



# Machine translation status of Indian scheduled languages: A survey

Nawaz Ali Lone<sup>1</sup> · Kaiser J. Giri<sup>1</sup> · Rumaan Bashir<sup>1</sup>

Received: 18 July 2022 / Revised: 22 November 2022 / Accepted: 6 April 2023 /

Published online: 28 April 2023

© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2023

## Abstract

Machine Translation (MT) being an emerging area of research has currently gained significant attention from researchers primarily because of incredible features offered by contemporary data scientific approaches such as machine learning, deep learning & transfer learning, etc. Machine Translation is the automatic conversion of a text while preserving its underlying information content form and across various languages. Various approaches having different levels of efficiency and accuracy that can be applied to perform machine translation from one language to another include Direct, Rule-based, and Corpus-based approaches. India being a diverse country in terms of its cultural heritage with a huge population, plenty of information related to different spheres of life is being exchanged between people of various regions & communities in different languages on daily basis. It is therefore imperative to come up with automated translation tools capable of translating among these languages. In this paper, a study of various languages spoken in India with special reference to those listed among scheduled languages and machine translation systems designed for them is presented. The languages with low translation status have also been highlighted. The work presented in this paper will be of great benefit to prospective researchers working in the area of machine translation. It will certainly help them in gaining a better insight about the previous research developments in this area and lead them towards the design, development & implementation of improved and effective translation systems in the future.

**Keywords** Machine translation · Indian scheduled languages · Language · Kashmiri · Corpus · Dogri

---

✉ Kaiser J. Giri  
kaiser.giri@islamicuniversity.edu.in

Nawaz Ali Lone  
lonenawazali@gmail.com

Rumaan Bashir  
rumaan.bashir@islamicuniversity.edu.in

<sup>1</sup> Department of Computer Science, Islamic University of Science & Technology, Kashmir, India