

Report of NEWS 2015 Machine Transliteration Shared Task

Rafael E. Banchs¹, Min Zhang², Xiangyu Duan², Haizhou Li¹, A. Kumaran³

¹ Institute for Infocomm Research, A*STAR, Singapore 138632
{rembanchs,hli}@i2r.a-star.edu.sg

² Soochow University, China 215006
{minzhang,xiangyuduan}@suda.edu.cn

³ Multilingual Systems Research, Microsoft Research India
a.kumaran@microsoft.com

Abstract

This report presents the results from the Machine Transliteration Shared Task conducted as part of The Fifth Named Entities Workshop (NEWS 2015) held at ACL 2015 in Beijing, China. Similar to previous editions of NEWS Workshop, the Shared Task featured machine transliteration of proper names over 14 different language pairs, including 12 different languages and two different Japanese scripts. A total of 6 teams participated in the evaluation, submitting 194 standard and 12 non-standard runs, involving a diverse variety of transliteration methodologies. Four performance metrics were used to report the evaluation results. Once again, the NEWS shared task on machine transliteration has successfully achieved its objectives by providing a common ground for the research community to conduct comparative evaluations of state-of-the-art technologies that will benefit the future research and development in this area.

1 Introduction

Names play an important role in the performance of most Natural Language Processing (NLP) and Information Retrieval (IR) applications. They are also critical in cross-lingual applications such as Machine Translation (MT) and Cross-language Information Retrieval (CLIR), as it has been shown that system performance correlates positively with the quality of name conversion across languages (Demner-Fushman and Oard 2002, Mandl and Womser-Hacker 2005, Hermjakob et al. 2008, Udupa et al. 2009). Bilingual dictionaries constitute the traditional source of information for name conversion across languages, however they offer very limited support due to the fact that, in most languages, names are continuously emerging and evolving.

All of the above points to the critical need for robust Machine Transliteration methods and systems. During the last decade, significant efforts has been conducted by the research community to address the problem of machine transliteration (Knight and Graehl 1998, Meng et al. 2001, Li et al. 2004, Zelenko and Aone 2006, Sproat et al. 2006, Sherif and Kondrak 2007, Hermjakob et al. 2008, Al-Onaizan and Knight 2002, Goldwasser and Roth 2008, Goldberg and Elhadad 2008, Klementiev and Roth 2006, Oh and Choi 2002, Virga and Khudanpur 2003, Wan and Verspoor 1998, Kang and Choi 2000, Gao et al. 2004, Li et al. 2009a, Li et al. 2009b). These previous works fall into three main categories: grapheme-based, phoneme-based and hybrid methods. Grapheme based methods (Li et al. 2004) treat transliteration as a direct orthographic mapping and only uses orthography-related features while phoneme-based methods (Knight and Graehl 1998) make use of phonetic correspondences to generate the transliteration. The hybrid approach refers to the combination of several different models or knowledge sources to support the transliteration generation process.

The first machine transliteration shared task (Li et al. 2009b, Li et al. 2009a) was organized and conducted as part of NEWS 2009 at ACL-IJCNLP 2009. It was the first time that common benchmarking data in diverse language pairs was provided for evaluating state-of-the-art machine transliteration. While the focus of the 2009 shared task was on establishing the quality metrics and on setting up a baseline for transliteration quality based on those metrics, the 2010 shared task (Li et al. 2010a, Li et al. 2010b) focused on expanding the scope of the transliteration generation task to about a dozen languages and on exploring the quality of the task depending on the direction of transliteration. In NEWS 2011 (Zhang et al. 2011a, Zhang et al. 2011b),

the focus was on significantly increasing the hand-crafted parallel corpora of named entities to include 14 different language pairs from 11 language families, and on making them available as the common dataset for the shared task. The NEWS 2015 Shared Task on Transliteration has been a continued effort for evaluating machine transliteration performance over such a common dataset following the NEWS 2012 (Zhang et al. 2012) and 2011 shared tasks.

In this paper, we present in full detail the results of the NEWS 2015 Machine Transliteration Shared Task. The rest of the paper is structured as follows. Section 2 provides a short review of the main characteristics of the machine transliteration task and the corpora used for it. Section 3 reviews the four metrics used for the evaluations. Section 4 reports specific details about participation in the 2015 edition of the shared task, and section 5 presents and discusses the evaluation results. Finally, section 6 presents our main conclusions and future plans.

2 Shared Task on Transliteration

Transliteration, sometimes also called Romanization, especially if Latin Scripts are used for target strings (Halpern 2007), deals with the conversion of names between two languages and/or script systems. Within the context of the Transliteration Shared Task, we are aiming not only at addressing the name conversion process but also its practical utility for downstream applications, such as MT and CLIR.

In this sense, we adopt the same definition of transliteration as proposed during the NEWS 2009 workshop (Li et al. 2009a). According to it, transliteration is understood as the “conversion of a given name in the source language (a text string in the source writing system or orthography) to a name in the target language (another text string in the target writing system or orthography” conditioned to the following specific requirements regarding the name representation in the target language:

- it is phonetically equivalent to the source name,
- it conforms to the phonology of the target language, and
- it matches the user intuition on its equivalence with respect to the source language name.

Following NEWS 2011 and NEWS 2012, the three back-transliteration tasks are maintained. Back-transliteration attempts to restore translit-

erated names back into their original source language. For instance, the tasks for converting western names written in Chinese and Thai back into their original English spellings are considered. Similarly, a task for back-transliterating Romanized Japanese names into their original Kanji strings is considered too.

2.1 Shared Task Description

Following the tradition of NEWS workshop series, the shared task in NEWS 2015 consists of developing machine transliteration systems in one or more of the specified language pairs. Each language pair of the shared task consists of a source and a target language, implicitly specifying the transliteration direction. Training and development data in each of the language pairs was made available to all registered participants for developing their transliteration systems.

At the evaluation time, a standard hand-crafted test set consisting of between 500 and 3,000 source names (approximately 5-10% of the training data size) was released, on which the participants were required to produce a ranked list of transliteration candidates in the target language for each source name. The system output is tested against a reference set (which may include multiple correct transliterations for some source names), and the performance of a system is captured in multiple metrics (defined in Section 3), each designed to capture a specific performance dimension.

For every language pair, each participant was required to submit at least one run (designated as a “standard” run) that uses only the data provided by the NEWS workshop organizers in that language pair; i.e. no other data or linguistic resources are allowed for standard runs. This ensures parity between systems and enables meaningful comparison of performance of various algorithmic approaches in a given language pair. Participants were allowed to submit one or more standard runs for each task they participated in. If more than one standard runs were submitted, it was required to name one of them as a “primary” run, which was the one used to compare results across different systems.

In addition, more than one “non-standard” runs could be submitted for every language pair using either data beyond the one provided by the shared task organizers, any other available linguistic resources in a specific language pair, or both. This essentially enabled participants to demonstrate the limits of performance of their systems in a given language pair.

2.2 Shared Task Corpora

Two specific constraints were considered when selecting languages for the shared task: language diversity and data availability. To make the shared task interesting and to attract wider participation, it is important to ensure a reasonable variety among the languages in terms of linguistic diversity, orthography and geography. Clearly, the ability of procuring and distributing a reasonably large (approximately 10K paired names for training and testing together) hand-crafted corpora consisting primarily of paired names is critical for this process. Following NEWS 2011, the 14 tasks shown in Tables 1.a-e were used (Li et al. 2004, Kumaran and Kellner 2007, MSRI 2009, CJKI 2010). Additionally, the test sets from NEWS 2012 (each of size 1K) were also used for evaluation purposes in this shared task.

The names given in the training sets for Chinese, Japanese, Korean, Thai, Persian and Hebrew languages are Western names and their respective transliterations; the Japanese Name (in English) → Japanese Kanji data set consists only of native Japanese names; the Arabic data set consists only of native Arabic names. The Indic data set (Hindi, Tamil, Kannada, Bangla) consists of a mix of Indian and Western names.

For all of the tasks chosen, we have been able to procure paired-name data between the source and the target scripts and were able to make them available to the participants. For some language pairs, such as the case of English-Chinese and English-Thai, there are both transliteration and back-transliteration tasks. Most of the tasks are just one-way transliteration, although Indian data sets contains a mixture of names from both Indian and Western origins.

3 Evaluation Metrics and Rationale

The participants have been asked to submit standard and, optionally, non-standard runs. One of the standard runs must be named as the primary submission, which was the one used for the performance summary. Each run must contain a ranked list of up to ten candidate transliterations for each source name. The submitted results are compared to the ground truth (reference transliterations) using four evaluation metrics capturing different aspects of transliteration performance. The four considered evaluation metrics are:

- Word Accuracy in Top-1 (ACC),
- Fuzziness in Top-1 (Mean F-score),
- Mean Reciprocal Rank (MRR), and
- Mean Average Precision (MAP_{ref}).

Task ID: EnCh			<i>data size</i>		
<i>Origin</i>	<i>Source</i>	<i>Target</i>	<i>Train</i>	<i>Dev</i>	<i>Test</i>
Western	English	Chinese	37K	2.8K	2.0K

Task ID: ChEn			<i>data size</i>		
<i>Origin</i>	<i>Source</i>	<i>Target</i>	<i>Train</i>	<i>Dev</i>	<i>Test</i>
Western	Chinese	English	28K	2.7K	2.7K

Table 1.a: Datasets provided by Institute for Infocomm Research, Singapore.

Task ID: EnKo			<i>data size</i>		
<i>Origin</i>	<i>Source</i>	<i>Target</i>	<i>Train</i>	<i>Dev</i>	<i>Test</i>
Western	English	Korean	7.0K	1.0K	0.6K

Task ID: EnJa			<i>data size</i>		
<i>Origin</i>	<i>Source</i>	<i>Target</i>	<i>Train</i>	<i>Dev</i>	<i>Test</i>
Western	English	Katakana	26K	2.0K	1.8K

Task ID: JnJk			<i>data size</i>		
<i>Origin</i>	<i>Source</i>	<i>Target</i>	<i>Train</i>	<i>Dev</i>	<i>Test</i>
Japanese	English	Kanji	10K	2.0K	0.6K

Task ID: ArEn			<i>data size</i>		
<i>Origin</i>	<i>Source</i>	<i>Target</i>	<i>Train</i>	<i>Dev</i>	<i>Test</i>
Arabic	Arabic	English	27K	2.5K	2.6K

Table 1.b: Datasets provided by the CJK Institute, Japan.

Task ID: EnHi			<i>data size</i>		
<i>Origin</i>	<i>Source</i>	<i>Target</i>	<i>Train</i>	<i>Dev</i>	<i>Test</i>
Mixed	English	Hindi	12K	1.0K	1.0K

Task ID: EnTa			<i>data size</i>		
<i>Origin</i>	<i>Source</i>	<i>Target</i>	<i>Train</i>	<i>Dev</i>	<i>Test</i>
Mixed	English	Tamil	10K	1.0K	1.0K

Task ID: EnKa			<i>data size</i>		
<i>Origin</i>	<i>Source</i>	<i>Target</i>	<i>Train</i>	<i>Dev</i>	<i>Test</i>
Mixed	English	Kannada	10K	1.0K	1.0K

Task ID: EnBa			<i>data size</i>		
<i>Origin</i>	<i>Source</i>	<i>Target</i>	<i>Train</i>	<i>Dev</i>	<i>Test</i>
Mixed	English	Bangla	13K	1.0K	1.0K

Task ID: EnHe			<i>data size</i>		
<i>Origin</i>	<i>Source</i>	<i>Target</i>	<i>Train</i>	<i>Dev</i>	<i>Test</i>
Western	English	Hebrew	9.5K	1.0K	1.0K

Table 1.c: Datasets provided by Microsoft Research India.

Task ID: EnTh			<i>data size</i>		
<i>Origin</i>	<i>Source</i>	<i>Target</i>	<i>Train</i>	<i>Dev</i>	<i>Test</i>
Western	English	Thai	27K	2.0K	2.0K

Task ID: ThEn			<i>data size</i>		
<i>Origin</i>	<i>Source</i>	<i>Target</i>	<i>Train</i>	<i>Dev</i>	<i>Test</i>
Western	Thai	English	25K	2.0K	1.9K

Table 1.d: Datasets provided by National Electronics and Computer Technology Center.

Task ID: EnPe			<i>data size</i>		
<i>Origin</i>	<i>Source</i>	<i>Target</i>	<i>Train</i>	<i>Dev</i>	<i>Test</i>
Western	English	Persian	10K	2.0K	2.0K

Table 1.e: Dataset provided by Sarvnaz Karimi / RMIT.

In the next subsections, we present a brief description of the four considered evaluation metrics. The following notation is further assumed:

- N : Total number of names (source words) in the test set,
- n_i : Number of reference transliterations for i -th name in the test set ($n_i \geq 1$),
- $r_{i,j}$: j -th reference transliteration for i -th name in the test set,
- $c_{i,k}$: k -th candidate transliteration (system output) for i -th name in the test set ($1 \leq k \leq 10$),
- K_i : Number of candidate transliterations produced by a transliteration system.

3.1 Word Accuracy in Top-1 (ACC)

Also known as Word Error Rate, it measures correctness of the first transliteration candidate in the candidate list produced by a transliteration system. $ACC = 1$ means that all top candidates are correct transliterations; i.e. they match one of the references, and $ACC = 0$ means that none of the top candidates are correct.

$$ACC = \frac{1}{N} \sum_{i=1}^N \left\{ \begin{array}{l} 1 \text{ if } \exists r_{i,j} : r_{i,j} = c_{i,1} \\ 0 \text{ otherwise} \end{array} \right\} \quad (\text{Eq.1})$$

3.2 Fuzziness in Top-1 (Mean F-score)

The Mean F-score measures how different, on average, the top transliteration candidate is from its closest reference. F-score for each source word is a function of Precision and Recall and equals 1 when the top candidate matches one of the references, and 0 when there are no common characters between the candidate and any of the references.

Precision and Recall are calculated based on the length of the Longest Common Subsequence (LCS) between a candidate and a reference:

$$LCS(c, r) = \frac{1}{2}(|c| + |r| - ED(c, r)) \quad (\text{Eq.2})$$

where ED is the edit distance and $|x|$ is the length of x . For example, the longest common subsequence between “abcd” and “afcde” is “acd” and its length is 3. The best matching reference, i.e. the reference for which the edit distance has the minimum, is taken for calculation. If the best matching reference is given by

$$r_{i,m} = \arg \min_j (ED(c_{i,1}, r_{i,j})) \quad (\text{Eq.3})$$

the Recall, Precision and F-score for the i -th word are calculated as:

$$R_i = \frac{LCS(c_{i,1}, r_{i,m})}{|r_{i,m}|} \quad (\text{Eq.4})$$

$$P_i = \frac{LCS(c_{i,1}, r_{i,m})}{|c_{i,1}|} \quad (\text{Eq.5})$$

$$F_i = 2 \frac{R_i \times P_i}{R_i + P_i} \quad (\text{Eq.6})$$

The lengths are computed with respect to distinct Unicode characters, and no distinctions are made for different character types of a language (e.g. vowel vs. consonant vs. combining diereses, etc.).

3.3 Mean Reciprocal Rank (MRR)

Measures traditional MRR for any right answer produced by the system, from among the candidates. $1/\text{MRR}$ tells approximately the average rank of the correct transliteration. MRR closer to 1 implies that the correct answer is mostly produced close to the top of the n -best lists.

$$RR_i = \left\{ \begin{array}{l} \min_j \frac{1}{j} \text{ if } \exists r_{i,j}, c_{i,k} : r_{i,j} = c_{i,k} \\ 0 \text{ otherwise} \end{array} \right\} \quad (\text{Eq.7})$$

$$MRR = \frac{1}{N} \sum_{i=1}^N RR_i \quad (\text{Eq.8})$$

3.4 Mean Average Precision (MAP_{ref})

This metric measures tightly the precision in the n -best candidates for i -th source name, for which reference transliterations are available. If all of the references are produced, then the MAP is 1. If we denote the number of correct candidates for the i -th source word in k -best list as $num(i, k)$, then MAP_{ref} is given by:

$$\text{MAP}_{ref} = \frac{1}{N} \sum_{i=1}^N \frac{1}{n_i} (\sum_{k=1}^{n_i} num(i, k)) \quad (\text{Eq.9})$$

4 Participation in the Shared Task

A total of six teams from six different institutions participated in the NEWS 2015 Shared Task. More specifically, the participating teams were from University of Alberta (UALB), Uppsala University (UPPS), Beijing Jiaotong University (BJTU), the National Institute of Information and Communications Technology (NICT), the Indian Institute of Technology Bombay (IITB) and the National Taiwan University (NTU).

Teams were required to submit at least one standard run for every task they participated in, and for both, NEWS 2011 and NEWS 2012, test sets. The former was used as a progress evaluation set and the latter as the official NEWS 2015 evaluation set. In total, we received 97 standard and 6 non-standard runs for each test set; i.e. 194 standard and 12 non-standard runs in total. Table 2 summarizes the number of standard runs, non-standard runs and teams participating per task.

Task	Std	Non	Teams Participating
EnCh	26	2	UALB, UPPS, BJTU, NICT, IITB, NTU
ChEn	20	2	UPPS, BJTU, NICT, IITB
EnKo	18	0	NICT, NTU
EnJa	6	0	UALB, NICT
JnJk	4	0	NICT
ArEn	6	0	UALB, NICT
EnHi	20	2	UALB, NICT, IITB
EnTa	20	2	UALB, NICT, IITB
EnKa	12	2	UALB, NICT, IITB
EnBa	18	0	UALB, NICT, IITB
EnHe	12	2	UALB, NICT, IITB
EnTh	10	0	UALB, NICT, IITB
ThEn	10	0	UALB, NICT, IITB
EnPe	12	0	UALB, NICT, IITB
	194	12	

Table 2: Number of standard (Std) and non-standard (Non) runs submitted, and teams participating in each task.

As seen from the table, the most popular task continues to be the transliteration from English to Chinese (Zhang et al. 2012), followed by Chinese to English, English to Hindi, and English to Tamil. Non-standard runs were only submitted for 6 of the 14 tasks.

5 Task Results and Analysis

Figure 1 summarizes the results of the NEWS 2015 Shared Task. In the figure, only F-scores over the NEWS 2012 evaluation test set (referred to as NEWS12/15) for all primary standard submissions are depicted. A total of 41 primary standard submissions were received.

As seen from the figure, with the exception of the English to Japanese Katakana, only transliteration tasks involving Arabic, Persian and the four considered Indian languages are consistently scored above 80%. For the rest of the languages, with the exception of Japanese Katakana and Hebrew, scores are consistently in the range from 60% to 80%. Notice also that, regardless the availability of training data, the English to Chinese transliteration task seems to be the more demanding one for state-of-the-art systems with respect to the considered metric.

Another interesting observation that can be derived from the figure, when looking to the language pairs English-Chinese and English-Thai, is that systems tend to perform slightly better for the case of back-transliteration tasks.

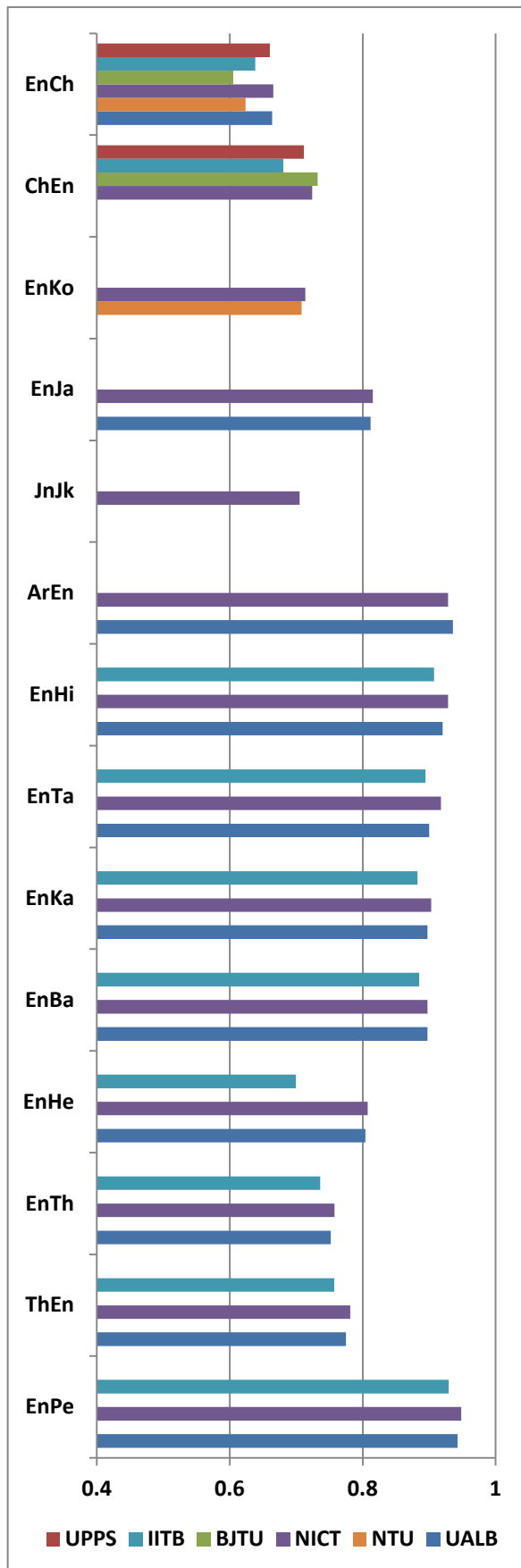


Figure 1: Mean F-scores (Top-1) on the evaluation test set (NEWS12/15) for all primary standard submissions and all transliteration tasks.

A much more comprehensive presentation of results for the NEWS 2015 Shared Task is provided in the Appendix at the end of this paper. There, resulting scores are reported for all received submissions, including standard and non-standard submissions, over both the progress test (NEWS11) and evaluation test (NEWS12/15), and the four considered evaluation metrics. All results are presented in 28 tables, each of which reports the scores for one transliteration task over one test set. In the tables, all primary standard runs are highlighted in bold-italic fonts.

Regarding the systems participating in this year evaluation, the UALB’s system (Nicolai et al. 2015) was based on multiple system combinations. They presented experimental results involving three different well-known transliteration approaches: DirecTL+ (Jiampoamarn et al. 2009), Sequitur (Bisani and Ney 2008) and SMT (Koehn et al. 2007). They showed error reductions of up to 20% over a baseline system by using system combination.

The UPPS’s system (Shao et al. 2015) implemented a phrase-based transliteration approach, which is enhanced with refined alignments produced by the M2M-aligner (Jiampoamarn et al. 2007). They also implemented a ranking mechanism based on a linear regression, showing a significant improvement on both EnCh and ChEn transliteration tasks.

The BJTU’s system (Wang et al. 2015a) implemented an SMT (Koehn et al. 2007) log linear model combination for transliteration, including standard SMT features such as a language model scores and forward and reverse phrase translation probabilities, as well as other specific transliteration features such as length of names and length of name penalties.

The NICT’s system (Finch et al. 2015) builds upon their previous SMT-based system used for NEWS 2012 (Finch et al. 2012). In this shared task, the previous system rescoring step is augmented with a neural network based transliteration model (Bahdanau et al. 2014). They showed significant improvements in 8 of the 14 transliteration tasks with respect to their 2012 system.

The ITTB’s system (Kunchukuttan and Bhattacharyya 2015) also followed the SMT approach to transliteration. In this case they include two specific preprocessing enhancements: the addition of word-boundary markers, and a language-independent overlapping character segmentation. They observed that word-boundary markers substantially improved transliteration accuracy, and overlapping segmentation showed some potential.

The NTU’s system (Wang et al. 2015b) is based on DirecTL+ with alignments generated by the M2M-aligner (Jiampoamarn et al. 2010). In preprocessing, they experimented with different grapheme segmentation methods for English, Chinese and Korean; while in post-processing, they evaluated two re-ranking approaches: orthography similarity ranking and web-based ranking.

As seen from the previous system descriptions, phrase-based SMT approaches are still predominant in the state-of-the-art for machine transliteration. Significant improvements are achieved by incorporating novel approaches in the preprocessing and post-processing stages, as well as by system combinations. Regarding pre-processing, the main focus was on segmentation, while in post-processing, using neural networks for rescoring provided the most significant gains.

Finally, figure 2 compares, in terms of Mean F-scores, the best primary standard submissions in NEWS 2012 with the ones in NEWS 2015.

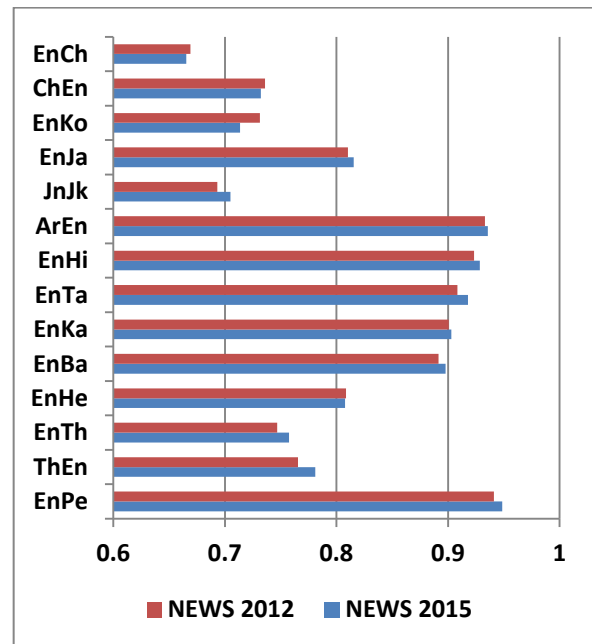


Figure 2: Mean F-scores (Top-1) on the evaluation test set (NEWS12/15) for the best primary standard submissions in 2012 and 2015.

As seen from the figure, in most of the considered transliteration tasks, some incremental improvements can be observed between the 2012 and 2015 shared tasks. The most significant improvements are in those tasks involving Japanese Katakana, Tamil, Bangla (Bengali) and Thai.

Regarding the observed drops in performance, only the one for the English to Korean Hangul task is significant. It is mainly due to the fact that the best performing system for this task in 2012 did not participate in the 2015 shared task.

6 Conclusions

The Shared Task on Machine Transliteration in NEWS 2015 has shown, once again, that the research community has a continued interest in this area. This report summarizes the results of the NEWS 2015 Shared Task.

We are pleased to report a comprehensive set of machine transliteration approaches and their evaluation results over two test sets: progress test (NEWS11) and evaluation test (NEWS12/15), as well as two conditions: standard runs and non-standard runs. While the standard runs allow for conducting meaningful comparisons across different algorithms, the non-standard runs open up more opportunities for exploiting a variety of additional linguistic resources.

Six teams from six different institutions participated in the shared task. In total, we received 97 standard and 6 non-standard runs for each test set; i.e. 194 standard and 12 non-standard runs in total. Most of the current state-of-the-art in machine transliteration is represented in the systems that have participated in the shared task.

Encouraged by the continued success of the NEWS workshop series, we plan to continue this event in the future to further promoting machine transliteration research and development.

Acknowledgments

The organizers of the NEWS 2015 Shared Task would like to thank the Institute for Infocomm Research (Singapore), Microsoft Research India, CJK Institute (Japan), National Electronics and Computer Technology Center (Thailand) and Sarvnaz Karim / RMIT for providing the corpora and technical support. Without those, the Shared Task would not be possible. We also want to thank all programme committee members for their valuable comments that improved the quality of the shared task papers. Finally, we wish to thank all participants for their active participation, which have made again the NEWS Machine Transliteration Shared Task a successful one.

References

- Y. Al-Onaizan, K. Knight. 2002. Machine transliteration of names in arabic text. In Proc. ACL-2002Workshop: Computational Approaches to Semitic Languages, Philadelphia, PA, USA.
- D. Bahdanau, K. Cho, Y. Bengio. 2014. Neural machine translation by jointly learning to align and translate. Cornell University Library, arXiv:1409.0473 [cs.CL]
- M. Bisani, H. Ney. 2008. Joint sequence models for grapheme-to-phoneme conversion. *Speech Communication*, 50(5):434–451.
- CJKI. 2010. CJK Institute. <http://www.cjk.org/>.
- D. Demner-Fushman, D.W. Oard. 2002. The effect of bilingual term list size on dictionary-based cross-language information retrieval. In Proc. 36-th Hawaii Int'l. Conf. System Sciences, volume 4, page 108.2.
- A. Finch, P. Dixon, E. Sumita. 2012. Rescoring a phrase-based machine transliteration system with recurrent neural network language models. In Proceedings of the 4th Named Entity Workshop (NEWS) 2012, pages 47–51, Jeju, Korea, July.
- A. Finch, L. Liu, X. Wang, E. Sumita. 2015. Neural Network Transduction Models in Transliteration Generation. In Proceedings of the 2015 Named Entities Workshop: Shared Task on Transliteration (NEWS 2015), Beijing, China.
- W. Gao, K.F. Wong, W. Lam. 2004. Phoneme-based transliteration of foreign names for OOV problem. In Proc. IJCNLP, pages 374–381, Sanya, Hainan, China.
- Y. Goldberg, M. Elhadad. 2008. Identification of transliterated foreign words in Hebrew script. In Proc. CICLing, volume LNCS 4919, pages 466–477.
- D. Goldwasser, D. Roth. 2008. Transliteration as constrained optimization. In Proc. EMNLP, pages 353–362.
- J. Halpern. 2007. The challenges and pitfalls of Arabic romanization and arabization. In Proc. Workshop on Comp. Approaches to Arabic Scriptbased Lang.
- U. Hermjakob, K. Knight, H. Daum. 2008. Name translation in statistical machine translation: Learning when to transliterate. In Proc. ACL, Columbus, OH, USA, June.
- S. Jiampojarn, G. Kondrak, T. Sherif. 2007. Applying many-to-many alignments and hidden markov models to letter-to-phoneme conversion. In proceedings of Human Language Technologies 2007: The Conference of the North American Chapter of the Association for Computational Linguistics; pages 372–379, Rochester, New York, April.
- S. Jiampojarn, A. Bhargava, Q. Dou, K. Dwyer, G. Kondrak. 2009. DirecTL: a language independent approach to transliteration. In Proceedings of the 2009 Named Entities Workshop: Shared Task on Transliteration (NEWS 2009), pages 28–31, Suntec, Singapore.
- S. Jiampojarn, C. Cherry, G. Kondrak. 2010. Integrating joint n-gram features into a discriminative training framework. In Proceedings of NAACL-2010, Los Angeles, CA, June. Association for Computational Linguistics.
- B.J. Kang, K.S. Choi. 2000. English-Korean automatic transliteration/ backtransliteration system and character alignment. In Proc. ACL, pages 17–18, Hong Kong.
- A. Klementiev, D. Roth. 2006. Weakly supervised named entity transliteration and discovery from multilingual comparable corpora. In Proc. 21st Int'l Conf Computational Linguistics and 44th Annual Meeting of ACL, pages 817–824, Sydney, Australia, July.

- K. Knight, J. Graehl. 1998. Machine transliteration. *Computational Linguistics*, 24(4).
- P. Koehn, H. Hoang, A. Birch, C. Callison-Burch, M. Federico, N. Bertoldi, B. Cowan, W. Shen, C. Moran, R. Zens, C. Dyer, O. Bojar, A. Constantin, E. Herbst. 2007. Moses: Open source toolkit for statistical machine translation. In *Proceedings of the 45th Annual Meeting of the Association for Computational Linguistics Companion Volume Proceedings of the Demo and Poster Sessions*, pages 177–180, Prague, Czech Republic.
- A. Kumaran, T. Kellner. 2007. A generic framework for machine transliteration. In *Proc. SIGIR*, pages 721–722.
- A. Kunchukuttan, P. Bhattacharyya. 2015. Data representation methods and use of mined corpora for Indian language transliteration. In *Proceedings of the 2015 Named Entities Workshop: Shared Task on Transliteration (NEWS 2015)*, Beijing, China.
- H. Li, M. Zhang, J. Su. 2004. A joint source-channel model for machine transliteration. In *Proc. 42nd ACL Annual Meeting*, pages 159–166, Barcelona, Spain.
- H. Li, A. Kumaran, V. Pervouchine, M. Zhang. 2009a. Report of NEWS 2009 machine transliteration shared task. In *Proc. Named Entities Workshop at ACL 2009*.
- H. Li, A. Kumaran, M. Zhang, V. Pervouchine. 2009b. ACL-IJCNLP 2009 Named Entities Workshop - Shared Task on Transliteration. In *Proc. Named Entities Workshop at ACL 2009*.
- H. Li, A. Kumaran, M. Zhang, V. Pervouchine. 2010a. Report of news 2010 transliteration generation shared task. In *Proc. Named Entities Workshop at ACL 2010*.
- H. Li, A. Kumaran, M. Zhang, V. Pervouchine. 2010b. Whitepaper of news 2010 shared task on transliteration generation. In *Proc. Named Entities Workshop at ACL 2010*.
- T. Mandl, C. Womser-Hacker. 2005. The effect of named entities on effectiveness in cross-language information retrieval evaluation. In *Proc. ACM Symp. Applied Comp.*, pages 1059–1064.
- H.M. Meng, W.K. Lo, B. Chen, K. Tang. 2001. Generate phonetic cognates to handle name entities in English-Chinese cross-language spoken document retrieval. In *Proc. ASRU*.
- MSRI. 2009. Microsoft Research India. <http://research.microsoft.com/india>.
- G. Nicolai, B. Hauer, M. Salameh, A. St Arnaud, Y. Xu, L. Yao, G. Kondrak. 2015. Multiple System Combination for Transliteration. In *Proceedings of the 2015 Named Entities Workshop: Shared Task on Transliteration (NEWS 2015)*, Beijing, China.
- J.H. Oh, K.S. Choi. 2002. An English-Korean transliteration model using pronunciation and contextual rules. In *Proc. COLING 2002*, Taipei, Taiwan.
- Y. Shao, J. Tiedemann, J. Nivre. 2015. Boosting English-Chinese Machine Transliteration via High Quality Alignment and Multilingual Resources. In *Proceedings of the 2015 Named Entities Workshop: Shared Task on Transliteration (NEWS 2015)*, Beijing, China.
- T. Sherif, G. Kondrak. 2007. Substringbased transliteration. In *Proc. 45th Annual Meeting of the ACL*, pages 944–951, Prague, Czech Republic, June.
- R. Sproat, T. Tao, C.X. Zhai. 2006. Named entity transliteration with comparable corpora. In *Proc. 21st Int'l Conf Computational Linguistics and 44th Annual Meeting of ACL*, pages 73–80, Sydney, Australia.
- R. Udupa, K. Saravanan, A. Bakalov, A. Bhole. 2009. “They are out there, if you know where to look”: Mining transliterations of OOV query terms for cross-language information retrieval. In *LNCS: Advances in Information Retrieval*, volume 5478, pages 437–448. Springer Berlin / Heidelberg.
- P. Virga, S. Khudanpur. 2003. Transliteration of proper names in cross-lingual information retrieval. In *Proc. ACL MLNER*, Sapporo, Japan.
- S. Wan, C.M. Verspoor. 1998. Automatic English-Chinese name transliteration for development of multilingual resources. In *Proc. COLING*, pages 1352–1356.
- D. Wang, X. Yang, J. Xu, Y. Chen, N. Wang, B. Liu, J. Yang, Y. Zhang. 2015a. A Hybrid Transliteration Model for Chinese/English Named Entities — BJTU-NLP Report for the 5th Named Entities Workshop. In *Proceedings of the 2015 Named Entities Workshop: Shared Task on Transliteration (NEWS 2015)*, Beijing, China.
- Y.C. Wang, C.K. Wu, R.T.H. Tsai. 2015b. NCU IISR English-Korean and English-Chinese Named Entity Transliteration Using Different Grapheme Segmentation Approaches. In *Proceedings of the 2015 Named Entities Workshop: Shared Task on Transliteration (NEWS 2015)*, Beijing, China.
- D. Zelenko, C. Aone. 2006. Discriminative methods for transliteration. In *Proc. EMNLP*, pages 612–617, Sydney, Australia, July.
- M. Zhang, A. Kumaran, H. Li. 2011a. Whitepaper of news 2011 shared task on machine transliteration. In *Proc. Named Entities Workshop at IJCNLP 2011*.
- M. Zhang, H. Li, A. Kumaran, M. Liu. 2011b. Report of news 2011 machine transliteration shared task. In *Proc. Named Entities Workshop at IJCNLP 2011*.
- M. Zhang, H. Li, A. Kumaran, M. Liu. 2012. Report of NEWS 2012 Machine Transliteration Shared Task. *Proceedings of the 50th Annual Meeting of the Association for Computational Linguistics*, pages 10–20, Jeju, Republic of Korea.

Appendix: Evaluation Results

Team	Submission	Test Set	ACC	F-score	MRR	MAP _{ref}
UPPS	Run-1	NEWS11	0.333000	0.673433	0.387373	0.320348
UPPS	Run-2	NEWS11	0.324000	0.681840	0.403869	0.311746
UPPS	Run-3	NEWS11	0.339500	0.694297	0.397274	0.326723
UPPS	Run-4	NEWS11	0.365000	0.708208	0.430950	0.351070
UPPS	Run-5	NEWS11	0.721500	0.869726	0.775266	0.717143
BJTU	Run-1	NEWS11	0.223500	0.628967	0.223500	0.212291
BJTU	Non-standard	NEWS11	0.224500	0.619581	0.224500	0.212253
NICT	Run-1	NEWS11	0.344500	0.694670	0.448921	0.335733
NICT	Run-2	NEWS11	0.213500	0.633107	0.250693	0.206071
IITB	Run-1	NEWS11	0.303000	0.671143	0.391680	0.292121
IITB	Run-2	NEWS11	0.177000	0.561347	0.212209	0.175762
NTU	Run-1	NEWS11	0.232500	0.630283	0.232500	0.219895
NTU	Run-2	NEWS11	0.292500	0.671896	0.292500	0.277193
UALB	Run-1	NEWS11	0.331500	0.687510	0.413785	0.321167

Table A1: Results for the English to Chinese transliteration task (EnCh) on Progress Test.

Team	Submission	Test Set	ACC	F-score	MRR	MAP _{ref}
UPPS	Run-1	NEWS12/15	0.325397	0.660252	0.384383	0.313092
UPPS	Run-2	NEWS12/15	0.310516	0.659662	0.396441	0.302924
UPPS	Run-3	NEWS12/15	0.335317	0.675918	0.396312	0.323261
UPPS	Run-4	NEWS12/15	0.373016	0.693169	0.436131	0.362703
UPPS	Run-5	NEWS12/15	0.655754	0.824023	0.735236	0.649278
BJTU	Run-1	NEWS12/15	0.193452	0.605230	0.193452	0.182230
BJTU	Non-standard	NEWS12/15	0.204365	0.604767	0.204365	0.195381
NICT	Run-1	NEWS12/15	0.312500	0.665305	0.432201	0.305466
NICT	Run-2	NEWS12/15	0.220238	0.627412	0.279823	0.216849
IITB	Run-1	NEWS12/15	0.280754	0.638436	0.371490	0.273775
IITB	Run-2	NEWS12/15	0.182540	0.545881	0.219496	0.180018
NTU	Run-1	NEWS12/15	0.235119	0.623692	0.235119	0.224172
NTU	Run-2	NEWS12/15	0.279762	0.645468	0.279762	0.265198
UALB	Run-1	NEWS12/15	0.314484	0.663729	0.406026	0.305790

Table A2: Results for the English to Chinese transliteration task (EnCh) on Evaluation Test.

Team	Submission	Test Set	ACC	F-score	MRR	MAP _{ref}
UPPS	Run-1	NEWS11	0.150044	0.754860	0.228409	0.149603
UPPS	Run-2	NEWS11	0.108561	0.731156	0.182593	0.108561
UPPS	Run-3	NEWS11	0.153133	0.768302	0.233491	0.152692
UPPS	Run-4	NEWS11	0.164607	0.772975	0.251579	0.164056
UPPS	Run-5	NEWS11	0.354369	0.833290	0.427793	0.353707
BJTU	Run-1	NEWS11	0.105031	0.746174	0.105031	0.104700
BJTU	Non-standard	NEWS11	0.151368	0.765544	0.151368	0.150927
NICT	Run-1	NEWS11	0.158429	0.769593	0.252760	0.158760
NICT	Run-2	NEWS11	0.115181	0.747071	0.176935	0.115512
IITB	Run-1	NEWS11	0.105914	0.727373	0.176256	0.105914
IITB	Run-2	NEWS11	0.048102	0.616609	0.083495	0.048102

Table A3: Results for the Chinese to English transliteration task (ChEn) on Progress Test.

Team	Submission	Test Set	ACC	F-score	MRR	MAP _{ref}
UPPS	Run-1	NEWS12/15	0.191364	0.711462	0.271377	0.187057
UPPS	Run-2	NEWS12/15	0.146222	0.712179	0.223034	0.143250
UPPS	Run-3	NEWS12/15	0.199215	0.752383	0.280989	0.194663
UPPS	Run-4	NEWS12/15	0.213935	0.745219	0.304566	0.212245
UPPS	Run-5	NEWS12/15	0.345437	0.805257	0.421142	0.345437
BJTU	Run-1	NEWS12/15	0.157017	0.732100	0.157017	0.150720
BJTU	Non-standard	NEWS12/15	0.157017	0.732100	0.157017	0.150720
NICT	Run-1	NEWS12/15	0.184495	0.723785	0.283272	0.181196
NICT	Run-2	NEWS12/15	0.136408	0.712954	0.205076	0.135509
IITB	Run-1	NEWS12/15	0.141315	0.680611	0.214933	0.140361
IITB	Run-2	NEWS12/15	0.077527	0.560009	0.107662	0.076927

Table A4: Results for the Chinese to English transliteration task (ChEn) on Evaluation Test.

Team	Submission	Test Set	ACC	F-score	MRR	MAP _{ref}
NICT	Run-1	NEWS11	0.364532	0.679477	0.459777	0.361248
NICT	Run-2	NEWS11	0.200328	0.588171	0.237128	0.196223
NTU	Run-1	NEWS11	0.318555	0.657583	0.318555	0.311166
NTU	Run-2	NEWS11	0.448276	0.725463	0.448276	0.439245
NTU	Run-3	NEWS11	0.274220	0.599985	0.274220	0.268883
NTU	Run-4	NEWS11	0.215107	0.570723	0.215107	0.209770
NTU	Run-5	NEWS11	0.042693	0.332851	0.042693	0.041461
NTU	Run-6	NEWS11	0.208539	0.527015	0.343186	0.204844
NTU	Run-7	NEWS11	0.499179	0.733029	0.539546	0.494253

Table A5: Results for the English to Korean Hangul transliteration task (EnKo) on Progress Test.

Team	Submission	Test Set	ACC	F-score	MRR	MAP _{ref}
NICT	Run-1	NEWS12/15	0.363810	0.713655	0.447686	0.363333
NICT	Run-2	NEWS12/15	0.188571	0.616223	0.231373	0.188095
NTU	Run-1	NEWS12/15	0.327619	0.707843	0.327619	0.326905
NTU	Run-2	NEWS12/15	0.445714	0.748241	0.445714	0.444762
NTU	Run-3	NEWS12/15	0.145714	0.522242	0.145714	0.145476
NTU	Run-4	NEWS12/15	0.174286	0.583525	0.174286	0.174048
NTU	Run-5	NEWS12/15	0.056190	0.375155	0.056190	0.056190
NTU	Run-6	NEWS12/15	0.195238	0.552238	0.334921	0.195000
NTU	Run-7	NEWS12/15	0.506667	0.761401	0.531746	0.505476

Table A6: Results for the English to Korean Hangul transliteration task (EnKo) on Evaluation Test.

Team	Submission	Test Set	ACC	F-score	MRR	MAP _{ref}
NICT	Run-1	NEWS11	0.412121	0.807197	0.549902	0.411983
NICT	Run-2	NEWS11	0.399449	0.796744	0.495723	0.399174
UALB	Run-1	NEWS11	0.424793	0.806944	0.551851	0.424656

Table A7: Results for the English to Japanese Katakana transliteration task (EnJa) on Progress Test.

Team	Submission	Test Set	ACC	F-score	MRR	MAP _{ref}
NICT	Run-1	NEWS12/15	0.407551	0.815306	0.528128	0.404163
NICT	Run-2	NEWS12/15	0.337851	0.784695	0.439676	0.337367
UALB	Run-1	NEWS12/15	0.396902	0.811790	0.524526	0.394966

Table A8: Results for the English to Japanese Katakana transliteration task (EnJa) on Evaluation Test.

Team	Submission	Test Set	ACC	F-score	MRR	MAP _{ref}
NICT	Run-1	NEWS11	0.451839	0.637787	0.548539	0.451839
NICT	Run-2	NEWS11	0.035026	0.320328	0.041669	0.035026

Table A9: Results for the English to Japanese Kanji transliteration task (JnJk) on Progress Test.

Team	Submission	Test Set	ACC	F-score	MRR	MAP _{ref}
NICT	Run-1	NEWS12/15	0.534247	0.704960	0.610474	0.437821
NICT	Run-2	NEWS12/15	0.031963	0.341456	0.042975	0.018189

Table A10: Results for the English to Japanese Kanji transliteration task (JnJk) on Evaluation Test.

Team	Submission	Test Set	ACC	F-score	MRR	MAP _{ref}
NICT	Run-1	NEWS11	0.527048	0.927201	0.627657	0.390565
NICT	Run-2	NEWS11	0.494204	0.891547	0.595463	0.370677
UALB	Run-1	NEWS11	0.617079	0.941314	0.718896	0.435967

Table A11: Results for the Arabic to English transliteration task (ArEn) on Progress Test.

Team	Submission	Test Set	ACC	F-score	MRR	MAP _{ref}
NICT	Run-1	NEWS12/15	0.529412	0.928260	0.655920	0.459441
NICT	Run-2	NEWS12/15	0.468858	0.914553	0.609188	0.405085
UALB	Run-1	NEWS12/15	0.596021	0.935767	0.711291	0.477259

Table A12: Results for the Arabic to English transliteration task (ArEn) on Evaluation Test.

Team	Submission	Test Set	ACC	F-score	MRR	MAP _{ref}
UALB	Run-1	NEWS11	0.488000	0.883368	0.603763	0.486000
UALB	Run-2	NEWS11	0.477000	0.881284	0.580941	0.475250
UALB	Non-standard	NEWS11	0.531000	0.901526	0.627492	0.530250
NICT	Run-1	NEWS11	0.474000	0.881670	0.583569	0.472750
NICT	Run-2	NEWS11	0.410000	0.855251	0.509001	0.409250
ITTB	Run-1	NEWS11	0.434000	0.870109	0.556714	0.432750
ITTB	Run-2	NEWS11	0.416000	0.860905	0.519782	0.413750
ITTB	Run-3	NEWS11	0.468000	0.873300	0.582212	0.465250
ITTB	Run-4	NEWS11	0.440000	0.867979	0.542456	0.439000
ITTB	Run-5	NEWS11	0.276000	0.814516	0.399723	0.275750
ITTB	Run-6	NEWS11	0.263000	0.806711	0.347494	0.263000

Table A13: Results for the English to Hindi transliteration task (EnHi) on Progress Test.

Team	Submission	Test Set	ACC	F-score	MRR	MAP _{ref}
UALB	Run-1	NEWS12/15	0.649000	0.920225	0.730004	0.642778
UALB	Run-2	NEWS12/15	0.558000	0.895520	0.657864	0.552889
UALB	Non-standard	NEWS12/15	0.559000	0.898486	0.661465	0.553750
NICT	Run-1	NEWS12/15	0.696000	0.928536	0.753320	0.690167
NICT	Run-2	NEWS12/15	0.641000	0.914722	0.702288	0.631861
ITTB	Run-1	NEWS12/15	0.603000	0.907403	0.693690	0.598472
ITTB	Run-2	NEWS12/15	0.584000	0.899290	0.671509	0.579417
ITTB	Run-3	NEWS12/15	0.621000	0.911463	0.710887	0.615083
ITTB	Run-4	NEWS12/15	0.599000	0.904661	0.686692	0.595639
ITTB	Run-5	NEWS12/15	0.303000	0.810614	0.411777	0.299222
ITTB	Run-6	NEWS12/15	0.295000	0.810716	0.382562	0.291972

Table A14: Results for the English to Hindi transliteration task (EnHi) on Evaluation Test.

Team	Submission	Test Set	ACC	F-score	MRR	MAP _{ref}
UALB	Run-1	NEWS11	0.476000	0.907893	0.597020	0.474500
UALB	Run-2	NEWS11	0.477000	0.906608	0.607133	0.476500
UALB	Non-standard	NEWS11	0.480000	0.907802	0.592867	0.479000
NICT	Run-1	NEWS11	0.474000	0.904289	0.591604	0.471750
NICT	Run-2	NEWS11	0.406000	0.879832	0.509260	0.403750
ITTB	Run-1	NEWS11	0.383000	0.875980	0.524588	0.382500
ITTB	Run-2	NEWS11	0.406000	0.887583	0.531215	0.405750
ITTB	Run-3	NEWS11	0.388000	0.875171	0.520549	0.386500
ITTB	Run-4	NEWS11	0.398000	0.883577	0.526356	0.398000
ITTB	Run-5	NEWS11	0.156000	0.774768	0.241502	0.156250
ITTB	Run-6	NEWS11	0.138000	0.770373	0.196521	0.138000

Table A15: Results for the English to Tamil transliteration task (EnTa) on Progress Test.

Team	Submission	Test Set	ACC	F-score	MRR	MAP _{ref}
UALB	Run-1	NEWS12/15	0.537000	0.900219	0.633817	0.536500
UALB	Run-2	NEWS12/15	0.585000	0.910232	0.679684	0.585000
UALB	Non-standard	NEWS12/15	0.528000	0.897556	0.620700	0.527250
NICT	Run-1	NEWS12/15	0.626000	0.917861	0.702626	0.625000
NICT	Run-2	NEWS12/15	0.584000	0.901983	0.649233	0.583500
ITTB	Run-1	NEWS12/15	0.521000	0.894533	0.630000	0.521000
ITTB	Run-2	NEWS12/15	0.542000	0.899063	0.640836	0.541750
ITTB	Run-3	NEWS12/15	0.520000	0.893332	0.626811	0.519750
ITTB	Run-4	NEWS12/15	0.543000	0.899800	0.643898	0.542750
ITTB	Run-5	NEWS12/15	0.142000	0.756809	0.228639	0.142250
ITTB	Run-6	NEWS12/15	0.139000	0.758239	0.190331	0.139250

Table A16: Results for the English to Tamil transliteration task (EnTa) on Evaluation Test.

Team	Submission	Test Set	ACC	F-score	MRR	MAP _{ref}
UALB	Run-1	NEWS11	0.434000	0.883839	0.548446	0.433000
UALB	Run-2	NEWS11	0.425000	0.885530	0.520380	0.423333
UALB	Non-standard	NEWS11	0.441000	0.893042	0.548766	0.439722
NICT	Run-1	NEWS11	0.412000	0.877273	0.526961	0.410889
NICT	Run-2	NEWS11	0.360000	0.858829	0.453091	0.358750
ITTB	Run-1	NEWS11	0.373000	0.867258	0.489983	0.372722
ITTB	Run-2	NEWS11	0.364000	0.864140	0.460513	0.362111

Table A17: Results for the English to Kannada transliteration task (EnKa) on Progress Test.

Team	Submission	Test Set	ACC	F-score	MRR	MAP _{ref}
UALB	Run-1	NEWS12/15	0.545000	0.897336	0.643426	0.543861
UALB	Run-2	NEWS12/15	0.475000	0.880163	0.582526	0.473861
UALB	Non-standard	NEWS12/15	0.491000	0.891682	0.600186	0.490361
NICT	Run-1	NEWS12/15	0.562000	0.902859	0.647315	0.561361
NICT	Run-2	NEWS12/15	0.546000	0.893315	0.611181	0.544611
ITTB	Run-1	NEWS12/15	0.498000	0.882556	0.600849	0.497611
ITTB	Run-2	NEWS12/15	0.505000	0.882445	0.590167	0.504361

Table A18: Results for the English to Kannada transliteration task (EnKa) on Evaluation Test.

Team	Submission	Test Set	ACC	F-score	MRR	MAP _{ref}
NICT	Run-1	NEWS11	0.479000	0.891883	0.592440	0.477000
NICT	Run-2	NEWS11	0.375000	0.852264	0.467951	0.373250
ITTB	Run-1	NEWS11	0.470000	0.887156	0.584836	0.467750
ITTB	Run-2	NEWS11	0.442000	0.881586	0.547812	0.441500
ITTB	Run-3	NEWS11	0.453000	0.876508	0.571826	0.452000
ITTB	Run-4	NEWS11	0.435000	0.880181	0.543165	0.434750
ITTB	Run-5	NEWS11	0.234000	0.799288	0.338209	0.233750
ITTB	Run-6	NEWS11	0.241000	0.809643	0.309816	0.240500
UALB	Run-1	NEWS11	0.509000	0.897792	0.619730	0.507500

Table A19: Results for the English to Bangla (Bengali) transliteration task (EnBa) on Progress Test.

Team	Submission	Test Set	ACC	F-score	MRR	MAP _{ref}
NICT	Run-1	NEWS12/15	0.483000	0.897317	0.590843	0.482667
NICT	Run-2	NEWS12/15	0.364000	0.847578	0.465787	0.361750
ITTB	Run-1	NEWS12/15	0.441000	0.885009	0.567487	0.439917
ITTB	Run-2	NEWS12/15	0.422000	0.876431	0.530971	0.420417
ITTB	Run-3	NEWS12/15	0.451000	0.882013	0.575119	0.449667
ITTB	Run-4	NEWS12/15	0.432000	0.875988	0.541814	0.430528
ITTB	Run-5	NEWS12/15	0.238000	0.796878	0.342817	0.238000
ITTB	Run-6	NEWS12/15	0.238000	0.806505	0.320520	0.235778
UALB	Run-1	NEWS12/15	0.492000	0.897661	0.608379	0.491028

Table A20: Results for the English to Bangla (Bengali) transliteration task (EnBa) on Evaluation Test.

Team	Submission	Test Set	ACC	F-score	MRR	MAP _{ref}
UALB	Run-1	NEWS11	0.622000	0.933084	0.725101	0.622000
UALB	Run-2	NEWS11	0.622000	0.936077	0.733577	0.622000
UALB	Non-standard	NEWS11	0.616000	0.934090	0.722406	0.616000
NICT	Run-1	NEWS11	0.609000	0.933595	0.715783	0.609000
NICT	Run-2	NEWS11	0.558000	0.918467	0.646346	0.558000
ITTB	Run-1	NEWS11	0.041000	0.739161	0.059080	0.041000
ITTB	Run-2	NEWS11	0.000000	0.008072	0.000000	0.000000

Table A21: Results for the English to Hebrew transliteration task (EnHe) on Progress Test.

Team	Submission	Test Set	ACC	F-score	MRR	MAP _{ref}
UALB	Run-1	NEWS12/15	0.173636	0.803924	0.252981	0.172273
UALB	Run-2	NEWS12/15	0.180000	0.805826	0.271303	0.179318
UALB	Non-standard	NEWS12/15	0.183636	0.805540	0.257166	0.181818
NICT	Run-1	NEWS12/15	0.179091	0.807675	0.257256	0.178636
NICT	Run-2	NEWS12/15	0.162727	0.796318	0.217959	0.160909
ITTB	Run-1	NEWS12/15	0.008182	0.699630	0.016538	0.008182
ITTB	Run-2	NEWS12/15	0.000000	0.006314	0.000000	0.000000

Table A22: Results for the English to Hebrew transliteration task (EnHe) on Evaluation Test.

Team	Submission	Test Set	ACC	F-score	MRR	MAP _{ref}
NICT	Run-1	NEWS11	0.387000	0.866853	0.488948	0.383153
NICT	Run-2	NEWS11	0.358500	0.800512	0.443323	0.354538
ITTB	Run-1	NEWS11	0.312000	0.841161	0.425847	0.310233
ITTB	Run-2	NEWS11	0.284500	0.837963	0.384735	0.281712
UALB	Run-1	NEWS11	0.410000	0.871492	0.519079	0.404424

Table A23: Results for the English to Thai transliteration task (EnTh) on Progress Test.

Team	Submission	Test Set	ACC	F-score	MRR	MAP _{ref}
NICT	Run-1	NEWS12/15	0.156958	0.757421	0.213140	0.156958
NICT	Run-2	NEWS12/15	0.131877	0.742635	0.195015	0.131877
ITTB	Run-1	NEWS12/15	0.118932	0.735916	0.185874	0.118932
ITTB	Run-2	NEWS12/15	0.102751	0.733311	0.149795	0.102751
UALB	Run-1	NEWS12/15	0.140777	0.751829	0.208695	0.140777

Table A24: Results for the English to Thai transliteration task (EnTh) on Evaluation Test.

Team	Submission	Test Set	ACC	F-score	MRR	MAP _{ref}
NICT	Run-1	NEWS11	0.276923	0.846328	0.425615	0.278711
NICT	Run-2	NEWS11	0.178462	0.807659	0.302689	0.180481
ITTB	Run-1	NEWS11	0.247692	0.830477	0.402999	0.250661
ITTB	Run-2	NEWS11	0.247692	0.833104	0.376071	0.248732
UALB	Run-1	NEWS11	0.272821	0.845536	0.432649	0.274439

Table A25: Results for the Thai to English transliteration task (ThEn) on Progress Test.

Team	Submission	Test Set	ACC	F-score	MRR	MAP _{ref}
NICT	Run-1	NEWS12/15	0.153722	0.781110	0.226355	0.153722
NICT	Run-2	NEWS12/15	0.129450	0.762891	0.189012	0.129450
ITTB	Run-1	NEWS12/15	0.115696	0.757194	0.197850	0.115696
ITTB	Run-2	NEWS12/15	0.101133	0.746325	0.161466	0.101133
UALB	Run-1	NEWS12/15	0.156149	0.774646	0.241982	0.156149

Table A26: Results for the Thai to English transliteration task (ThEn) on Evaluation Test.

Team	Submission	Test Set	ACC	F-score	MRR	MAP _{ref}
UALB	Run-1	NEWS11	0.381500	0.860210	0.517000	0.375188
UALB	Run-2	NEWS11	0.360500	0.853419	0.476237	0.354408
NICT	Run-1	NEWS11	0.359500	0.852437	0.471200	0.354309
NICT	Run-2	NEWS11	0.329000	0.837196	0.425778	0.324021
ITTB	Run-1	NEWS11	0.342000	0.844966	0.468104	0.336937
ITTB	Run-2	NEWS11	0.335000	0.847316	0.453176	0.333686

Table A27: Results for the English to Persian transliteration task (EnPe) on Progress Test.

Team	Submission	Test Set	ACC	F-score	MRR	MAP _{ref}
UALB	Run-1	NEWS12/15	0.683301	0.942521	0.782315	0.658255
UALB	Run-2	NEWS12/15	0.710173	0.949957	0.807624	0.690791
NICT	Run-1	NEWS12/15	0.696737	0.948468	0.789989	0.682543
NICT	Run-2	NEWS12/15	0.565259	0.911092	0.668964	0.550183
ITTB	Run-1	NEWS12/15	0.619962	0.929311	0.740966	0.604472
ITTB	Run-2	NEWS12/15	0.622841	0.931697	0.723980	0.610456

Table A28: Results for the English to Persian transliteration task (EnPe) on Evaluation Test.