

Daniil Sorokin

UKP Lab, Hochschulstraße 10
64289 Darmstadt, Germany
✉ sorokin@ukp.informatik.tu-darmstadt.de
📄 daniilsorokin.github.io

Keywords: natural language processing, machine learning, semantic analysis, knowledge bases

Work Experience

- 04.2015 – now **Researcher**, *UKP Lab, TU Darmstadt, Germany*.
- Neural network approaches to knowledge base question answering
 - Teaching the course “Foundations of Natural Language Processing with Python”
 - Bachelor and master student supervision
 - Organizing a regular research seminar and guest talks
 - Candidate evaluation for PhD- and PostDoc-level positions
- 11.2018 – (01.2019) **Research Intern**, *Factmata Ltd., London, UK*.
- Automatic fact checking and claim validation
- 01.2014 – 03.2015 **Researcher**, *Research center SFB 833, University of Tübingen, Germany*.
- Automatic semantic processing of German and English noun compounds
- 05.2012 – 12.2013 **Research Assistant**, *Computational Linguistics department, University of Tübingen, Germany*.
- Teaching assistant for the course “Programming with Java”
 - Development of RESTful web services for linguistic resources
- 10.2010 – 08.2011 **Developer**, *PROMT Ltd., Saint Petersburg, Russia*.
- Stress- and usability-tests of machine translation tools
 - Construction of multi-lingual parallel corpora from web data

Education

- 04.2015 – (06.2019) **PhD student in Computer Science**, *TU Darmstadt, Germany*.
Supervisor: Prof. Dr. Iryna Gurevych, UKP Lab
Topic: Lexical-Semantic and World Knowledge for Semantic Parsing
- 10.2011 – 01.2014 **M.A. in Computational Linguistics**, *University of Tübingen, Germany*,
GPA: 4.0 (German 1.1 / 1.0).
Thesis: A Machine Learning Approach to the Semantic Classification of German Noun Compounds
- 09.2007 – 07.2011 **B.A. in Computational Linguistics**, *Saint Petersburg State University, Russia*, GPA: 4.0 (Russian 4.9 / 5.0).
Thesis: Automatic Identification of Morphological Types of Geographical Names and Their Representation in a Geo-information System
- Summer schools Machine Learning Summer School (2016, Max Plank Institute), The European Summer School in Logic, Language and Information (ESSLLI 2014)

Skills

Programming languages	Main: Python, Java	Experience	Javascript, XML, CSS, \LaTeX , with: SPARQL, SQL, R, Haskell, Scala
Main frameworks	PyTorch, NumPy, Keras, TensorFlow, Flask, Weka, Stanford CoreNLP		
Languages	English, German, Russian	Misc.	Climbing, cycling, skiing
Awards	Outstanding Reviewer (ACL, 2018), Teaching Award in Computer Science (TU Darmstadt, 2017), Best Challenge Paper (ESWC, 2017), Fully-funded Master’s scholarship (DAAD, 2011)		

Main Publications

Daniil Sorokin and Iryna Gurevych. “Interactive Instance-based Evaluation of Knowledge Base Question Answering”. In: *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing (EMNLP 2018): System Demonstrations*. Brussels, Belgium: Association for Computational Linguistics, Nov. 2018, pp. 114–119. URL: <http://aclweb.org/anthology/D18-2020>.

Daniil Sorokin and Iryna Gurevych. “Mixing Context Granularities for Improved Entity Linking on Question Answering Data across Entity Categories”. In: *Proceedings of the Seventh Joint Conference on Lexical and Computational Semantics (*SEM 2018 @ NAACL 2018)*. New Orleans, LA, U.S.: Association for Computational Linguistics, June 2018, pp. 65–75. URL: <http://aclweb.org/anthology/S18-2007>.

Daniil Sorokin and Iryna Gurevych. “Modeling Semantics with Gated Graph Neural Networks for Knowledge Base Question Answering”. In: *Proceedings of COLING 2018, the 27th International Conference on Computational Linguistics*. Santa Fe, NM, U.S.: Association for Computational Linguistics, Aug. 2018, pp. 3306–3317. URL: <http://aclweb.org/anthology/C18-1280>.

Daniil Sorokin and Iryna Gurevych. “Context-Aware Representations for Knowledge Base Relation Extraction”. In: *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing (EMNLP 2017)*. Copenhagen, Denmark: Association for Computational Linguistics, Sept. 2017, pp. 1785–1790. URL: <http://aclweb.org/anthology/D17-1188>.

Daniil Sorokin and Iryna Gurevych. “End-to-end Representation Learning for Question Answering with Weak Supervision”. In: *ESWC 2017 Semantic Web Challenges*. Portoroz, Slovenia, 2017.

Daniil Sorokin, Corina Dima, and Erhard Hinrichs. “Classifying Semantic Relations in German Nominal Compounds using a Hybrid Annotation Scheme”. In: *Journal of Cognitive Science* 16.3 (2015), pp. 260–285.

Google Scholar profile: <https://scholar.google.de/citations?user=PRX-B-oAAAAJ>

Question Answering Demo: <http://semanticparsing.ukp.informatik.tu-darmstadt.de:5000/question-answering/>