William K. Roberts

Pankstrasse 62, 13357 Berlin, Germany

Tel.: +49 176 4345 8016 ■ Email: wildwilhelm@gmail.com

Skills:

Natural language processing and machine learning

- Natural language toolkit (NLTK, contributor), scikit-learn, word2vec, WordNet, machine translation, semantic role labelling, word sense disambiguation, HPSG.
- Parsing, lemmatisation, morphological analysis. Corpus construction and annotation.

Software development and devops

- Programming languages: Python (contributor), Javascript, Shell, C/C++, Java, Go, Perl, PHP, Ruby.
- Devops: Kubernetes, Docker, Helm, Terraform, Gitlab administration including CI/CD.
- Software: git, MySQL, Redis, Nginx, RabbitMQ, PostgreSQL, MongoDB.
- Web frameworks: Flask, Django, Angular, React.

Work Experience:

Devops engineer, freelance, Realtainment, Berlin (Sept. 2019 – present)

- Dockerized and migrated a legacy high traffic e-commerce website to Kubernetes; created highly available MySQL, ElasticSearch and Redis clusters; and designed and implemented automated tools for data collection and backup.

Software architect, freelance, Ventura Travel, Berlin (May 2018 – Dec. 2018)

- Designed a new business administration system using microservice architecture in Kubernetes, and implemented the first component, a high-availability authentication, user management, and access control server written in Javascript.

Research assistant,
QTLeap: Quality translation by deep language engineering approaches (June 2015 – Nov. 2016)

- Created word embedding models for 8 European languages, and employed these to create automatic thesaurus models, as well as for automatic acquisition of non-compositional multiword expressions.
- Conducted experiments on machine translation of multiword expressions for English-Spanish, Spanish-English, and English-Basque language pairs.
- Contributed to deliverable reports and presented results, both internally, as well as to the European Commission.

Student research workshop chair,
Annual meeting of the Association for Computational Linguistics (2016)

- Created the website for the workshop, organised program committee, coordinated the student mentoring program and the travel grant process, solicited article submissions from students, oversaw peer review, made arrangements for a 3-hour poster session with 22 student presenters, and aided with the publication of the workshop proceedings.

University lecturer,
Institut für Anglistik and Amerikanistik, Humboldt University of Berlin (Apr. 2012 – Apr. 2018)

- Designed and taught undergraduate courses in *Introductory Semantics*, *Discourse Representation Theory*, and a seminar in *Verb Semantics*; created, assigned, and graded term papers and final examinations.

Research assistant,
NLX Lab, Department of Computer Science, University of Lisbon (Sept. 2011 – Feb. 2012)

- Semi-automatically extended lexicon and lexical type hierarchy of an HPSG grammar for Portuguese (LXGram).

Research assistant,
German Research Centre for Artificial Intelligence (DFKI), Saarbrücken (Sept. 2008 – May 2011)

- Developed Django-based web applications for sharing and searching linguistic resources (Meta-Share) and server infrastructure for machine translation (MT Server Land).
- Created language analysis software and maintained a web-based interface for semi-automatically building bilingual terminologies in several Western European languages for translation at the European Patent Office.

Voice-over-IP software developer, Namzak Labs, Kingston, Canada (Sept. 2005 – Jul. 2007)

- Worked in a small team of developers to design, implement and debug an Internet telephony system called Arrowfone, using a wide variety of programming tools.
- Improved the Arrowfone voice-over-IP client running on Symbian OS Nokia smartphones and ported the software to run on Windows PCs.
- Maintained, tested and extended the Arrowfone directory server, a highly scalable multithreaded Linux daemon, over a one year period.

Video game developer, Queen's University, Kingston, Canada (May 2005 – Aug. 2005)

- Designed, implemented and debugged a prototype 3D video game controlled by a stationary exercise bicycle for a new computer science research project. The program was written in C++ to be OS-independent, and was built using open-source tools and libraries.
- Assisted in the creation and documentation of a set of tools for use in new university courses on video game development and digital animation.

Education:

PhD, Computational Semantics
Humboldt-Universität zu Berlin, Berlin, Germany (Apr. 2013 – present)

- Thesis topic: Unsupervised lexical acquisition for German verbs.
- Supervisor: Prof. Markus Egg

Master of Science, Computational Linguistics
Universität des Saarlandes, Saarbrücken, Germany (Oct. 2007 – May 2011)

- Master's thesis "Integrating syntax and semantics for word sense disambiguation"
- Thesis supervisor: Dr. Valia Kordoni
- Final average 1,3 (A-)

Bachelor of Science (Honours), Computer Science and Psychology Double Major
Queen's University, Kingston, Ontario, Canada (Sept. 2000 – May 2007)

- Bachelor's thesis "Neural networks to recognise images"
- Thesis supervisor: Dr. Roger Browse
- Weighted average 85.1% (GPA 3.7)

Other Experience:

Native English speaker, with excellent spoken and written German and French

European driving licence

Publications:

Markus Egg, Helena Prepens, and Will Roberts. Annotation and automatic classification of aspectual categories. In: *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*. Association for Computational Linguistics, pp. 3335–3341, July 2019.

Will Roberts and Markus Egg. A large automatically-acquired all-words list of multiword expressions scored for compositionality. In: *Proceedings of the Eleventh International Conference on Language Resources and Evaluation (LREC 2018)*. European Language Resources Association, May 2018.

He, He, Tao Lei, and Will Roberts, eds. Proceedings of the ACL 2016 student research workshop. Association for Computational Linguistics, August 2016.

Will Roberts and Markus Egg. Integration of automatically-acquired multiword expressions in a hybrid machine translation system. PARSEME sixth general meeting, Struga, Macedonia, 7–8 April 2016.

Will Roberts and Markus Egg. A comparison of selectional preference models for automatic verb classification. In: *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*. Association for Computational Linguistics, pp. 511–522, October 2014.

Will Roberts, Markus Egg, and Valia Kordoni. Subcategorisation acquisition from raw text for a free word-order language. In: *Proceedings of the 14th Conference of the European Chapter of the Association for Computational Linguistics (EACL 2014)*. Association for Computational Linguistics, pp. 298–307, April 2014.

Egg, Markus, Will Roberts and Valia Kordoni. Multiword expression identification for German. PARSEME second general meeting, Athens, Greece, 10–11 March 2014.

Will Roberts and Valia Kordoni. Using verb subcategorization for word sense disambiguation. In: *Proceedings of the Eight International Conference on Language Resources and Evaluation (LREC'12)*, European Language Resources Association, May 2012.

T. C. Nicholas Graham and Will Roberts. Toward quality-driven development of 3D computer games. In: *Proceedings of the Thirteenth International Workshop on Design, Specification and Verification of Interactive Systems (DSV-IS 2006)*, Springer LNCS, pp. 248–261, July 2006.