Linguamatics
an IQVIA company

How natural language processing unlocks data to improve patient safety, quality, and reporting

A white paper

Linguamatics
an IQVIA company

Executive summary

As the healthcare industry evolves to embrace value-based payment models, patient safety improvement initiatives, and adherence to quality measures, natural language processing (NLP) has become increasingly important for its ability to unlock the value of unstructured data.

Healthcare enterprises have access to vast amounts of clinical, administrative, and billing data, especially now that most records are stored electronically. However, as much as 80% of this important data is stored in an unstructured format that is difficult to access and, thus, to glean actionable insights from. As a result, quality measure reporting and quality initiatives often rely on teams of nurses doing manual chart review to identify the vital information trapped in clinical notes.

Additionally, many healthcare enterprises struggle with data integrity and accuracy, as important patient data in electronic health records (EHRs) and other information systems is often missing or incomplete. A report from the ECRI Institute on top patient safety concerns for healthcare organizations (HCOs) found that data integrity was the second-most significant technology hazard, following only hazards caused by alarm fatigue.[1]

While the U.S. health system has made progress in recent years, patient safety remains a problem that all HCOs must prioritize. For example, an estimated 1.7 million healthcare-associated infections occur each year in the U.S. leading to 99,000 deaths. Further, adverse medication events cause more than 770,000 injuries and deaths each year at a cost as high as $5.6 billion annually, according to statistics cited by the Center for Patient Safety.[2]

With the pressures of value-based care and increased regulatory reporting, how can HCOs improve the efficiency of these laborious tasks?

To aid in this effort—and capture additional details from unstructured data—more HCOs are today embracing NLP tools. NLP tools make unstructured data usable by automating the identification and extraction of key concepts from large volumes of clinical documentation. Findings are transformed into structured data to streamline chart review and identification of high-risk patients, and hence improve the efficiency of quality initiatives and quality measure reporting.

As healthcare continues to evolve, the use of NLP to reduce manual efforts, extract key concepts from clinical notes, and quickly find at-risk patients is invaluable for meeting reporting requirements, improving the quality of care, and optimizing clinical outcomes.

[1] https://www.ecri.org/EmailResources/HRC/WebinarResources/ECRI%20Institute%20Top%2010%20Patient%20Safety%20Concerns%202015_Webcast%20Jun.18.15.pdf

[2] https://www.centerforpatientsafety.org/facts-stats/

Introduction

Several industry trends have coalesced to drive greater use of NLP in healthcare. The first is simply the sheer volume of healthcare data that exists today compared with years ago, as a result of the digitization of health records.[3] In addition to clinical notes and specialist reports from radiology and pathology departments in EHRs, a significant quantity of data is now coming from patient-reported information in portals and emails.

This influx of data has spurred HCOs to seek out ways to collect, analyze, and use this information to improve operations and patient care. In the absence of artificial intelligence (AI) tools such as NLP, health systems must resort to expensive, manual chart reviews by nurses and clinicians.

Additionally, healthcare's ongoing evolution from fee-for-service to value-based care has made it more important than ever to track patient safety, close care gaps, and improve quality efforts for reporting, so that HCOs can be appropriately reimbursed under value-based contracting arrangements. As HCOs are increasingly reimbursed based on the outcomes they produce, it has also become essential to address issues related to social determinants of health (SDoH), which often significantly impact health and outcomes. However, critical SDoH information—such as social isolation, food insecurity, transport issues, lifestyle choices, and living conditions—is not always readily available because it is trapped in clinical notes. Its existence is more commonly known after the information has already been a contributing negative health factor, which presents a challenge in delivering timely, actionable information to clinicians. This missing or incomplete SDoH information can lead to patient safety concerns, as clinicians may develop treatment plans that do not account for important social factors in patients' health.

Finally, advances in AI-based technologies have made these solutions more accessible to HCOs of varying sizes. AI-based technologies such as Linguamatics NLP platform are increasingly sophisticated and don't necessarily require teams of expensive data scientists who may take months to build usable algorithms.

HCOs currently leverage Linguamatics NLP in a wide variety of use cases to advance patient safety, streamline operations, and improve quality of care and reporting initiatives.

Patient safety

Patient safety is the No. 1 concern for HCOs. Errors are inevitable in an age when long clinical hours and skeleton crews are increasingly becoming standard operating practice. The possible consequences of errors and patient safety issues remain ominous, but NLP workflows can help reduce the likelihood of error and improve patient safety. Here are a few specific use cases that illustrate how NLP deployments are making a positive impact on patient safety, quality, and reporting.

[3] https://www.buildingbetterhealthcare.co.uk/technical/article_page/Comment_Health_networks__delivering_the_future_of_healthcare/94931

Linguamatics
an IQVIA company

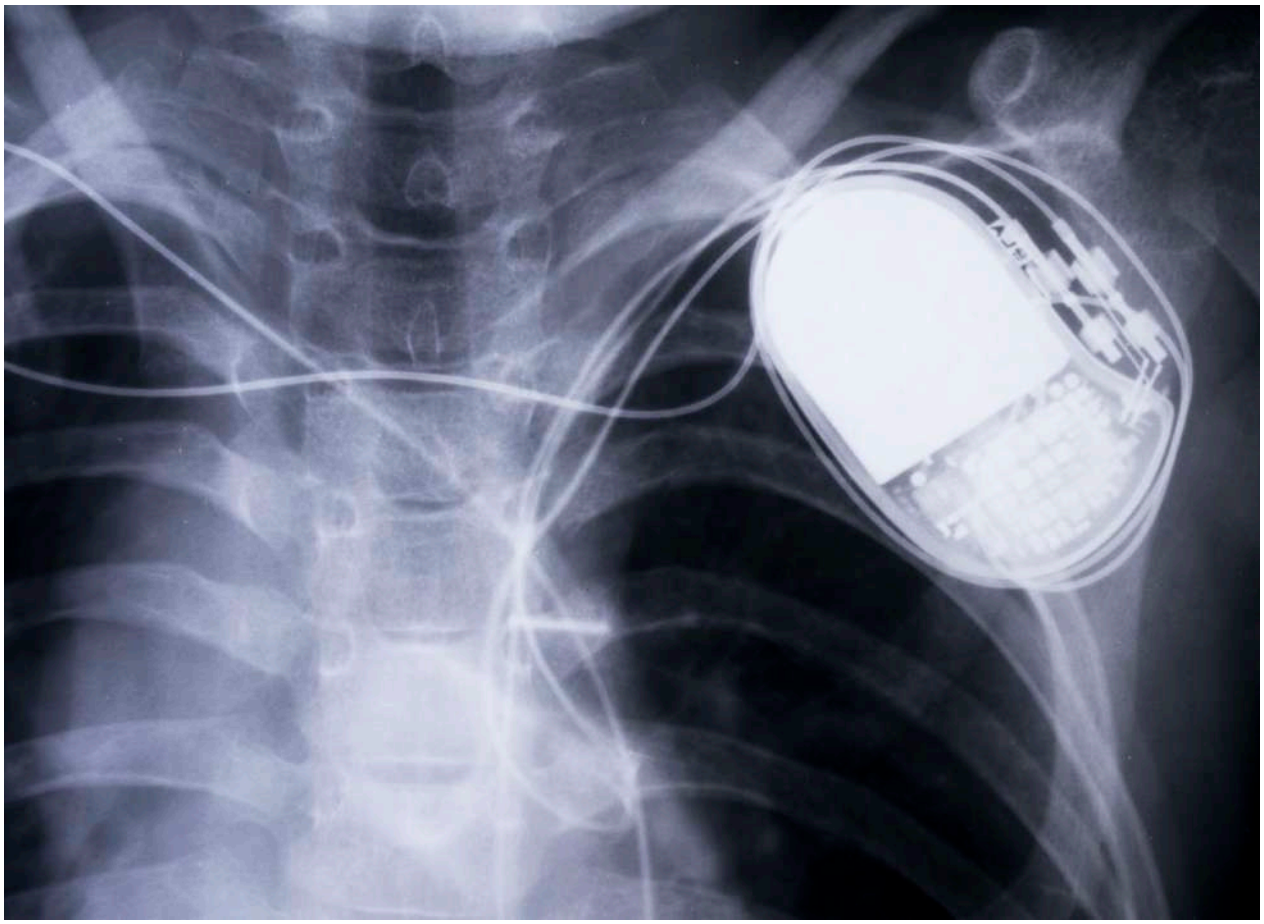
Case study: How an accountable care organization improved clinical documentation and closed care gaps

An accountable care organization (ACO) serving Medicare, Medicaid, and commercial populations needed ready access to clinical notes and data, in order to improve clinical care and report on care activities to payers. The ACO is required to track and report on activities that demonstrate improvements in the delivery of quality care, alongside any financial savings, to show that the care provided is consistently high value. For example, heart failure reporting requires information on the ejection fraction for every patient in the covered population. To accomplish these reporting and improvement goals, the ACO must identify at-risk patients to minimize gaps in care quality, and include these individuals in safety-net programs.

In the past, the ACO had difficulty obtaining certain clinical information, including the comprehensive identification of patients with certain medical conditions, and the tracking of quality metrics for some groups of patients, because the required information was captured in clinical notes. The ACO's clinicians frequently input clinical details in narrative form as free text, particularly when structured fields for such information are unavailable, but also due to an industry-acknowledged problem with clinician burnout. The narrative is necessary to capture clinical complexity and nuance, and is frequently preferred to structured data entry; however, it creates data-completeness challenges because the free-text information does not automatically translate into structured information on patient problem lists.

Recognizing that it needed to eliminate the inefficiencies of manual chart reviews to streamline reporting processes, the ACO identified Linguamatics NLP as a technology that could help users identify critical clinical data hidden within the free-text fields of clinical reports and clinician progress notes. They implemented an NLP pipeline to improve manual chart review that resulted in a significant reduction of charts being reviewed—with manual chart review, nurses had to review 1,000 charts to identify one care gap, but Linguamatics NLP reduced manual review to only six charts for each successful care gap identified.

Since implementing the Linguamatics NLP-based AI solution, the ACO has been able to close information gaps, which in turn has allowed closure of care gaps. For example, the ACO leveraged NLP to identify 92 patients who were documented in the narrative as having chronic obstructive pulmonary disease or congestive heart failure, but whose conditions were not entered into a structured format. These patients became eligible for the ACO's population-based disease management programs, and chronic disease-related care gaps were closed. In addition, the ACO could now recognize that those patients needed risk adjustment for care-budgeting purposes. By identifying these patients, the ACO received $75,000 to $150,000 per disease area in additional annual risk-adjusted revenue to support the care of those individuals. Nurses also experienced a nearly 200-fold increase in the efficiency of their chart reconciliation process, problem lists were more accurate and complete, and ACO reporting simplified.



Case study: Identifying patient safety issues and improving long-term outcomes

NLP also offers the potential to improve the safety and effectiveness of medical devices. By combining real world evidence of symptoms or conditions present before and after device or medication usage, along with captured adverse events, medical device and drug manufacturers can gain a greater understanding of how devices perform outside of a controlled clinical trial.

For example, a population-scale project by a large health system mined EHR data to evaluate heart failure device performance—letting the manufacturer know how to improve its implantable products as part of a contract research project, and helping the health system's own clinicians make better data-driven decisions on treatment.

The health system used Linguamatics NLP technology to extract this previously inaccessible data from 11 years of almost 34 million clinician notes from both inpatient and outpatient encounters for its cardiac patients. They extracted key cardiology measures (ejection fraction measurements; symptoms such as shortness of breath, fatigue, and palpitations; and New York Heart Association classifications) that they can now analyze as discrete data sets. The team was able to complete its project on time and at a 95–99% accuracy level. Most importantly, the whole analysis was completed in three months by two analysts with no previous experience in NLP—in contrast with an improbable, manual task that would have taken over 55 full-time-equivalent years to complete without NLP.

Linguamatics
an IQVIA company

Case study: Patient safety nets identify high-risk individuals in real time

The ACO mentioned previously has also identified opportunities to close patient safety gaps with real-time monitoring of patient documentation using Linguamatics NLP. For example, radiology reports sometimes include incidental findings such as lung nodules. Although not always the case, there are times when their presence indicates something serious, such as early-stage lung cancer. In an emergency room or urgent care situation, findings may not be properly handed off for further evaluation, resulting in a later-stage diagnosis, worse outcomes, and potential litigation. As a result of its NLP implementation, the ACO now checks all radiology reports for evidence of pulmonary nodules and other incidental findings, and at-risk patients are flagged for follow-up with care coordinators.

Case study: Reducing complications from surgery-site infections

A United Kingdom-based academic medical center adopted Linguamatics NLP to advance translational research and clinical practice to improve quality and safety. Specifically, the medical center has leveraged NLP to identify patients who may have surgical-site infections (SSIs), which are among the most common healthcare-associated infections and account for about 16% of all healthcare-associated infections in England. SSIs are globally linked with increased mortality rates, readmission rates, length of stay, and healthcare costs.

The medical center set out to develop a decision support tool that would assist clinicians in predicting cases of SSI, but quickly discovered that much of the needed information was embedded in unstructured data within patient clinical records. Rather than manually reviewing a vast amount of records, leadership decided to adopt NLP to help determine which patients had an SSI and classify the type of SSI.

Leveraging the NLP solution, the medical center created algorithms that mine unstructured clinical notes to identify and categorize SSIs based on the U.S. Center for Disease Control and Prevention's classifications of superficial incisional, deep incisional, and organ/space SSIs. The medical center is also using Linguamatics NLP to identify complex social situations that may influence patient outcomes; rapidly shift through imaging reports to identify cases of complex conditions, such as invasive fungal infections; and better understand the impact of early signs, symptoms, and diagnoses on patient outcomes.

Case study: Cohort identification for high-risk patient groups

NLP approaches can also be leveraged to help population-level cohort selection, to identify at-risk populations and find patients requiring follow-up due to a product recall.

For example, The Centers for Medicare & Medicaid Services has developed a code for lung cancer screening that targets 55 to 77 year-olds who are current or past smokers, have no lung cancer diagnosis, and have more than 30 pack-years of smoking. While these details have structured fields in modern EHRs, certain critical risk factors such as smoking pack-years will change over time, and may only be updated in the narrative. Linguamatics NLP can be used to extract these factors to derive a much deeper understanding of clinical risk, and invite people who meet the factors for CT screenings to identify early signs of lung cancer.

Addressing social determinants of health in population risk stratification

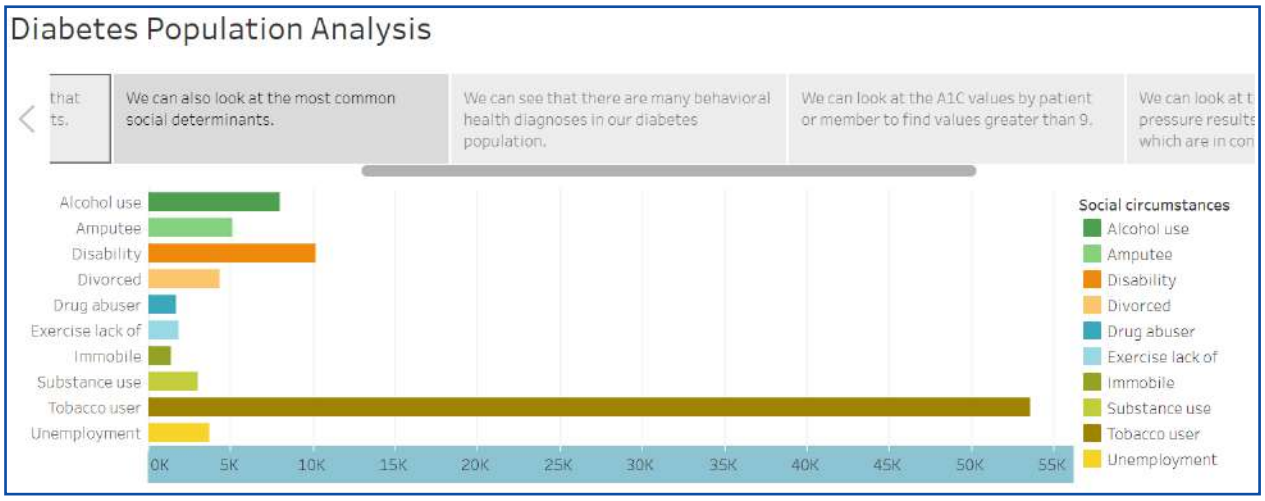
Today's HCOs increasingly realize that, in addition to diagnoses, procedures, and similar medical data, effective population health management requires details on SDoH, including factors such as lifestyle choices, living conditions, and health behaviors.

Consider, for example, how outcomes are negatively impacted when an HCO is unaware that a patient has mobility issues, lives alone, and lacks reliable transportation, resulting in difficulty traveling to medical appointments or picking up medications at the pharmacy. NLP techniques are able to extract these features from nurse notes and enable them to feed into predictive risk models. To ensure the patient receives the care required for optimal outcomes, the HCO can connect the patient with social service organizations, to arrange for transportation to medical appointments and the at-home delivery of medications. SDoH details also help providers identify patients who are more likely to be readmitted to the hospital within 30 days of discharge due to a lack of family and social support, or because of a history of risky health behaviors.

Additionally, payers also utilize NLP technologies to extract member insights from medical records and call center notes in order to improve population risk stratification. One top-five payer has implemented the Linguamatics NLP platform within an automated workflow to integrate NLP results with its structured claims data in Hadoop systems; this improves its ability to stratify risk for individuals with congestive heart failure. The payer uses NLP to mine risk factors such as mentions of disease and disease severity such as Body Mass Index, along with lifestyle factors like smoking and SDoH factors such as social isolation. Data is extracted and transformed into a structured format that is uploaded into Hadoop HIVE to create detailed insights for population risk stratification.

> Sometimes physicians focus excessively on the "medical" problems and don't pay enough attention to the context that people live in and the social aspects that influence their health. Our study [utilizing Linguamatics NLP] once again highlights the importance of knowing this information in order to provide patients our very best care.
>
> Leslie Lenert, M.D., MS, Chief Research Information Officer, MUSC and Director, MUSC's Biomedical Informatics Center

Figure 1: Social determinants of health can be extracted from clinical documentation using Linguamatics NLP. The resulting data can be loaded into a data warehouse to support patient dashboards, or used in predictive risk models.



Streamlining quality reporting

A 2016 Health Affairs article on the cost of reporting quality measures revealed that U.S physician practices spend more than $15.4 billion annually to accomplish this task.[4] This estimate was based on time spent and the salaries of the qualified individuals involved in the process.

Quality reporting involves measures for physicians, Preferred Provider Organizations, payers, and others. The Healthcare Effectiveness Data and Information Set is one of the more widely used performance improvement tools created by the National Committee for Quality Assurance (NCQA). NCQA is committed to continually developing patient-centric metrics that encourage clinicians to personalize care for patients, motivating them to take an active role in improving their own health. Achieving adherence with such advanced measures is likely to require the use of NLP techniques to find linguistic patterns that reveal the underlying unstructured data needed to personalize care.

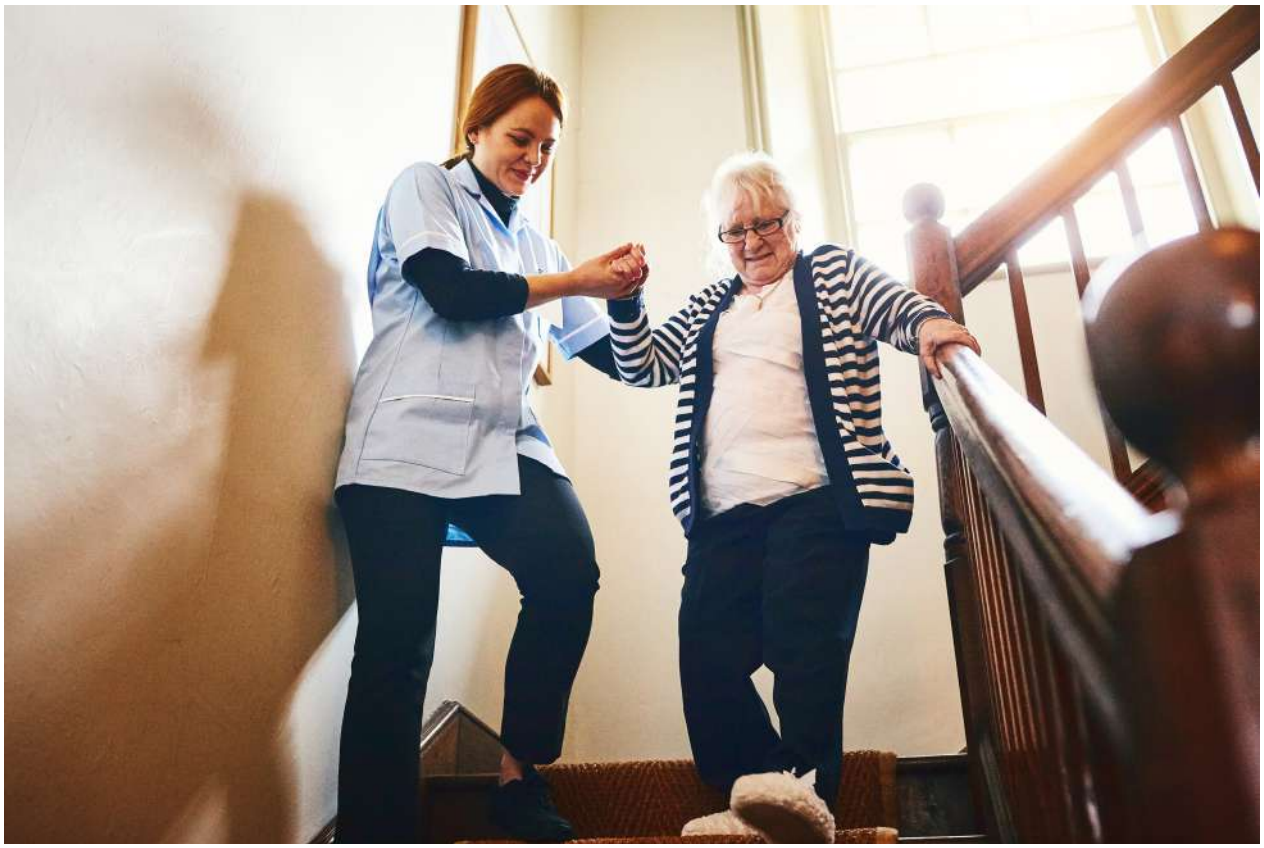
Prior to this advanced patient-centric initiative, other NLP efforts have been used for quality metrics such as fall risk screening. Falls are a leading cause of injury among elderly patients, with the risk of falls increasing as patients age. Clinical screening of fall risk can significantly prevent falls, but traditionally many patients who have experienced a fall do not discuss the event with their providers. With the onset of value-based care, fall risk screenings have increasingly become a high-priority quality measure for payers and providers. Although a particular medical code exists for this screening measure, coded data alone often does not capture clinical activities associated with this measure.

Researchers affiliated with an academic medical center sought to discover whether quality reporting that relies on coded administrative data alone could accurately and completely measure a provider's performance in fall risk screenings.

[4] https://www.healthaffairs.org/doi/10.1377/hlthaff.2015.1258

To analyze the issue, they developed an NLP solution with Linguamatics that identified fall risk screening documented in the clinical notes portion of patients' records, without coded fall risk screening data present in the record.

The study's results showed that, out of 144 eligible patients, the Linguamatics NLP solution identified 59 patients with fall risk screening documented only in the clinical notes and not coded. These 59 patients had previously been reported by their primary care physicians as having no fall risk. The researchers concluded that the results supported the concept that using both structured coded data and clinical narratives for quality reporting is superior to the current reporting approach based on administrative coded data alone.



Summary: From unstructured data to improved outcomes

As the prior use cases demonstrate, NLP is capable of delivering a significant, positive impact on patient safety, quality, and reporting, as well as many other important functions in the industry. As digitized healthcare data proliferates, so does the problem of unstructured data, in which EHRs contain information that is valuable, but is essentially held under lock and key. NLP represents the key that unlocks this vast amount of unstructured data—the first step toward transforming data into actionable insights that enable HCOs to make better decisions, leading to better outcomes and more comprehensive, cost-effective care.

Linguamatics
an IQVIA company

About Linguamatics

Linguamatics, an IQVIA company, delivers market-leading NLP-based AI solutions for high-value knowledge discovery and decision support from text. We empower our customers to speed up drug development and improve patient outcomes by breaking down data silos, boosting innovation, enhancing quality, and reducing risk and complexity.

Our award-winning NLP platform is proven across multiple real-world use cases. Linguamatics has been trusted for over 15 years to deliver actionable insights that address your most pressing bench-to-bedside challenges with quantifiable ROI.

Our customers include 18 of the top 20 global pharmaceutical companies; the US Food and Drug Administration (FDA); and leading cancer institutes, hospitals, and academic research centres. Linguamatics NLP has been deployed by organizations in pharmaceuticals, biotechnology, healthcare, chemicals and agrochemicals, Government, and academia.

Linguamatics is committed to excellence in healthcare informatics and is a corporate member of AMIA and HIMSS. The company operates globally, with headquarters in Cambridge, UK, and a U.S. office near Boston, MA.

If you are interested in learning more about Linguamatics NLP, please contact us at:
enquiries@linguamatics.com

Linguamatics UK
324 Cambridge Science Park
Milton Road
Cambridge CB4 0WG
Tel: +44 (0)1223 651910

Linguamatics US
500 Nickerson Road, Suite 125
Marlborough
MA 01752
Tel: +1 617 674 3256