

Experience the power of **DISCOVER** with natural language processing for text mining



ONTOFORCE

Transform even more data
into knowledge, even faster

DISCOVER, ONTOFORCE's powerful data and knowledge platform, built exclusively for the life sciences industry, now comes with advanced **natural language processing (NLP)** functionality for text mining.

With DISCOVER's enhanced NLP capability, you can now extract valuable insights and knowledge from unstructured text data, including scientific literature, news articles, clinical trials, and more.

DISCOVER's NLP technology comes with a powerful named entity recognition (NER) and an extraction engine which allows you to quickly find and analyze information about specific biological entities and their relationships (such as diseases, genes, proteins, and more), populations, interventions, and outcomes that are mentioned in large volumes of text data.

All with one single solution.



Benefits

One-stop-shop

Uncover hidden insights from data - whether it's structured or unstructured.

Pre-processed enhanced public data out-of-the-box

Begin working with enhanced public data immediately.

Process private data with pre-configured yet adjustable annotators

Enhance your private data with out-of-the-box, customizable annotators that can easily be finetuned.

ONTOFORCE's team of experts at your disposal

Our team of data scientists, machine learning experts, and bioinformaticians are available to support with implementation and user roll-out.

Modular and flexible

Available as packs: no full-blown solution required, thus making it more convenient for you from a deployment and licensing perspective. Simply license and run the packs you need for either a single use case or multiple.

Enhance DISCOVER with natural language processing packs

1. AVAILABLE OUT-OF-THE-BOX FOR PUBLIC DATA:

Uncover insights and knowledge that were previously hidden in the text of articles and publications. Our NLP models have been trained to extract PICOS, biomedical entities, and relationships from plain text. These are used to annotate and enrich DISCOVER's public data, which is available via remote data subscription (RDS) packages that are ready to ingest without the need for any additional deployments.

The NLP enhancement for public data uses two models:

- The **PICOS model**: retrieve population, intervention, comparison, outcomes, and study type data from the contents (such as descriptions and inclusion/exclusion criteria) of publications and clinical studies.
- The **Biomedical Entities model**: distills biomedical data from the descriptions of clinical trials and literature.

Often, the publicly available clinical studies do not specifically annotate patient demographic, intervention, outcome, or biomarker information in a structured way, meaning important knowledge or studies may be missed. Through NLP, DISCOVER can reach this siloed knowledge, add it to the knowledge graph and make it available for search and exploration.

2. APPLY ON YOUR PRIVATE DATA:

Find more hidden information inside your own private data.

The NLP functionality for private data comes with the following three models:

- The **PICOS model**: retrieve patient population, intervention, comparison, outcomes, and study type data from your internal documents as well as from the data in your own internal repositories such as clinical trial management systems.
- The **Biomedical Entities model**: distills biomedical data from your internal document repositories.
- The **Informed Consent model**: extracts relevant classification for primary and secondary usage of samples and data from consent forms to facilitate data reusability.

Available NLP enhanced packs for DISCOVER

	Data sources	Data types	Processing	Implementation
NLP Enhanced Clinical Pack - Public data	Clinical studies Publications	Annotators for <ul style="list-style-type: none">• Biomedical entities• PICOS criteria	Pre-processed	Out-of-the-box, via RDS packages
NLP Enhanced Clinical Pack - Private data	Internal document repositories such as clinical trial management systems Informed consent forms	Annotators for <ul style="list-style-type: none">• Biomedical entities• PICOS criteria• Informed consent forms	Allows for the processing of private data only	Customizations to your own data ingestion pipeline

Both NLP enhanced packs can be deployed for a single use case such as Clinical Trial Design or Clinical Insights, or for multiple use cases at once.

NLP enhanced packs for R&D, Regulatory, and Cross Functional Intelligence (Competitive Intelligence, External Innovation and more) use cases are coming soon. Contact us to learn more about the availability of these packs.

Get in touch to schedule a demo

 ontoforce.com
 info@ontoforce.com

