

# Oliver Ferschke

[ˈoləvər ˈfɛrʃkə]

Pittsburgh, PA, USA  
☎ +1 (412) 979 2440  
✉ [oliver.ferschke@gmail.com](mailto:oliver.ferschke@gmail.com)  
🌐 [www.ferschke.com](http://www.ferschke.com)  
Citizenship: German  
Age: 34  
US visa: O-1A  
Last updated: September 11th, 2017

## Summary of Qualifications

- Eight years of experience in interdisciplinary research with a focus on natural language processing.
- Experience in hands-on processing, representation, and manipulation of unstructured text datasets.
- Proficiency in applied machine learning with a particular focus on text classification.
- Experience with NLP for medical and clinical applications
- Experienced with build tools (Maven, Gradle, Ant, Ivy) and continuous integration systems (Jenkins).
- Background in computer-supported education and intelligent tutoring.
- Interested in ambient conversational interfaces.
- Well-versed in effectively presenting to both expert and novice audiences.

## Professional Experience

- Since 9/2016 **Senior NLP Engineer**,  
(ongoing) *M\*Modal*,  
Pittsburgh, PA, USA.  
Research and development in the area of ambient conversational interfaces and conversational agents for the health care sector.
- 10/2014–8/2016 **Postdoctoral Research Associate**,  
*Language Technologies Institute, School of Computer Science*,  
Carnegie Mellon University, Pittsburgh, PA, USA.  
Research in natural language processing with a focus on online conversation analysis, text classification, learning analytics, and knowledge management in online collaboration environments. Lead developer of DiscourseDB, a general purpose discourse representation framework and co-director of a working group on technology based discussion support in online education
- 1/2014–9/2014 **Research Scientist**,  
*German Institute for International Educational Research (DIPF)*,  
Frankfurt am Main, Germany.  
Research in technology-based assessment with a focus on automatic short answer grading.
- 1/2010–12/2013 **Doctoral Researcher**,  
*Ubiquitous Knowledge Processing Lab*, Technische Universität Darmstadt, Germany.  
Foundational research in natural language processing with a focus on automatic quality assessment of collaboratively created texts, computer supported collaborative work, discourse analysis, and machine learning for NLP.

## Degrees

- 2010 -2014 **Dr.-Ing. (PhD) in Computer Science**,  
*Technische Universität Darmstadt*, Germany,  
Thesis: The Quality of Content in Open Online Collaboration Platforms: Approaches to NLP-supported Information Quality Management in Wikipedia.  
Grade: magna cum laude (very good)

- 2005 - 2009 **Magister Artium in English Linguistics**,  
*Julius-Maximilians University of Würzburg, Germany,*  
 Thesis: Semantic Relations in WordNet and the BNC.  
 Equivalent to a Master's degree.  
 Final grade 1.4 on a 5-grade scale (very good)
- 2004 - 2009 **Staatsexamen in Computer Science and English (double major)**,  
*Julius-Maximilians University of Würzburg, Germany.*  
 Government licensing examination for future teachers in college preparatory schools.  
 Equivalent to a Master's degree.  
 Final grade 1.9 on a 5-grade scale (good)

## Open Source Software Development

### **DiscourseDB**,

*A comprehensive and generic discourse representation that alleviates cross-platform discourse analysis.*,

Role: Lead Developer.

<https://github.com/discoursedb>

### **DKPro Text Classification Framework**,

*A UIMA-based text classification framework for simplifying machine learning experiments with textual data,*

Role: Co-Lead Developer.

<https://dkpro.github.io/dkpro-tc/>

### **Java Wikipedia Programming Library**,

*A Java-based interface for efficient random access to Wikipedia and its revision history,*

Role: Co-Lead Developer.

<https://dkpro.github.io/dkpro-jwpl/>

### **Java API for text n-gram frequency data in Web1T format**,

*A Java tool for efficiently searching n-gram data in the Web 1T 5-gram corpus format,*

Role: Co-Lead Developer.

<https://github.com/dkpro/jweb1t>

### **Darmstadt Knowledge Processing Repository (DKPro)**,

*A large collection of UIMA-based software components for NLP,*

Role: Developer.

<https://dkpro.github.io/dkpro-core/>

## Honours and Awards

- 8/2016 Recipient of an O-1A visa from the United States of America for individuals with an extraordinary ability in the sciences
- 8/2015 Participation grant for the 15th annual GAIN conference funded by the German Academic International Network
- 6/2015 Early career research fellowship for the participation in the CSCL 2015 early career workshop funded by the National Science Foundation
- 10/2013 Travel grant for early career researchers in the digital humanities from the Volkswagen Foundation  
 Acceptance rate: 17%
- 1/2011–12/2013 Three-year PhD scholarship from the state of Hesse, Germany within the research excellence program LOEWE

---

## Leadership

- Since 11/2014 **Working Group “Discussion Affordances for Natural Collaborative Exchange”**,  
*Co-Director*  
<http://dance.cs.cmu.edu>.
- 1/2011–12/2013 **LOEWE Research Priority Program “Digital Humanities”**,  
*Member of the board of directors.*
- 8/2008–12/2010 **German Informatics Society (GI)**,  
*Member of the student advisory board.*

---

## Skills

- Natural Languages **German (native), English (fluent), French (basic).**
- Tools/Frameworks **UIMA, Maven, Gradle, Artifactory, Jenkins, Eclipse, Hibernate, Spring, Git, SVN, Mercurial, MySQL, MongoDB.**
- Programming Languages **Java, SQL, LaTeX, XML, C#.**

---

## Publications

- 2017 Carolyn P. Rosé, Iris Howley, Miaomiao Wen, Diyi Yang, and **Oliver Ferschke**. 'Assessment of Discussion in Learning Contexts', in Alina von Davier, Mengziao Zhu, Patrick Kyllonen (Eds.): *Innovative Assessment of Collaboration*, New York, Springer, in press.
- 2016 Ryan S. Baker, Yuan Wang, Luc Paquette, Vincent Alevan, Octav Popescu, Jonathan Sewall, Carolyn P. Rosé, Gaurav Tomar, **Oliver Ferschke**, Jing Zhang, Michael Cennamo, Stephanie Ogden, Therese Condit, Jose W. Diaz, Scott Crossley, Danielle McNamara, Denise Comer, Collin Lynch, Rebecca Brown, Tiffany Barnes and Yoav Bergner. 'A MOOC on Educational Data Mining', in Samira Elatia, Donald Ipperciel, Osmar Zaiane (Eds.): *Data Mining and Learning Analytics: Applications in Educational Research*, Chapter 4. Wiley. October 2016.
- Carolyn P. Rosé and **Oliver Ferschke**. 'Technology Support for Discussion Based Learning: From Computer Supported Collaborative Learning to the Future of Massive Open Online Courses', in: *International Journal of Artificial Intelligence in Education*. June 2016.
- Yohan Jo, Gaurav Tomar, **Oliver Ferschke**, Carolyn P. Rosé, Dragan Gasevic: 'Expediting Support for Social Learning with Behavior Modeling', in: *Proceedings of the The 9th International Conference on Educational Data Mining*, Raleigh, NC, USA. June 2016
- Carolyn P. Rosé Dragan Gasevic, Pierre Dillenbourg, Yohan Jo, Gaurav Tomar, Oliver Ferschke, Gijbert Erkens, Anouschka van Leeuwen, Jeroen Janssen, Mieke Brekelmans, Jennifer Tan, Elizabeth Koh, Imelda Caleon, Christin Jonathan and Simon Yang: 'Analytics of social processes in learning contexts: A multi-level perspective', in: *Proceedings of the 12th International Conference of the Learning Sciences (ICLS 2016)*. Invited Symposium. Singapore, June 2016.
- Yohan Jo, Gaurav Tomar, **Oliver Ferschke**, Carolyn P. Rosé, and Dragan Gašević. 'Pipeline for Expediting Learning Analytics and Student Support from Data in Social Learning', in: *Proceedings of the Sixth Learning Analytics and Knowledge Conference*, Edinburgh, UK, April 2016. Poster.
- 2015 Miaomiao Wen, **Oliver Ferschke**, and Carolyn P. Rosé: 'Towards Support of Collaborative Reflection, Help Exchange, and Group Learning in MOOCs', in: *Proceedings of the 2nd Workshop on Learning with MOOCs*. New York, NY, USA, October 2015.

Yuan Wang, Ryan Baker, Vincent Alevan, Carolyn P. Rosé, Jonathan Sewall, Octav Popescu, Gaurav Tomar, **Oliver Ferschke**, and Michael Cennamo: 'Interactive Activities in MOOCs', in: *Proceedings of the 2nd Workshop on Learning with MOOCs*. New York, NY, USA, October 2015.

**Oliver Ferschke**, Iris Howley, Gaurav Tomar, Diyi Yang, Carolyn P. Rosé: 'Fostering Discussion across Communication Media in Massive Open Online Courses', in: *Proceedings of the 11th International Conference on Computer Supported Collaborative Learning*, pp. 459–466, Gothenburg, Sweden, 2015.

Carolyn P. Rosé, **Oliver Ferschke**, Gaurav Tomar, Diyi Yang, Iris Howley, Vincent Alevan, George Siemens, Matthew Crosslin, Dragan Gasevic and Ryan Baker: 'Challenges and Opportunities of Dual-Layer MOOCs: Reflections from an edX Deployment Study', in: *Proceedings of the 11th International Conference on Computer Supported Collaborative Learning*, Interactive Event, pp. 848–851, Gothenburg, Sweden, 2015.

**Oliver Ferschke**, Diyi Yang, Gaurav Tomar, Carolyn P. Rosé: 'Positive Impact of Collaborative Chat Participation in an edX MOOC', in: C. Conati et al. (Eds.): *Proceedings of the 17th International Conference on Artificial Intelligence in Education*, LNAI 9112, pp. 115–124, Madrid, Spain, 2015.

Iris Howley, Gaurav Tomar, Diyi Yang, **Oliver Ferschke**, Carolyn P. Rosé: 'Alleviating the Negative Effect of Up and Downvoting on Help Seeking in MOOC Discussion Forums', in: *Proceedings of the 17th International Conference on Artificial Intelligence in Education*, Short Paper, pp. 629–632, Madrid, Spain, 2015.

**Oliver Ferschke**, Gaurav Tomar, Carolyn P. Rosé: 'Adapting Collaborative Chat for Massive Open Online Courses: Lessons Learned', in: *Proceedings of the 4th Workshop on Intelligent Support for Learning in Groups*. Workshop at AIED 2015, pp. 13–18, Madrid, Spain, 2015.

**Oliver Ferschke**, Diyi Yang, Carolyn P. Rosé: 'A Lightly Supervised Approach to Role Identification in Wikipedia Talk Page Discussions', in: *Wikipedia, a Social Pedia: Research Challenges and Opportunities*. Workshop at ICWSM 2015, pp. 43–47, Oxford, UK, 2015.

2014 Neil Dhruva, **Oliver Ferschke** and Iryna Gurevych: 'Solving Open-Domain Multiple Choice Questions with Textual Entailment and Text Similarity Measures', in: *CLEF 2014 Labs and Workshop, Notebook Papers*, p. 1375–1385, Sheffield, UK, September 2014.

Lucie Flekova, **Oliver Ferschke** and Iryna Gurevych: 'UKPDIPF: A Lexical Semantic Approach to Sentiment Polarity Prediction In Twitter Data', in: *Semeval-2014 Task 9: Sentiment Analysis in Twitter. Proceedings of the 8th International Workshop on Semantic Evaluation*, p.704–710, Dublin, Ireland, August 2014.

**Oliver Ferschke**: 'The Quality of Content in Open Online Collaboration Platforms: Approaches to NLP-supported Information Quality Management in Wikipedia', PhD Thesis, Technische Universität Darmstadt, July 2014.

Johannes Daxenberger, **Oliver Ferschke**, Iryna Gurevych and Torsten Zesch: 'DKPro TC: A Java-based Framework for Supervised Learning Experiments on Textual Data', in: *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics. System Demonstrations*, p. 61–66, Baltimore, MD, USA, June 2014.

Lucie Flekova, **Oliver Ferschke** and Iryna Gurevych: 'What Makes a Good Biography? Multidimensional Quality Analysis Based on Wikipedia Article Feedback Data', in: *Proceedings of the 23rd International World Wide Web Conference (WWW 2014)*, pp. 855–866, Seoul, Korea, April 2014.

- 2013 **Oliver Ferschte**, Iryna Gurevych and Marc Rittberger: 'The Impact of Topic Bias on Quality Flaw Prediction in Wikipedia', in: *Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pp. 721–730, Sofia, Bulgaria, August 2013.
- Oliver Ferschte**, Johannes Daxenberger and Iryna Gurevych: 'A Survey of NLP Methods and Resources for Analyzing the Collaborative Writing Process in Wikipedia', in Iryna Gurevych and Jungi Kim (Eds.): *The People's Web Meets NLP: Collaboratively Constructed Language Resources*, Chapter 5, pp. 121–160, Springer, April 2013.
- 2012 **Oliver Ferschte**, Iryna Gurevych and Marc Rittberger. 'FlawFinder: A Modular System for Predicting Quality Flaws in Wikipedia - Notebook for PAN at CLEF 2012', in: *CLEF 2012 Labs and Workshop, Notebook Papers*, Online Proceedings, Rome, Italy, September 2012.
- Johannes Daxenberger, **Oliver Ferschte** and Iryna Gurevych. 'Wikipedia-based Corpora for Analyzing Revisions, Discussions and Text Quality in Collaborative Writing', in: *Workshop on Automatic Processing of Non-Standard Data Sources in Corpus-Based Research (Extended Abstract)*, Cologne, August 2012.
- Oliver Ferschte**, Iryna Gurevych und Yevgen Chebotar. 'Behind the Article: Recognizing Dialog Acts in Wikipedia Talk Pages', in: *Proceedings of the 13th Conference of the European Chapter of the ACL (EACL 2012)*, pp. 777–786, Avignon, France, April 2012.
- 2011 **Oliver Ferschte**, Torsten Zesch and Iryna Gurevych. 'Wikipedia Revision Toolkit: Efficiently Accessing Wikipedia's Edit History', in: *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies. System Demonstrations*, pp. 97–102, Portland, OR, USA, June 2011.
- 2009 **Oliver Ferschte**. 'Semantic Relations in WordNet and the BNC', M.A. Thesis, Universität Würzburg, Januar 2009.