

8th Workshop on Asian Translation (WAT 2021)

Held online due to COVID-19

Bangkok, Thailand
5 – 6 August 2021

ISBN: 978-1-7138-3390-1

Printed from e-media with permission by:

Curran Associates, Inc.
57 Morehouse Lane
Red Hook, NY 12571



Some format issues inherent in the e-media version may also appear in this print version.

Copyright© (2021) by the Association for Computational Linguistics and
the Asian Federation of Natural Language Processing
All rights reserved.

Printed with permission by Curran Associates, Inc. (2021)

For permission requests, please contact the Association for Computational Linguistics
at the address below.

Association for Computational Linguistics
209 N. Eighth Street
Stroudsburg, Pennsylvania 18360

Phone: 1-570-476-8006
Fax: 1-570-476-0860

acl@aclweb.org

Additional copies of this publication are available from:

Curran Associates, Inc.
57 Morehouse Lane
Red Hook, NY 12571 USA
Phone: 845-758-0400
Fax: 845-758-2633
Email: curran@proceedings.com
Web: www.proceedings.com

Table of Contents

Overview of the 8th Workshop on Asian Translation

Toshiaki Nakazawa, Hideki Nakayama, Chenchen Ding, Raj Dabre, Shohei Higashiyama, Hideya Mino, Isao Goto, Win Pa Pa, Anoop Kunchukuttan, Shantipriya Parida, Ondřej Bojar, Chenhui Chu, Akiko Eriguchi, Kaori Abe, Yusuke Oda and Sadao Kurohashi 1

NHK's Lexically-Constrained Neural Machine Translation at WAT 2021

Hideya Mino, Kazutaka Kinugawa, Hitoshi Ito, Isao Goto, Ichiro Yamada and Takenobu Tokunaga 46

Input Augmentation Improves Constrained Beam Search for Neural Machine Translation: NTT at WAT 2021

Katsuki Chousa and Makoto Morishita 53

NICT's Neural Machine Translation Systems for the WAT21 Restricted Translation Task

Zuchao Li, Masao Utiyama, Eiichiro Sumita and Hai Zhao 62

Machine Translation with Pre-specified Target-side Words Using a Semi-autoregressive Model

Seiichiro Kondo, Aomi Koyama, Tomoshige Kiyuna, Toshio Hirasawa and Mamoru Komachi .. 68

NECTEC's Participation in WAT-2021

Zar Zar Hlaing, Ye Kyaw Thu, Thazin Myint Oo, Mya Ei San, Sasiporn Usanavasin, Ponrudee Netisopakul and Thepchai Supnithi 74

Hybrid Statistical Machine Translation for English-Myanmar: UTYCC Submission to WAT-2021

Ye Kyaw Thu, Thazin Myint Oo, Hlaing Myat Nwe, Khaing Zar Mon, Nang Aeindray Kyaw, Naing Linn Phyto, Nann Hwan Khun and Hnin Aye Thant 83

NICT-2 Translation System at WAT-2021: Applying a Pretrained Multilingual Encoder-Decoder Model to Low-resource Language Pairs

Kenji Imamura and Eiichiro Sumita 90

Rakuten's Participation in WAT 2021: Examining the Effectiveness of Pre-trained Models for Multilingual and Multimodal Machine Translation

Raymond Hendy Susanto, Dongzhe Wang, Sunil Yadav, Mausam Jain and Ohnmar Htun 96

BTS: Back TranScription for Speech-to-Text Post-Processor using Text-to-Speech-to-Text

chanjun park, Jaehyung Seo, Seolhwa Lee, Chanhee Lee, Hyeonseok Moon, Sugyeong Eo and Heuiseok Lim 106

Zero-pronoun Data Augmentation for Japanese-to-English Translation

Ryokan Ri, Toshiaki Nakazawa and Yoshimasa Tsuruoka 117

Evaluation Scheme of Focal Translation for Japanese Partially Amended Statutes

Takahiro Yamakoshi, Takahiro Komamizu, Yasuhiro Ogawa and Katsuhiko Toyama 124

TMU NMT System with Japanese BART for the Patent task of WAT 2021

Hwichan Kim and Mamoru Komachi 133

System Description for Transperfect

Wiktor Stribizew, Fred Bane, José Conceição and Anna Zaretskaya 138

<i>Bering Lab's Submissions on WAT 2021 Shared Task</i>	
Heesoo Park and Dongjun Lee	141
<i>NLPHut's Participation at WAT2021</i>	
Shantipriya Parida, Subhadarshi Panda, Ketan Kotwal, Amulya Ratna Dash, Satya Ranjan Dash, Yashvardhan Sharma, Petr Motliceck and Ondřej Bojar	146
<i>Improved English to Hindi Multimodal Neural Machine Translation</i>	
Sahinur Rahman Laskar, Abdullah Faiz Ur Rahman Khilji, Darsh Kaushik, Partha Pakray and Sivaji Bandyopadhyay	155
<i>IITP at WAT 2021: System description for English-Hindi Multimodal Translation Task</i>	
Baban Gain, Dibyanayan Bandyopadhyay and Asif Ekbal	161
<i>ViTA: Visual-Linguistic Translation by Aligning Object Tags</i>	
Kshitij Gupta, Devansh Gautam and Radhika Mamidi	166
<i>TMEKU System for the WAT2021 Multimodal Translation Task</i>	
Yuting Zhao, Mamoru Komachi, Tomoyuki Kajiwara and Chenhui Chu	174
<i>Optimal Word Segmentation for Neural Machine Translation into Dravidian Languages</i>	
Prajit Dhar, Arianna Bisazza and Gertjan van Noord	181
<i>Itihasa: A large-scale corpus for Sanskrit to English translation</i>	
Rahul Aralikkatte, Miryam de Lhoneux, Anoop Kunchukuttan and Anders Søgaard	191
<i>NICT-5's Submission To WAT 2021: MBART Pre-training And In-Domain Fine Tuning For Indic Languages</i>	
Raj Dabre and Abhisek Chakrabarty	198
<i>How far can we get with one GPU in 100 hours? CoAStAL at MultiIndicMT Shared Task</i>	
Rahul Aralikkatte, Héctor Ricardo Murrieta Bello, Daniel Hershovich, Marcel Bollmann and Anders Søgaard	205
<i>IIT Hyderabad Submission To WAT 2021: Efficient Multilingual NMT systems for Indian languages</i>	
Sourav Kumar, Salil Aggarwal and Dipti Sharma	212
<i>Language Relatedness and Lexical Closeness can help Improve Multilingual NMT: IITBombay@MultiIndicNMT WAT2021</i>	
Jyotsana Khatri, Nikhil Saini and Pushpak Bhattacharyya	217
<i>Samsung R&D Institute Poland submission to WAT 2021 Indic Language Multilingual Task</i>	
Adam Dobrowolski, Marcin Szymański, Marcin Chochowski and Paweł Przybyśz	224
<i>Multilingual Machine Translation Systems at WAT 2021: One-to-Many and Many-to-One Transformer based NMT</i>	
Shivam Mhaskar, Aditya Jain, Aakash Banerjee and Pushpak Bhattacharyya	233
<i>IITP-MT at WAT2021: Indic-English Multilingual Neural Machine Translation using Romanized Vocabulary</i>	
Ramakrishna Appicharla, Kamal Kumar Gupta, Asif Ekbal and Pushpak Bhattacharyya	238
<i>ANVITA Machine Translation System for WAT 2021 MultiIndicMT Shared Task</i>	
Pavanpankaj Vegi, Sivabhavani J, Biswajit Paul, Chitra Viswanathan and Prasanna Kumar K R	244