# PRIVACY PRESERVED PATIENTS REPORT SUMMARY

**M. Krishnaveni** Research Scholar Department of Computer Science and Engineering Avanthi Institute of Engineering and Technology Cherukupally (Village), Near Thagarapuvalasa Bridge, Vizianagaram (Dist)-531162
**Dr. Gandi Satyanarayana,** Professor and Head Department of Computer Science and Engineering Avanthi Institute of Engineering and Technology Cherukupally (Village), Near Thagarapuvalasa Bridge, Vizianagaram (Dist)-531162
**Dr. Akula Chandrashekar** Professor and Director Department of Computer Science and Engineering Avanthi Institute of Engineering and Technology Cherukupally (Village), Near Thagarapuvalasa Bridge, Vizianagaram (Dist)-531162

**Abstract**
Syntactic parsing (NLP) is a fundamental operation of Natural Language Processing. However, there hasn't been much research into syntactic parsing in the area of medicine. Several methods were thoroughly investigated in this research paper to improve the efficiency of syntactic parsing of medical text, including (1) building a pair of clinical tree banks of discharge summaries and updates by developing annotation guidelines that manage absent components of clinical sentences; (2) retraining four cutting-edge parsers, including the Stanford parser, Berkeley parser, Charniak parser, and Bikel parser, using clinical tree banks, and nonclinical tree banks greatly improved the functionality of existing parsers, according to our investigation. Clinical tree banks greatly improved the functionality of existing parsers, according to our investigation. Our suggested PP attachment tactics improved accuracy by 2.35% on the MiPACQ corpus and 1.77% on the I2b2 corpus. Our technique produced synchronization precisions of 94.9% and 90.3% for the MiPACQ and i2b2 corpora, respectively. To demonstrate the success of the revised parsing approaches, we used the outcomes of the parsers for two external NLP tasks: semantic role naming and temporal relation extraction. The results of the experiments showed that using the parse tree data from our optimized parsers enhanced efficiency for both tasks, with improvements in F-measure of 3.26% and 1.5% for semantic role labeling and temporal relation extraction, respectively. To extract, transport, and load patient data in a secure manner, we encrypt data using the AES technique.

**Keywords**:
Voice recognition, text to text summarization, TF-IDF vectorizer, sentence tokenization, Natural Language Tool Kit (NLTK), K-means clustering, and accuracy are all terms related to natural language processing (NLP).

## Introduction

It is vital to integrate, organize, and make use of enormous amounts of genetic, pharmacological, cellular, tissue-based, environmental, and clinical information in order to progress healthcare. The Clinical Data Architecture (CDA) is an XML-based standard exchange mechanism for patient data that is projected to improve the exploitation and accessibility of information in electronic health records (EHR), consequently considerably enhancing biological research and treatment. Patient reports in the EHR primarily communicate patient information using natural language. Despite being a useful source of information, patient reports are difficult for computer systems to appropriately retrieve due to text. This is a significant impediment to the widespread deployment of effective therapeutic applications.

Natural Language Processing (NLP) systems are automated systems that manage information in texts by utilizing linguistic expertise.NLP applications in healthcare include decision support, infectious disease surveillance, research studies, automated encoding, quality assurance, indexing patient information, and billing systems. In recent years, NLP systems have been used in hospital settings to