

Yan Shao (邵研)

Personal Information

Born January 31, 1990, Ningxia, China
Nationality: Chinese (China PR)
E-mail: yan.shao@lingfil.uu.se
Web: www.shaoyan.se
Tel: +46 (0) 7606 99989
Office address: Rum 9-2032, Thunbergsvgen, 3 H, Engelska Parken, Uppsala, Sweden
Post address: Box 635, 751 26, Uppsala, Sweden

Education

PhD - Computational Linguistics 2014.11 – 2018.6
Department of Linguistics and Philology
Uppsala University, Sweden

Master of Arts - Language Technology 2012.9 – 2014.9
Department of Linguistics and Philology
Uppsala University, Sweden

Bachelor of Arts - Applied and Computational Linguistics 2008.9 – 2012.7
Department of Chinese Language and Literature
Yuanpei Program
Peking University, China

Yinchuan No.9 Middle School, Ningxia, China 2005.9 – 2008.6

Working Experience

Department of Linguistics and Philology, Uppsala University, Sweden
- *Researcher* 2018.6 –

- *Director:* Professor Joakim Nivre, joakim.nivre@lingfil.uu.se

Ubiquitous Knowledge Processing (UKP) Lab, TU Darmstadt, Darmstadt, Germany
- *Research Assistant* 2015.6 – 2015.11

- *Director:* Professor Iryna Gurevych, gurevych@ukp.informatik.tu-darmstadt.de
- *Project:* Query-based summarisation for Question-Answering System

Barcelona Vibes/Stoke Travel, Barcelona, Spain
- *IT Manager* 2013.6 – 2013.9

- *Director:* Stuart Jones, stuart@barcelonavibes.com
- *Project:* Web design and SEO work

Institute of Computer Science and Technology of Peking University, Beijing, China
- *Research Intern* 2009.6 – 2011.6

- *Director:* Professor Wan Xiaojun, wanxiaojun@icst.pku.edu.cn
- *Project:* Text classification and sentiment analysis

Skills

Programming Languages: C#, Java, Python
Operating Systems: Linux/Unix, Windows, Mac OS
Deep Learning Libraries: TensorFlow, Pytorch, Numpy
Editors: Microsoft Office, Pycharm, L^AT_EX
Web Design: Wordpress, html, css

Language Skills: Chinese (native), English (fluent), Swedish (intermediate), Spanish (basic)

Github Projects

Joint Word Segmentation and POS Tagging System: github.com/yanshao9798/tagger
Universal (Multilingual) Word/Morpheme Segmenter: github.com/yanshao9798/segmenter

Research Interests

Deep learning
Sequence processing, segmentation, tagging and named entity recognition
Automatic summarisation
Machine transliteration
Syntactic parsing

Teaching

Machine Learning (master's course)
Basic Text Analysis (bachelor's course)
Machine Translation (bachelor/master's course)

Scholarships and Awards

Feedzai Scholarship, Feedzai	2016
Erasmus Traineeship Grant, EU Lifelong Learning Programme	2015
Chinese Government Scholarship, China Scholarship Council	2014
Erasmus Placement Scholarship, EU Lifelong Learning Programme	2013
Chefredaktor Axel Johansson's Scholarship, Uplands Nation	2013
Anders Wall Scholarship, Uppsala University	2013

Other Research and Social Activities

<i>Reviewer/Program Committee member:</i> ACL, COLING, TALLIP	
<i>Invited speech:</i> Statistical study of journey to the west (with Professor Nicholas Koss)	2012.6
<i>Dialectic study:</i> Chaozhou, China (supervised by Professor Li Xiaofan)	2011.7

Hobbies

Guitar, Writing, Travelling

Publications

- **Yan Shao**, Christian Hardmeier, Joakim Nivre (2018). Universal Word Segmentation: Implementation and Interpretation. *Transactions of Association of Computational Linguistics*, (To appear).
- **Yan Shao**, Christian Hardmeier, Jörg Tiedemann, Joakim Nivre (2017). Character-Based Joint Segmentation and POS Tagging for Chinese using Bidirectional RNN-CRF. In *Proceedings of the 8th International Joint Conference on Natural Language Processing*, pages 173–183, Taipei, Taiwan, China.
- **Yan Shao**, Christian Hardmeier, Joakim Nivre (2017). Recall is the Proper Evaluation Metric for Word Segmentation. In *Proceedings of the 8th International Joint Conference on Natural Language Processing*, pages 86–90, Taipei, Taiwan, China.
- **Yan Shao** (2017). Cross-Lingual Word Segmentation and Morpheme Segmentation as Sequence Labelling. In *Proceedings of MLP 2017: The First Workshop on Multi-Language Processing in a Globalising World*, pages 75–80, Dublin, Ireland.
- Miryam de Lhoneux, **Yan Shao**, Ali Basirat, Eliyahu Kiperwasser, Sara Stymne, Yoav Goldberg, Joakim Nivre (2017). From Raw Text to Universal Dependencies - Look, No Tags! In *Proceedings of the CoNLL 2017 Shared Task: Multilingual Parsing from Raw Text to Universal Dependencies*, pages 207–217, Vancouver, Canada.
- **Yan Shao**, Joakim Nivre (2016). Applying Neural Networks to English-Chinese Named Entity Transliteration. In *Proceedings of the Sixth Named Entity Workshop, joint with 54th ACL*, pages 73–77, Berlin, Germany.
- **Yan Shao**, Christian Hardmeier, Joakim Nivre (2016). Multilingual Named Entity Recognition using Hybrid Neural Networks. In *Proceedings of the 6th Swedish Conference on Language Technology*, Umeå, Sweden.
- **Yan Shao**, Jörg Tiedemann, Joakim Nivre (2015). Boosting English-Chinese Machine Transliteration via High Quality Alignment and Multilingual Resources. In *Proceedings of the Fifth Named Entity Workshop, joint with 53rd ACL and the 7th IJCNLP*, pages 56–60, Beijing, China.
- **Yan Shao**, Jörg Tiedemann (2014). Machine Transliteration of Names from Different Language Origins into Chinese. In *Proceedings of the 5th Swedish Conference on Language Technology*, Uppsala, Sweden.
- Liqiang Guo, Fangzhou Zhai, **Yan Shao**, Xiaojun Wan (2010). PKUTM at TREC 2010 Blog Track. In *The 19th Text REtrieval Conference (TREC 2010) Proceedings*.