

NIKO SCHENK

Computational Linguist (PhD)

@ nikoschenk@gmail.com Berlin, Germany english-linguistics.de/nschenk/ Google Scholar
in linkedin.com/in/niko-schenk-88927a41/ github.com/nikoschenk orcid.org/0000-0002-6081-6023



EXPERIENCE

Language Engineer II, Alexa AI



Amazon Development Center Germany

October 2020 – Ongoing Berlin, Germany

- Improved Alexa's performance and functionalities through self-developed experiments and various accuracy initiatives.
- Successfully released multiple new system versions for French.

Postdoctoral Researcher in NLP



Applied Computational Linguistics Lab, Goethe University Frankfurt Research Assistant & PhD Candidate between 2013–2019

September 2019 – October 2020 Frankfurt, Germany

- Developed a state-of-the-art Chinese implicit discourse parser.

Computational Linguist Intern



Google

July 2019 – September 2019 Zurich, Switzerland

- Implemented a semantic topic framework for dialog modeling NLU.

Research Intern



SONY Deutschland GmbH, Stgt Tech Center, Speech & Sound Group

October 2010 – March 2011 Stuttgart, Germany

- Successful patent grant for multi-language sentiment analysis.

Computational Linguist Intern



IBM Deutschland GmbH, Text Analysis, Information Management

August 2009 – January 2010 Böblingen, Germany

- Analysis of user-generated content and integration into IBM product.

OPEN SOURCE CONTRIBUTIONS

Self-Developed Design & Implementation of *Beta Writer*

The First Entirely Machine-Generated Research Book (406k downloads)

April 2019 In cooperation with **SPRINGER NATURE**

GSoC Mentor 2018 & 2019



Google Summer of Code

Machine Translation & Automated Analysis of Cuneiform Languages

Summer 2018, 2019 Frankfurt/Toronto/Mumbai, Delhi

- Mentored **Bakhtiyar Syed** and **Ravneet Punia** who implemented the first **Sumerian** machine translation and semantic role labeling system.

PROGRAMMING LANGUAGES

Java ●●●●●

Python ●●●●●

FRAMEWORKS/TOOLS

Git Keras Deeplearning4J

Hugging Face Transformers

EDUCATION

PhD in Computer Science / NLP

Goethe University Frankfurt, Germany

January 2015 – September 2019

Thesis: *Retrieving Implicit Relations from Text: Hidden Semantics and Natural Language Processing*

Final grade: **summa cum laude** (excellent)

M.A. in Computational Linguistics

University of Tübingen, Germany

April 2011 – March 2013

Thesis: *A Suffix-Array Based Algorithm for the Extraction of Enhanced N-Grams*

Final grade: 1.0 (A)

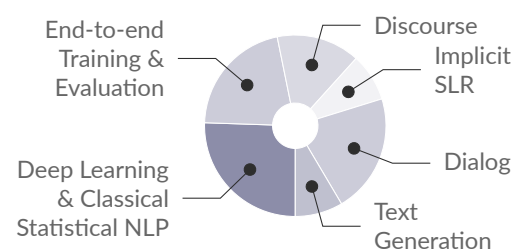
B.A. in Computational Linguistics

University of Tübingen, Germany

October 2007 – October 2010

Thesis: *A Finite-State Approach to Sentiment Analysis*, Final grade: 1.3 (A-)

RESEARCH INTERESTS



SELECTED PUBLICATIONS

Journal Article

- Chiarcos, Christian et al. (2018a). "Annotating a Low-Resource Language with LLOD Technology: Sumerian Morphology and Syntax". In: *Information* 9.11, p. 290. DOI: 10.3390/info9110290. URL: <https://doi.org/10.3390/info9110290>.

Conference Proceedings

- Punia, Ravneet et al. (Dec. 2020). "Towards the First Machine Translation System for Sumerian Transliterations". In: *Proceedings of the 28th International Conference on Computational Linguistics. (Short Papers, 26.2% acceptance rate)*. Barcelona, Spain (Online): International Committee on Computational Linguistics, pp. 3454–3460. DOI: 10.18653/v1/2020.coling-main.308. URL: <https://www.aclweb.org/anthology/2020.coling-main.308>.
- Chiarcos, Christian and Niko Schenk (2019). "CoNLL-Merge: Efficient Harmonization of Concurrent Tokenization and Textual Variation". In: *2nd Conference on Language, Data and Knowledge (LDK 2019)*. Ed. by Maria Eskevich et al. Vol. 70. OpenAccess Series in Informatics (OASICS). Dagstuhl, Germany: Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik, 7:1–7:14. ISBN: 978-3-95977-105-4. DOI: 10.4230/OASICS.LDK.2019.7. URL: <http://drops.dagstuhl.de/opus/volltexte/2019/10371/pdf/OASICS-LDK-2019-7.pdf>.
- – (2018). "The ACoLi CoNLL Libraries: Beyond Tab-Separated Values". In: *Proceedings of the Eleventh International Conference on Language Resources and Evaluation, LREC 2018, Miyazaki, Japan, May 7-12, 2018*. URL: <http://aclweb.org/anthology/L18-1090>.
- Chiarcos, Christian et al. (2018b). "Towards a Linked Open Data Edition of Sumerian Corpora". In: *Proceedings of the Eleventh International Conference on Language Resources and Evaluation, LREC 2018, Miyazaki, Japan, May 7-12, 2018*. URL: <http://www.lrec-conf.org/proceedings/lrec2018/pdf/862.pdf>.
- Schenk, Niko and Christian Chiarcos (2017). "Resource-Lean Modeling of Coherence in Commonsense Stories". In: *Proceedings of the 2nd Workshop on Linking Models of Lexical, Sentential and Discourse-level Semantics*. Valencia, Spain: Association for Computational Linguistics, pp. 68–73. URL: <http://aclweb.org/anthology/W17-0910>.
- Schenk, Niko, Samuel Rönnqvist, and Christian Chiarcos (2017). "A Recurrent Neural Model with Attention for the Recognition of Chinese Implicit Discourse Relations". In: *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Short Papers, 18% acceptance rate)*. Vancouver, Canada: Association for Computational Linguistics, pp. 256–262. DOI: 10.18653/v1/P17-2040. URL: <http://www.aclweb.org/anthology/P17-2040>.
- Schenk, Niko and Christian Chiarcos (2016). "Unsupervised Learning of Prototypical Fillers for Implicit Semantic Role Labeling". In: *NAACL HLT 2016, The 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, San Diego California, USA, June 12-17, 2016. (Short Papers, 25% acceptance rate)*, pp. 1473–1479. URL: <http://aclweb.org/anthology/N/N16/N16-1173.pdf>.
- Schenk, Niko et al. (2016). "Do We Really Need All Those Rich Linguistic Features? A Neural Network-Based Approach to Implicit Sense Labeling". In: *Proceedings of the CoNLL-16 shared task. Winner of Chinese Explicit Sense Classification Subtask*. Berlin, Germany: Association for Computational Linguistics, pp. 41–49. DOI: 10.18653/v1/K16-2005. URL: <http://aclweb.org/anthology/K16-2005>.
- Chiarcos, Christian and Niko Schenk (2015a). "A Minimalist Approach to Shallow Discourse Parsing and Implicit Relation Recognition". In: *Proceedings of the 19th Conference on Computational Natural Language Learning: Shared Task, CoNLL 2015, Beijing, China, July 30-31, 2015*, pp. 42–49. URL: <http://aclweb.org/anthology/K/K15/K15-2006.pdf>.
- – (2015b). "Memory-Based Acquisition of Argument Structures and its Application to Implicit Role Detection". In: *Proceedings of the SIGDIAL 2015 Conference, The 16th Annual Meeting of the Special Interest Group on Discourse and Dialogue, 2-4 September 2015, Prague, Czech Republic. Best Paper Award Nominee*, pp. 178–187. URL: <http://aclweb.org/anthology/W/W15/W15-4626.pdf>.
- Eger, Steffen, Niko Schenk, and Alexander Mehler (2015). "Towards Semantic Language Classification: Inducing and Clustering Semantic Association Networks from Europarl". In: *Proceedings of the Fourth Joint Conference on Lexical and Computational Semantics, *SEM 2015, June 4-5, 2015, Denver, Colorado, USA.*, pp. 127–136. URL: <http://aclweb.org/anthology/S/S15/S15-1014.pdf>.

Machine-Generated Book (Algorithm Design & Implementation with 406,000 downloads)

- Writer, Beta (2019). "Cathode Materials, Samples, Pristine, Layered, Doping, Discharge Capacity". In: *Lithium-Ion Batteries: A Machine-Generated Summary of Current Research*. Cham: Springer International Publishing, pp. 73–161. ISBN: 978-3-030-16800-1. DOI: 10.1007/978-3-030-16800-1_2. URL: https://doi.org/10.1007/978-3-030-16800-1_2.

Patent

- Eggink, Jana et al. (Apr. 2014). *Method for determining a sentiment from a text*. US Patent App. 14/111,101. URL: <https://patents.google.com/patent/US20140114648A1/en>.