

# What is Natural Language Processing (NLP)?

July 2022



## NLP is a form of AI which processes language.

Natural Language Processing (NLP) has existed for more than 50 years and has roots in the field of linguistics. It has a variety of real-world applications in a number of fields, including medical research.

NLP is opening up new possibilities for how we can evaluate free speech to help identify and predict psychiatric illness in patients. As NLP and machine learning capabilities become more sophisticated, automated speech analysis can offer predictive power that exceeds today's standard clinical evaluation methodologies and scoring systems

**NLP is a branch of Artificial Intelligence (AI) that endeavours to develop the automated computational processing of human (natural) language. Whether the language is spoken or written, natural language processing uses artificial intelligence to take real-world input, process it, and make sense of it in a way a computer can understand.**

That is to say develop machines that are able to read, understand, extract information from, and even converse in, human language. Most modern forms of NLP utilise Machine Learning (ML). The process of ML involves developing algorithms that allow computers to learn from (typically large) datasets and perform tasks with increasingly greater accuracy without being explicitly programmed.

## Why is NLP important?

### Large volumes of textual data

Considering the staggering amount of unstructured data that is generated every day, from medical records to social media, automation is critical to analysing text and speech data efficiently. NLP has various applications in health care today, from assisting with clinical documentation and coding, supporting clinical decision-making, to supporting mental health treatment.

We can access massive quantities of unstructured, text-heavy qualitative data and need a way to efficiently process it. This is especially true in psychiatric health records, where up to 70-80% of the information available exists as unstructured clinical notes. NLP is important because it helps resolve ambiguity in language and adds useful numeric structure to the data. The technology to analyse unstructured text actively learns from the data as it comes in by combining machine learning with human direction to generate new insights.

## How do we use NLP models at Akrivia Health?

Our team of experts trained NLP to use a state-of-the-art language model called BERT (Bidirectional Encoder Representations from Transformers). This model has been

pretrained on a vast text dataset with the key innovation of using self-attention. This allows the model to focus on different semantic aspects within a sentence and gives it a deeper sense of language context compared to other neural language models.

To systematically extract information from the free text, we follow a three-step procedure:

- Model 1 - Named Entity Recognition
- Model 2 - Contextual Classification
- Model 3 - Relationship Extraction

The combination of the three models gives the outputs for all our concepts in our schema.

## Benefits of Natural Language Processing

Since between 70-80% of essential patient data lies in unstructured clinical notes, gaining immediate access to clinical information is difficult. Researchers can use artificial intelligence methods to sort through unstructured data by using natural language processing. The data can then provide valuable insights into patient care, research efforts, and disease diagnosis.

An example of the use of NLP:

The aim of the study was to identify a group of patients presumed to have Difficult-to-treat-depression (DTD) in UK specialist mental health National Health Service (NHS) Trusts and to examine demographic, disease and treatment data through

the analysis of secondary-care mental health records ('real-world data'). An NLP model was used to analyse anonymised electronic health records (EHRs) of five specialist mental health National Health Service (NHS) Trusts in the United Kingdom. Data on disease characteristics, comorbidities and treatment histories were extracted from structured fields and using natural language algorithms from unstructured fields. (<https://akriviahealth.com/insight/difficult-to-treat-depression/>)

Without accurate and systematic case identification, population management and research, understanding psychiatric conditions and other complex conditions is not possible.

NLP can support with identifying the correct treatment options for patients, as well as assisting in disease diagnosis.

**Natural Language processing holds the power to improve how we live and work. It can help bring progress to areas that have been slow or difficult to change without the partnership between human and technology.**

*Our approach to accelerating research is grounded in the principle of data protection by design and default. The technologies and methodologies we develop and used are designed to both enable research whilst simultaneously protecting privacy and confidentiality. This is true innovation.*



Email: [contact@akriviahealth.com](mailto:contact@akriviahealth.com)  
 Telephone: 0330 332 4878  
[www.akriviahealth.com](http://www.akriviahealth.com)

CRISal Health Limited t/a Akrivia Health, a company registered in England and Wales, whose registered offices are Clarendon House, Cornmarket Street, Oxford, United Kingdom, OX1 3HJ, company number: 11892096

**Change your trajectory today.**