iJRASET
International Journal For Research in Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 9 **Issue:** XI **Month of publication:** November 2021

DOI: https://doi.org/10.22214/ijraset.2021.38742

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.429
Volume 9 Issue XI Nov 2021- Available at www.ijraset.com

Transliteration: A Magnetic Analysis

Dr. Rudra Prasad Mishra
Faculty in Odia, Ravenshaw University

Abstract: ***Machine transliteration is an important problem in an increasingly multilingual world as it plays a critical role in many downstream applications such as machine translation or cross-lingual information retrieval systems. There is now a vast amount of information accessible via the Internet where a lot of regional and cultural information is put on the World Wide Web in different languages and scripts. There are more that six thousand living languages in the world. Adding to the diversity is the fact that some languages are written in different scripts in different regions of the world. The multitude of foreign languages and mutually incomprehensible scripts of the same language pose a barrier to information exchange as we cannot all learn every language or script in use worldwide. Therefore, if we can get around the language barrier or at least the script barrier, we can access much more of the world's culture and can explore its abundant richness.***
Keywords: Transliteration, Translation. Cross-lingual, Multilingual, Language, Script

I. INTRODUCTION

Transliteration is the process of representing word or text written in a writing system of one language using the alphabet or writing system of another language so that the pronunciation is as close as possible to the original word or text. Machine transliteration is an automatic method for converting words in one language into phonetically equivalent words in another language. Generally, transliteration of proper nouns and technical terms is a significant and challenging problem in many multi-lingual text and speech processing applications. Machine transliteration plays an important role in natural language application such as cross language information retrieval and machine translation, especially when handling proper nouns and technical terms. Even when hand-crafted translation lexicons, used for machine translation and cross-lingual information retrieval provide significant coverage of the words encountered in the text, a considerable portion of the tokens not covered by the lexicons are proper names and technical terms . Also in cross lingual information retrieval systems, proper names and technical terms are very much important, since they carry the most distinctive information in a query. The poor translations adversely affect the performance of such systems. Automatic name transliteration plays vital role in many of the natural language processing tasks like machine translation, cross lingual information retrieval, multilingual named entity and term processing, corpus alignment, automatic bilingual dictionary compilation.

Transliteration of technical terms and proper names, such as personal, location and organization names, from one language into another language with approximate phonetic equivalents is called as named entity transliteration. The phonetic transliteration from the source language to the target language is termed as forward transliteration. The reverse process of recalling a word in source language from a list of transliterations is called as backward transliteration. Since a source language word can have more than one valid phonetically equivalent transliteration in the target language, it also becomes necessary to generate a list of possible equivalent transliterations in target language for a given source language word.

Machine transliteration is an important problem in an increasingly multilingual world as it plays a critical role in many downstream applications such as machine translation or cross-lingual information retrieval systems. There is now a vast amount of information accessible via the Internet where a lot of regional and cultural information is put on the World Wide Web in different languages and scripts. There are more than six thousand living languages in the world. Adding to the diversity is the fact that some languages are written in different scripts in different regions of the world. The multitude of foreign languages and mutually incomprehensible scripts of the same language pose a barrier to information exchange as we cannot all learn every language or script in use worldwide. Therefore, if we can get around the language barrier or at least the script barrier, we can access much more of the world's culture and can explore its abundant richness.

Language transliteration is also one of the important areas in natural language processing. Transliteration is the process of mapping a written word from a language-script pair to another language-script pair . Machine Transliteration is the computer based conversion of a character or word from one language/script to another without losing its phonological characteristics. Machine transliteration emerged as a part of machine translation to deal with proper names and technical terms that are translated with preserved pronunciation. With the advent of new technology and the open access to huge information through the Web, it has become increasingly common to adopt foreign words into one’s language.

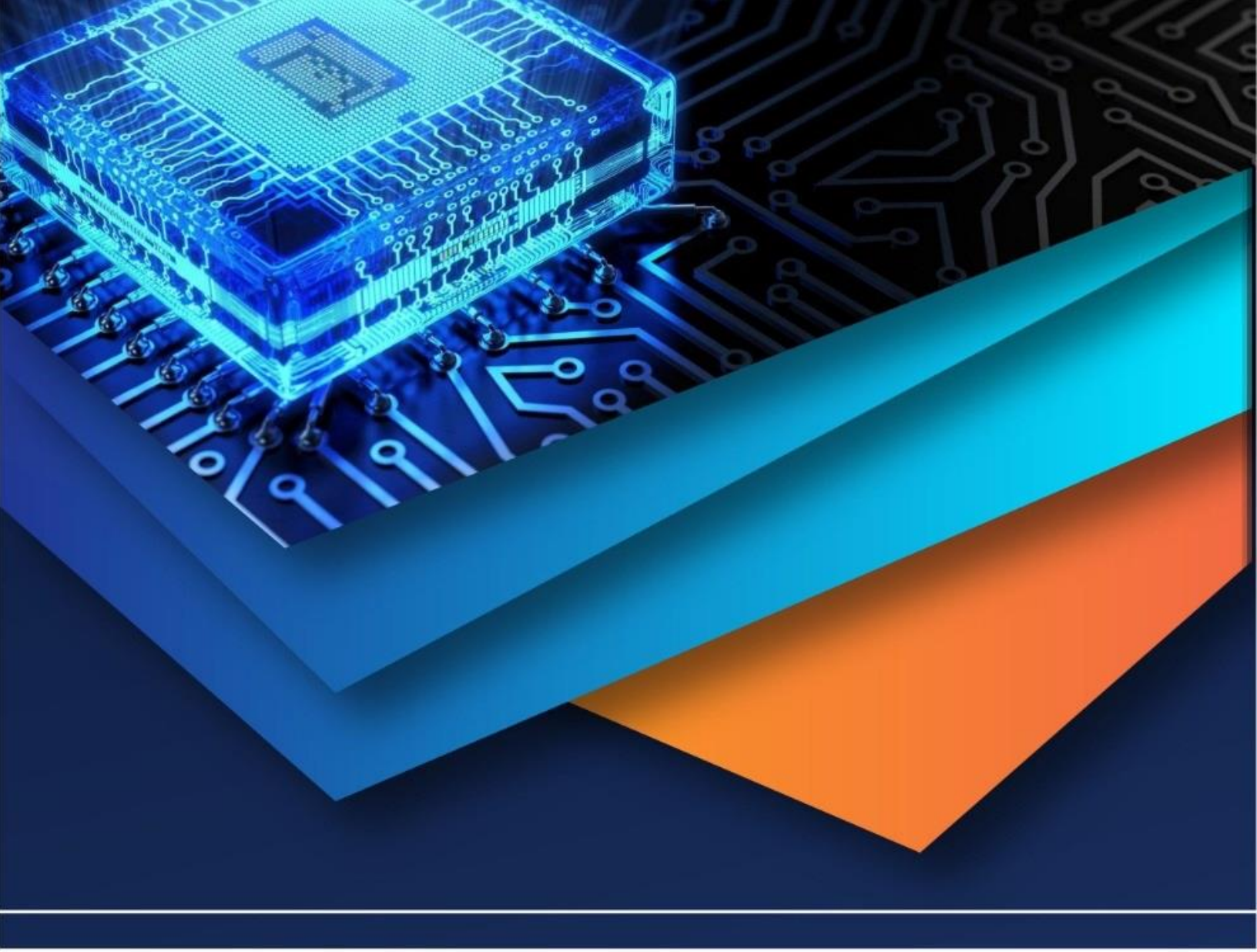
©IJRASET: All Rights are Reserved

This phonetic "translation" of foreign words is called transliteration. In other words, Machine transliteration is the process of automatically transforming the script of a word from a source language to a target language, while preserving pronunciation. Transliteration is particularly used to translate proper names and technical terms from languages using Roman alphabet into ones using non-Roman alphabet such as from English to Korean, Japanese, Chinese or Indian languages. Machine transliteration is complicated when the language pairs have different alphabet and sound inventories. Machine transliteration is normally used to support machine translation (MT) and cross-language information retrieval (CLIR).For example, machine transliteration can assist query translation in CLIR, where proper names and technical terms frequently appear in source language queries. In the area of machine translation, transliteration helps preventing translation errors when translations of proper names and technical terms are not registered in the translation dictionary. Multilingual chat applications, talking translators, and real-time translation of emails and websites are some examples of the modern commercial applications of machine transliteration and translation. Machine transliteration can, therefore, improve the performance of MT and CLIR.

REFERENCES

[1] Daniel Jurafsky, James H Martin (2000). Speech and Language Processing – An Introduction to Natural Language Processing, Computational Linguistics and Speech Recognition. Prentice Hall. Page 799.
[2] Helen M. Meng, Wai-Kit Lo, Berlin Chen, Karen Tang (2001). Generate Phonetic Cognates to Handle Name Entities in English-Chinese crosslanguage spoken document retrieval. In proceedings of ASRU 2001, Italy.
[3] Jong-Hoon Oh, Key-Sun Choi, Hitoshi Isahara (2006). A Comparison of Different Machine Transliteration Models. Journal of Artificial Intelligence Research, California.
[4] Li Haizhou, Zhang Min, Su Jian (2004). A Joint Source-Channel Model for Machine Transliteration. In proceedings of the 42nd Annual Meeting of Association of Computational Linguistics, Barcelona.
[5] Zhang Min, Li Haizhou, Su Jian (2004) Direct Orthographical Mapping for Machine Transliteration. In proceedings of the 20th International conference on Computational Linguistics, Geneva.

©IJRASET: All Rights are Reserved



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)