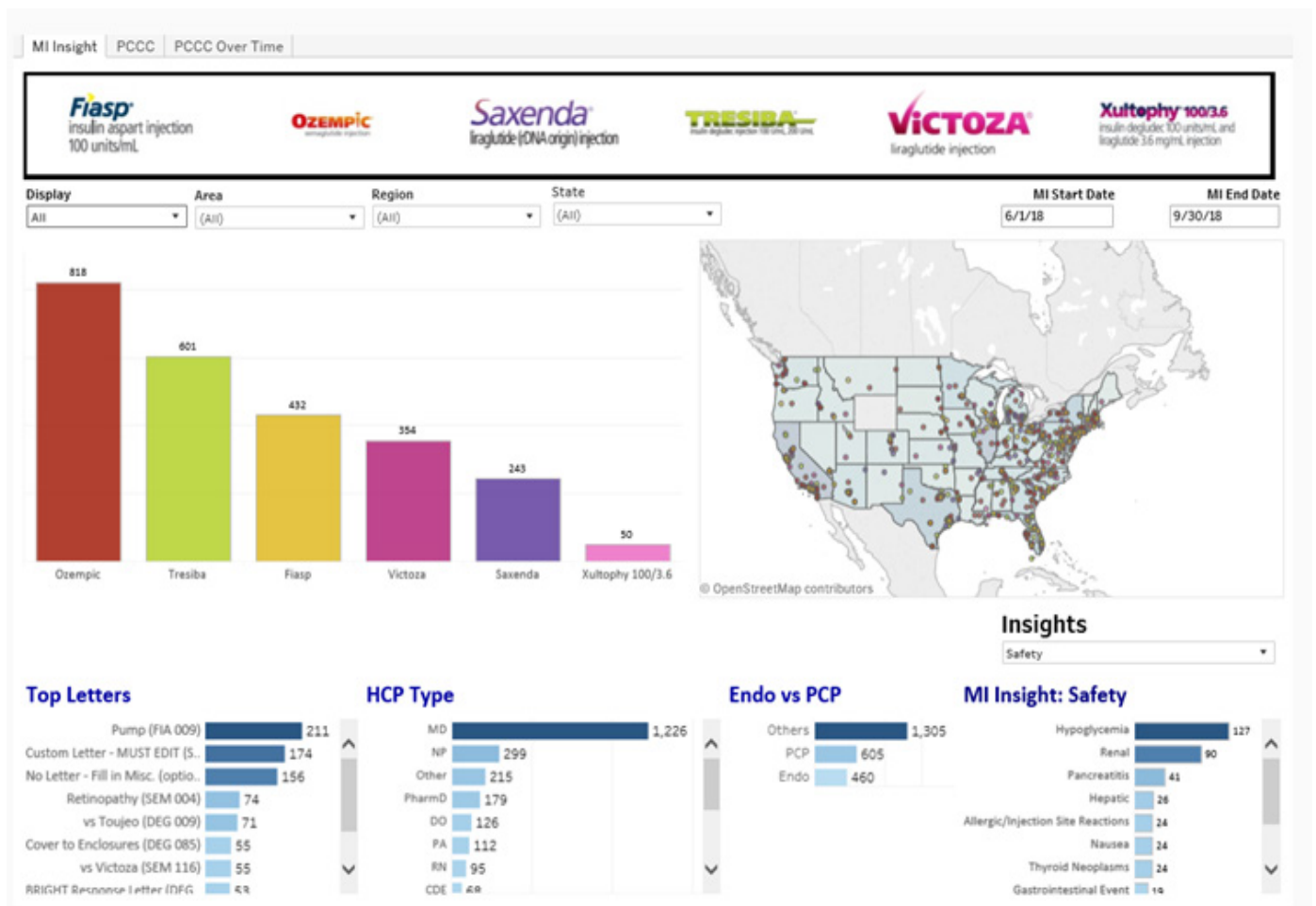
RWD source files from the data lake, runs the NLP platform workflow, and then automatically deposits the extracted results back into the data lake for analysis and visualization via Tableau. This enables enhanced access to the RWD information, and opens up future possibilities of using AWS' rich suite of advanced analytical tools (e.g. voice recognition for phone messages) in synergy with Linguamatics and Tableau solutions.

Success

Novo Nordisk had already exploited the power of NLP platform text mining to save time and extract actionable insights in a series of individual projects (see box).

To access RWD, Novo Nordisk was able to use the platform to efficiently extract and present structured information from three disparate RWD structured and unstructured data sources, and make it available

Novo Nordisk RWD Medical Insights Dashboard



This dashboard shows what questions are being asked as a set of products by healthcare professionals across the US. The system visualizes NLP platform-extracted data from medical information requests, field medical affairs and call center feeds (PCCC in the image). This interactive dashboard enables the stakeholders to generate their own insights and understand real world conversations about their product.

to product teams and business users via interactive Tableau-based dashboards, for exploration and insight generation.

Novo Nordisk successfully migrated the Linguamatics NLP text mining workflow and Tableau dashboards

to the Novo Nordisk AWS cloud-based big data visualization and analytics platform, which then provides improved performance, scalability, centralized data management in a data lake, advanced analytics and the ability to serve a growing global user population.

Linguamatics NLP platform use cases

Novo Nordisk use the power of Linguamatics NLP text mining across a number of use cases, including:

- Effective publication gap analysis. Using the NLP platform to mine PubMed extracts, a task that previously took around six weeks to complete across multiple staff, can now be done in a few hours and on a monthly basis. The time saved is equivalent to savings of over approximately \$100,000 per year.
- In a Health Economics Outcomes Research (HEOR) ethnographic insight study, the NLP platform is used to mine transcripts of patient/caregiver advisory boards and consumer councils, and extract valuable insights from hundreds of pages of conversational text.
- In a social media project to identify key opinion leaders for obesity, the NLP platform was used to mine Twitter feeds. Previously the medical affairs teams paid tens of thousands of dollars to outsource this work, but the NLP platform enables Novo Nordisk to bring these capabilities in-house for repeated use.
- In an ongoing study of clinical trial protocol deviations, the NLP platform has replaced an inconsistent manual process run quarterly with extracted information made available via interactive dashboards for deeper insights.
- In another project, Novo Nordisk uses Linguamatics NLP to uncover competitive insights from Dow Jones DNA news data.