

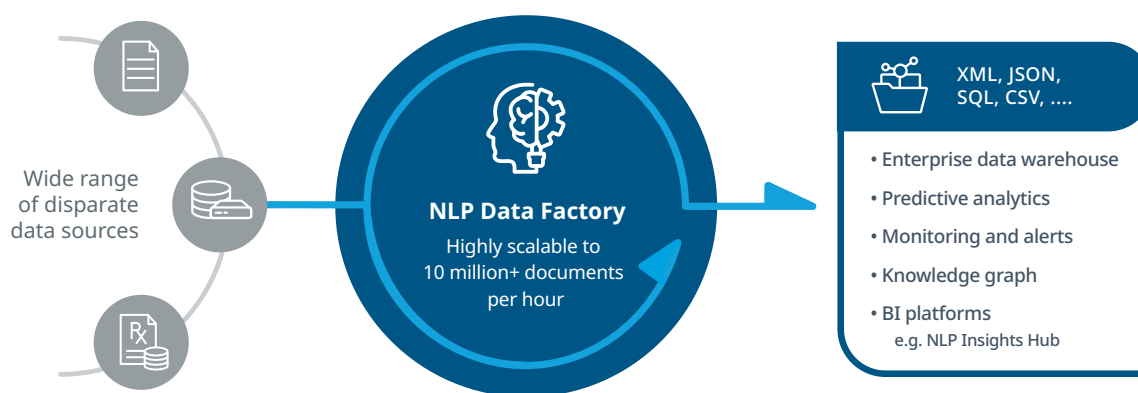
# IQVIA Natural Language Processing (NLP) Data Factory

*Deliver comprehensive, systematic value from textual data across your business*

The importance of data in driving clinical and commercial outcomes is absolute, yet many life science and healthcare organizations leave 80 percent of potential insights untapped, locked away in unstructured and semi structured formats. Manual abstraction of these vast amounts of data is time-consuming and unsustainable, and niche point solutions can't extract and integrate information across multiple business areas.

The IQVIA NLP Data Factory accurately surfaces and normalizes features of interest at scale, in an automated, robust and easily configurable pipeline. Transform efficiency and productivity across the enterprise with reproducible extraction, enrichment and delivery of critical data.

## Automate critical data extraction, enrichment and delivery



## YOUR CHALLENGES

- Missing important information hidden in vast amounts of unstructured and semi structured text
- Staff spend huge amounts of time and expense on manual data curation – including inconsistencies in structured data where the same data is recorded in many ways
- Expensive point solutions are narrow in scope and lack critical domain knowledge
- Lacking the tools to transform and prepare data for analysis and use at scale

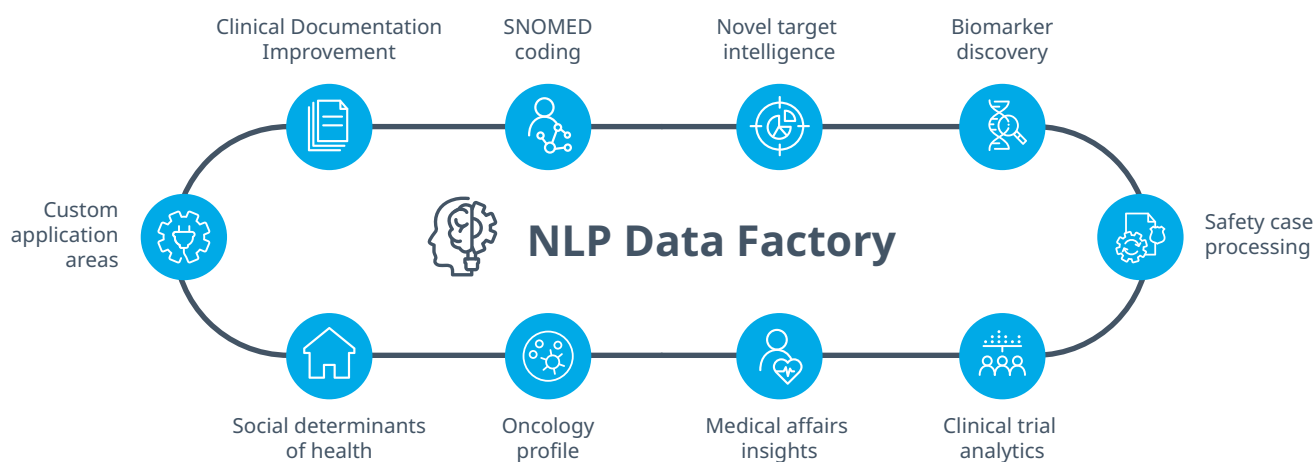
## YOUR NLP DATA FACTORY SOLUTION

- Mobilizes value at scale from the key data your organization has generated or invested in
- Automates NLP workflows which process millions of documents and data fields across a range of business lines, normalizing data and enhancing metadata for easy consumption
- Easily and efficiently scales to add new use cases to the robust NLP Data Factory pipeline
- Delivers ready-to-use data and solutions that improve business outcomes

## NLP DATA FACTORY KEY FEATURES

- Effective robust solution for extracting value from textual data - think of it as intelligent ETL (Extract -Transform-Load) of free text.
- World-class NLP to surface and transform data from a wide range of formats to common data standards outputs
- Integrated OCR to ensure scanned text is not left behind
- Embeds seamlessly into existing enterprise architectures and workflows
- Deployable on-premise, in cloud environments or with a hybrid implementation
- Recognizes and normalizes complex constructs such as cancer staging
- Seamlessly incorporates trained ML models

## Examples where NLP Data Factory can transform your business



## TECHNICAL OVERVIEW

**Industry-proven NLP technology.** The NLP Data Factory uses IQVIA NLP technologies to power the normalization and standardization of input data. This blend of methods combines rule-based queries, machine learning, terminology matching, pattern extraction and relationship identification to ensure the highest possible accuracy for the task in hand.

**Fast, scalable, architecture.** The components required to power the NLP Data Factory has been developed together to optimize the efficiency of the system. An internal orchestration component is designed to parallelize incoming data, ensuring that scaling is effective and matches the availability of resources.

Contact [nlp@iqvia.com](mailto:nlp@iqvia.com) for more information about how your organization can repeatedly harness and utilize your heterogeneous unstructured data at scale with the NLP Data Factory.

**Flexible NLP framework.** The flexible nature of the IQVIA NLP query engines ensures that new modules can be dropped into the NLP Data Factory with no effort. These modules can be used right away or can be tuned further using a powerful browser-based query editing tool.

**Easy deployment via Kubernetes.** Components in the system are containerized for simpler management. Furthermore the full system is deployed using Kubernetes or equivalent, allowing for simpler installation, easier service monitoring and automated scaling of the system.