

# Machine Translation Approaches and Survey for Indian Languages

Antony P. J.\*

## Abstract

The term Machine Translation is a standard name for computerized systems responsible for the production of translations from one natural language into another with or without human assistance. It is a sub-field of computational linguistics that investigates the use of computer software to translate text or speech from one natural language to another. Many attempts are being made all over the world to develop machine translation systems for various languages using rule-based as well as statistically based approaches. Development of a full-fledged bilingual machine translation (MT) system for any two natural languages with limited electronic resources and tools is a challenging and demanding task. In order to achieve reasonable translation quality in open source tasks, corpus based machine translation approaches require large amounts of parallel corpora that are not always available, especially for less resourced language pairs. On the other hand, the rule-based machine translation process is extremely time consuming, difficult, and fails to analyze accurately a large corpus of unrestricted text. Even though there has been effort towards building English to Indian language and Indian language to Indian language translation system, unfortunately, we do not have an efficient translation system as of today. The literature shows that there have been many attempts in MT for English to Indian languages and Indian languages to Indian languages. At present, a number of government and private sector projects are working towards developing a full-fledged MT for Indian languages. This paper gives a brief description of the various approaches and major machine translation developments in India.

**Keywords:** Corpus, Computational Linguistics, Statistical Approach, Interlingua Approach, Dravidian Languages0.

---

\* Professor and Head, Department of ISE, St. Joseph Engineering College, Mangalore, VTU.  
E-mail: antonyjohn@gmail.com