

# Alexander M. Fraser

## Curriculum Vitae



---

### Personal Data

Current Position	Professor for Data Analytics & Statistics, Technical University of Munich
Work Address	School of Computation, Information and Technology Bildungscampus 2, 74076 Heilbronn, Germany
	Phone +49 7131 264 18 912 (Sekretariat) Email <a href="mailto:contact102@alexanderfraser.de">contact102@alexanderfraser.de</a> Web <a href="http://alexfraser.github.io">alexfraser.github.io</a>
Citizenship	USA
Family Status	Married, no children
Languages	English (native)   German, French, Spanish (fluent)   Arabic (proficient)

---

### Research Interests

<b>Natural Language Processing</b>	machine translation, language models, cross-lingual transfer learning, parsing of syntax and discourse, morphology, information retrieval, evaluation
<b>Machine Learning</b>	representation learning, structured prediction, domain adaptation, semi-supervised learning, unsupervised learning, distributed algorithms for machine learning

---

### Education

12/2007	<b>PhD in Computer Science</b> University of Southern California, Los Angeles (CA), USA <ul style="list-style-type: none"><li>Thesis Title: Improved Word Alignments for Statistical Machine Translation</li><li>Advisor: Daniel Marcu</li><li>Committee: Eduard Hovy, Gareth James, Kevin Knight, Daniel Marcu, Paul Rosenbloom</li></ul>
12/2002	<b>Masters of Science in Computer Science</b> University of Southern California, Los Angeles (CA), USA <ul style="list-style-type: none"><li>GPA 3.86/4.00</li></ul>
06/1995	<b>BA in Computer Science and Economics (Concentration in Linguistics)</b> University of Massachusetts, Boston (MA), USA <ul style="list-style-type: none"><li>GPA 3.72/4.00</li></ul>

## Professional Experience

---

- 04/2024 – present **W3 Professor for Data Analytics & Statistics (Chair, tenured)**  
School of Computation, Information and Technology (CIT)  
Technical University of Munich, Germany
- 09/2017 – 03/2024 **W2 Professor of Information and Language Processing (~Full, tenured)**  
Center for Information and Language Processing (CIS), (tenure track 2017-2020)  
Ludwig-Maximilians-Universität München, Germany
- 03/2014 – 08/2014 **Visiting Professor for the W3 Professor of Linguistic Computer Science**  
Institute of Computational Linguistics (ICL),  
Ruprecht-Karls-Universität Heidelberg
- 04/2013 - 08/2017 **Akademischer Rat (~ Assistant professor, permanent)**  
Center for Information and Language Processing (CIS),  
Ludwig-Maximilians-Universität München, Germany
- 10/2012 – 09/2017 **Special Faculty Member**  
Department of Computer Science, University of Maryland,  
College Park (MD), USA
- 10/2009 – 04/2013 **Senior Researcher, Leader of Statistical Machine Translation Group**  
Institute for Natural Language Processing (IfNLP/IMS),  
University of Stuttgart, Germany
- 11/2007 – 09/2009 **Researcher**, Institute for Natural Language Processing (IfNLP/IMS),  
DFG Collaborative Research Centre 732: Incremental Specification in Context  
University of Stuttgart, Germany
- 09/2000 – 09/2007 **Graduate Research Assistant** (except dates listed below)  
University of Southern California, Information Science Institute (ISI),  
Marina del Rey (CA), USA
- 06/2003 – 05/2004 **Research Scientist**, Language Weaver, Marina del Rey (CA), USA
- 06 – 08/2002 **Research Scientist**, Department of Speech and Language Processing,  
05 – 08/2001 BBN Technologies, Cambridge (MA), USA  
02 – 08/2000
- 09/1995 – 01/1998 **Technical Director**, SatelLife, Boston (MA), USA (last position held)  
03/1994 – 12/1994
- 05/1987 – 01/1990 **Programmer / Analyst**, The Forum Corporation, Boston (MA), USA (last position held)

## Honors / Awards

- European Research Council Advanced Grant, Reserve List, current (2023 call)
- European Research Council Proof of Concept, Data for Multilingual Learning, 2022
- Best unsupervised Upper Sorbian / German machine translation, WMT 2020, unsupervised track
- Best English to German machine translation, WMT 2017 (news), constrained track
- ERC Starting Grant, Domain Adaptation for Statistical Machine Translation, 2014
- Best German to English machine translation, ACL-WMT 2009 constrained track
- 1000 Euro award for open source, distributed implementation of LEAF alignment model (from dissertation) at Euromatrix Second Machine Translation Marathon. 2009
- Full sponsorship for participation in Euromatrix Second Machine Translation Marathon (one week). 2009
- Word alignments used in 1st or 2nd best ranked Chinese to English and Arabic to English machine translation systems, NIST-MT 2006, 2008, 2009
- Best Romanian/English word alignment, ACL-WPT 2005 constrained track
- Best Arabic to English machine translation, NIST-MT 2003 constrained and unconstrained tracks
- Full sponsorship to spend 8 weeks in Baltimore at the Johns Hopkins University CLSP Summer Workshop 2003: Syntax for Statistical Machine Translation, led by Franz Och and Daniel Gildea. 2003
- SIGIR 2002 Student Travel Award. 2002
- Best English to Arabic cross-language retrieval, TREC 2001, TREC 2002
- Best Arabic information retrieval, TREC 2001
- University of Southern California - Charles Lee Powell Graduate Fellowship, full support and additional stipend for 3 years of PhD studies. 2000-2003
- University of Massachusetts, Boston - Graduated magna cum laude. 1995
- University of Massachusetts, Boston - Juan Carlos Merlo Prize, awarded annually to a single graduating mathematics or computer science student for outstanding achievement. 1995
- University of Massachusetts, Boston - Honors in Computer Science. 1995
- University of Massachusetts, Boston - Honors in Economics. 1995

## Grants

- Principal Investigator, Munich Center for Machine Learning (MCML), 2024-2025, 72000 Euros
- Principal Investigator (with Sabine Schulte im Walde, U Stuttgart), DFG Sachbeihilfe, "MeTRapher - Learning to Translate Metaphors", 3 years, 2024-2027, 319518 Euros (LMU)
- Principal Investigator, ERC Proof of Concept Grant, 18 months, 2023-2025, 150000 Euros
- Principal Investigator (with Frauke Kreuter), Deutsche Bundesbank, "Cooperation for explorative generation of structured data on sustainability", 3 years, 2023-2026, 397886 Euros

- Principal Investigator (with Viktor Hangya), DFG Sachbeihilfe, “Transfer Learning for Hate Speech Detection”, 3 years, 2023-2025, 336600 Euros
- Principal Investigator, LMUexcellent Investment Fund, 1 year, 2022-2023, 88000 Euros
- Principal Investigator, Munich Center for Machine Learning (MCML), 2022-2024, 144000 Euros
- Principal Investigator, Cambridge-LMU Strategic Partnership, 2021-2022, funding for multiple joint workshops and training seminars at Cambridge and at LMU for Prof. Anna Korhonen’s ERC project LEXICAL (Cambridge) and my ERC project DASMT
- Principal Investigator, Google Cloud Compute Credits for the FLORES 101 Large-Scale Multilingual Machine Translation Task at WMT, Meta, 2021, 4000 USD
- Principal Investigator (with Helmut Schmid), DFG Project (Sachbeihilfe), “Machine Translation of German without Parallel Corpora”, 2020, 581700 Euros
- Principal Investigator, Cambridge-LMU Strategic Partnership, 2019, funded two joint workshops at Cambridge and at LMU for my ERC project DASMT and Prof. Anna Korhonen’s ERC project LEXICAL (Cambridge)
- Principal Investigator, LMUexcellent Investment Fund, 1 year, 2018-2019, 50000 Euros
- Principal Investigator, ERC Starting Grant, 2015-2021, 5 years, 1.2 million Euros
- Principal Investigator, Horizon2020 Innovation Action “Health in my Language (HimL)” with U. Edinburgh, Charles U. Prague, 3 other organizations, 2015-2018, 500000 Euros
- Principal Investigator (with Knight, Koehn, Schmid, Uszkoreit) on Dagstuhl Seminar on Statistical Techniques for Translating to Morphologically Rich Languages, 2014, facilitated meeting of 60 researchers in Germany for one week
- Principal Investigator (with Schütze) on DFG Sachbeihilfe, “Models of Morphosyntax for Statistical Machine Translation - Phase II”, 2 years, 2013-2015, 300000 Euros
- Principal Investigator on Johns Hopkins Summer Workshop on Domain Adaptation for Machine Translation (with Daume, Carpuat, Quirk), 2012, facilitated research by 13 people for 6 weeks on site
- Deputy Investigator, European Commission FP7 STREP, “TTC: Terminology Extraction, Translation Tools and Comparable Corpora”. Coordinator Beatrice Daille, Nantes; Stuttgart PI Ulrich Heid, 2010-2012, 1 full-time position allocated to me for 1 1/2 years
- Principal Investigator (with Schütze) on DFG Sachbeihilfe, “Models of Morphosyntax for Statistical Machine Translation”, 3 years, 2009-2012, 517000 Euros
- Principal Investigator on Small Business Innovation Research grant to Language Weaver, “Linguistic Error Analysis of Statistical Machine Translation”. 50000 USD, 2004
- Principal Investigator (sole implementing PI) on grant for design, implementation and quality assurance of Language Weaver Arabic/English hybrid statistical machine translation system, 2 years, 2003-2004, amount confidential, approx. 5 full-time positions

## Teaching

At TUM, I teach in the Bachelors and Masters Programs in Information Engineering (at the multi-university Bildungscampus in Heilbronn). At LMU Munich, I was the coordinator of the Masters degree in Computational Linguistics. At LMU, I typically taught a standard full teaching load (9 SWS), but after becoming a professor I received a reduction (during the rest of my ERC Starting Grant, and later through German Excellence Funding). I also organized the institute reading group (together with Schütze and Plank), on deep learning approaches for natural language processing and machine translation.

- “Discrete Probability Theory”, TUM, Lecture and Exercise, **Every Summer Semester starting in 2024**
- “Seminar on Large Language Models”, TUM, Seminar, **Every Summer Semester starting in 2024**
- “Informationsextraktion”, LMU München, Lecture, **Every Winter Semester from 2014/2015 to 2023/2024**
- “Seminar zur Informationsextraktion (I)”, LMU München, Seminar/Exercise, **Every Winter Semester from 2014/2015 to 2023/2024**
- “Seminar zur Informationsextraktion (II)”, LMU München, Seminar/Exercise, **Every Winter Semester from 2014/2015 to 2023/2024**
- “Seminar on Large Language Models”, LMU München, Seminar, Winter Semester 2023/2024
- “Erweiterungsmodul: Machine Translation”, LMU München, Lecture and Exercise, Summer Semester 2023
- “Erweiterungsmodul: Machine Translation”, LMU München, Lecture and Exercise, Summer Semester 2021
- “Erweiterungsmodul: Machine Translation”, LMU München, Lecture and Exercise, Summer Semester 2020
- “Einführung in die Computerlinguistik”, with Robert Zangenfeind, LMU München, Lecture and Exercise, Winter Semester 2019/2020
- “Lexicon, Syntax, Semantics: Structured Prediction and Domain Adaptation”, LMU München, Seminar, Winter Semester 2017/2018
- “Erweiterungsmodul: Machine Translation”, LMU München, Lecture and Exercise, Summer Semester 2017
- “Computational Morphology and Electronic Dictionaries”, LMU München, Lecture and Exercise, Summer Semester 2017
- “Lexicon, Syntax, Semantics: Statistical Machine Translation and Neural Machine Translation”, LMU München, Seminar, Winter Semester 2016/2017
- “Seminar: Advanced Neural Machine Translation”, LMU München, Seminar, Summer Semester 2016
- “Computational Morphology and Electronic Dictionaries”, LMU München, Lecture and Exercise, Summer Semester 2016
- “Seminar: Neural Machine Translation”, LMU München, Seminar, Winter Semester 2015/2016
- “Lexicon, Syntax, Semantics: Word Sense Disambiguation and Statistical Machine Translation”, LMU München, Seminar, Winter Semester 2015/2016
- “Intensive Introduction to Neural Machine Translation”, LMU München, Seminar, Summer Semester 2015
- “Lexicon, Syntax, Semantics: Word Sense Disambiguation and Statistical Machine Translation”, LMU München, Seminar, Winter Semester 2014/2015
- “Information Extraction”, DAAD International Summer School in Advanced Language Engineering, University of Colombo, Sri Lanka, September 2014
- “Programmieren II: Fortgeschrittene (Java)”, Universität Heidelberg, Lecture and Exercise, Summer Semester 2014

- “Statistical Machine Translation”, Universität Heidelberg, Lecture, Summer Semester 2014
- “Linguistic Knowledge for Statistical Machine Translation”, Universität Heidelberg, Proseminar, Summer Semester 2014
- “Informationsextraktion”, LMU München, Lecture, Winter Semester 2013/2014
- “Seminar zur Informationsextraktion”, LMU München, Seminar, Winter Semester 2013/2014
- “Statistical Machine Translation”, DAAD Summer School in Advanced Language Engineering, Kathmandu University, Nepal, September 2012
- “Statistical Machine Translation”, with Andreas Maletti, Seminar (Hauptseminar), University of Stuttgart, Winter Semester 2011/2012
- “Machine Learning for Natural Language Processing” Seminar (Hauptseminar), with Stefan Riezler and Hinrich Schütze, Joint Seminar: Heidelberg University and University of Stuttgart, Winter Semester 2010/2011
- “Statistical Machine Translation”, with Helmut Schmid, Seminar (Lektürekurs), University of Stuttgart, Summer Semester 2010
- “Parsing and Machine Translation”, with Helmut Schmid, Seminar (Hauptseminar), University of Stuttgart, Summer Semester 2009
- “Data-driven and Hybrid Approaches to Machine Translation”, with Martin Forst, intensive two week course at German Society for Linguistics Fall School in Konstanz, 2009
- One week intensive course on Statistical Machine Translation at 9th Summer School of the European Masters in Language and Speech Technology Programme, Stuttgart, 2008

## Advising

### Post-docs

- Dr. Daryna Dementieva, 06/2024-, DFG, TUM; research on dealing with toxic content
- Dr. Shu Okabe, 05/2024-, ERC, TUM; research on data for multilingual learning and multilingual embeddings
- Dr. Lukas Edman, 02/2024-, DFG Unsup, MCML; research on character-level large language models
- Dr. Marion Weller-Di Marco, 09/2019-, LMUExcellent, DFG Unsup, research on morphological generalization in language models and translation
- Dr. Viktor Hangya, 01/2017-06/2024, MCML, ERC; research on multilingual embeddings for domain adaptation
- Dr. Jindřich Libovický, 09/2019-12/2021, research on character-based machine translation and multilingual representations
- Dr. Matthias Huck, 04/2016 to 03/2020, research on domain adaptation and translation to morphologically rich languages
- Dr. Fabienne Braune, 01/2016 to 12/2017, research on multilingual embeddings for mining comparable corpora
- Dr. Aleš Tamchyna, 01/2017 to 04/2017, research on translation to morphologically rich languages (jointly supervised with Prof. Jan Hajic, Charles University Prague)

- Dr. Tsuyoshi Okita, 01/2016 to 12/2016, research on neural models for automatic post-editing
- Dr. Liane Guillou, 04/2016 to 09/2016, research on discourse for translation

#### Ph.D. Advising

- Sarah Ball (with Frauke Kreuter). 11/2023-, Fellow in Konrad Zuse School of Excellence in Reliable AI; Moral Reasoning in Large Language Models
- Faeze Ghorbanpour (with Viktor Hangya). 09/2023-, DFG Transfer Learning for Hate Speech Detection; Hate Speech Detection, Transfer Learning
- Irina Bigoulaeva (with Iryna Gurevych, TU Darmstadt). 10/2021-, TU Darmstadt funding; Transfer Learning, Hate Speech Detection
- Katharina Hämmerl (with Jindrich Libovicky, Charles U Prague). 07/2021-, MCML, DFG Unsup; Multilingual Representations
- Prisca Piccirilli (with Sabine Schulte im Walde, U Stuttgart). 07/2021-, Studienstiftung Scholarship, Joint Stuttgart and LMU DFG funding; Metaphor Translation
- Wen (Lavine) Lai. 04/2021-, LMU/CSC; Low Resource NMT, Domain Adaptation
- Sophie Henning 01/2021-, Bosch, with Annemarie Friedrich, research on calibration
- Alexandra Chronopoulou, 09/2019-, MCML, ERC, DFG Unsup, research on transfer and multilingual learning
- Dr. rer. nat. Dario Stojanovski, 06/2017-06/2021, Modeling Contextual Information in Neural Machine Translation. First permanent position after graduation: Microsoft.
- Dr. phil. Anita Ramm (née Gojun) (with Sabine Schulte im Walde), 05/2013 to 03/2018, Modeling the position and inflection of verbs in English to German machine translation. First position after graduation: BeOne Stuttgart GmbH.
- Dr. phil. Marion Weller-Di Marco (with Sabine Schulte im Walde), 12/2011 to 09/2017, Using Morpho-Syntactic and Semantic Information to Improve Statistical Machine Translation. First position after graduation: Post-doc at University of Amsterdam with Khalil Sima'an.
- Dr. rer. nat. Fabienne Braune (with Andreas Maletti), 10/2009-11/2015, Decoding Strategies for Syntax-based Statistical Machine Translation. Winner of the **European Association for Machine Translation Best Thesis Award** for the best machine translation dissertation in Europe in 2015. First position after graduation: Post-doc at LMU Munich with Alexander Fraser.
- Dr. phil. Fabienne Cap (née Fritzinger), 10/2009-10/2014, Morphological Processing of Compounds for Statistical Machine Translation. First position after graduation: Post-doc (VINNMER Marie Curie Incoming) at Uppsala University with Joakim Nivre.
- Dr. rer. nat. Nadir Durrani (with Helmut Schmid), 04/2008-11/2012, A Joint Translation Model with Integrated Reordering. Winner of the **German Society for Computational Linguistics award** for the best natural language processing / computational linguistics dissertation for the years 2011 to 2013. First position after graduation: Post-doc at University of Edinburgh with Philipp Koehn.
- Dr. rer. nat. Hassan Sajjad (with Helmut Schmid), 04/2008-10/2012, Statistical Models for Unsupervised, Semi-supervised and Supervised Transliteration Mining. First position after graduation: Qatar Computing Research Institute. Currently: Associate Professor, Dalhousie University, Nova Scotia, Canada

#### Visiting Ph.D. Students

- Denis Peskov, University of Maryland, College Park, 01/2020 to 10/2020

- Simon Rieß, LMU Data Science, 10/2019 to 12/2019
- Viktor Hangya, University of Szeged, 10/2016 to 12/2016
- Aleš Tamchyna, Charles University Prague, 10/2015 to 05/2016
- Marija Brkic, University of Zagreb, Croatia. 09/2010 to 02/2011

### External Ph.D. Committees

- Ph.D. Committee Member: David Adelani. Saarland University, Germany. 06/2023 (Dietrich Klakow)
- Ph.D. Committee Member: Patrick Schramowski. TU Darmstadt, Germany. 03/2023 (Kristian Kersting)
- Ph.D. Committee Member: Gongbo Tan. Uppsala University, Sweden. 12/2020 (Joakim Nivre)
- Ph.D. Committee Member: Andreas Guta. RWTH Aachen, Germany. 09/2020 (Hermann Ney)
- Ph.D. Committee Member: Minh-Quang Pham. LIMSI, Paris, France. 06/2020 (Francois Yvon)
- Ph.D. Reviewer: Duygu Ataman. University of Trento, Italy. 09/2019 (Marcello Federico)
- Ph.D. Reviewer: Ekaterina Vylomova. University of Melbourne, Australia. 11/2018 (Tim Baldwin)
- Ph.D. Committee Member: Matthias Huck. RWTH Aachen, Germany. 08/2018 (Hermann Ney)
- Ph.D. Committee Member: Peyman Passban. Dublin City University, Ireland. 09/2017 (Andy Way)
- Ph.D. Committee Member: Arianna Bisazza. University of Trento, Italy. 04/2013 (Marcello Federico)
- Ph.D. Committee Member: Jagadeesh Jagarlamudi. University of Maryland, College Park, MD, USA. 04/2013 (Hal Daume)
- Examiner, “Licentiate of Philosophy” (similar to American Ph.D. qualification exam): Sara Stymne, Department of Computer and Information Science, Linköping University, Sweden. 12/2009

### Masters/Bachelors Advising

I have directly advised 39 Bachelors level and 29 Masters level students myself, mostly at LMU Munich and the University of Stuttgart, but also at University of Ljubljana and Université de Fribourg. I have acted as Prüfer and/or second supervisor (mostly together with a PhD or PostDoc supervised by me) for many more. Marion Weller’s Bachelors Thesis (Studienarbeit) in Computational Linguistics, *Separate Morphologiebehandlung als Methode zur Verbesserung statistischer maschineller Übersetzung* (Separate Morphological Processing as a Method for Improving Statistical Machine Translation), completed 05/2009, was the winner of **Informatik-Verbund-Stuttgart Award** (given yearly for the two best Bachelors Theses in computer science-related fields in Stuttgart). A good number of the student works I have supervised have resulted in joint conference publications which passed peer review, see my publications list.

### Invited Presentations

- Invited Talk, Tübingen AI Center (University of Tübingen, Max Planck IIS), Germany, 2023
- Invited Talk, St. Mary’s University, San Antonio, Texas, USA, 2023
- Invited Talk, University of Kassel, Germany, 2023



- Invited Talk, University of Münster, Germany, 2022
- Invited Talk, Bosch Artificial Intelligence, Germany, 2022
- Keynote, “Bilingual Induction and Pseudo Parallel Corpora”, Workshop on Building and Using Comparable Corpora (BUCC), Marseille, France, 2022
- Invited Talk, Technical University of Munich, Germany, 2022
- Invited Talk, German Society for Computational Linguistics and Language Technology, 2021
- Invited Talk, University of Cambridge, UK, 2021
- Invited Talk, University of Bamberg, Germany, 2020
- Invited Talk, Uppsala University, Sweden, 2020
- Invited Talk, Karlsruhe Institute of Technology, Germany, 2020
- Invited Talk, TU Wien, Austria, 2020
- Invited Talk, University of Regensburg, Germany, 2020
- Invited Talk, University of Vienna, Austria, 2019
- Panel, “Open Problems in Machine Translation”, ACL Conference on Machine Translation, Florence, Italy, 2019
- Invited Talk, TU Wien, Austria, 2019
- Invited Talk, University of Potsdam (Digital Engineering), Germany, 2019
- Keynote, “Translation to Morphologically Rich Languages”, Baltic Human Language Technologies, University of Tartu, Estonia, 2018
- Invited Talk, Heidelberg University, Germany, 2018
- Invited Talk, Dublin City University, Ireland, 2017
- Invited Talk, University of Potsdam (Linguistics), Germany, 2017
- Keynote, “Translation to Morphologically Rich Languages”, European Association for Machine Translation (EAMT 2016), Riga, Latvia, 2016
- Invited Talk, University of Hamburg, Germany, 2015
- Invited Talk, Universität Zürich, Switzerland, 2014
- Invited Talk, Machine Translation Marathon, Prague, Czech Republic, 2013
- Keynote, “Improving Translation to Morphologically Rich Languages”, La 20ème conférence du Traitement Automatique du Langage Naturel (TALN), Sables d’Olonnes, France, 2013
- Invited Talk, Karlsruher Institut für Technologie, Germany, 2013
- Invited Talk, Institute for Natural Language Processing, University of Stuttgart, Germany, 2012
- Invited Talk, Ubiquitous Knowledge Processing Lab, TU Darmstadt, Germany, 2012
- Panel, “Arabic Natural Language Processing Technologies Round-Table”, Qatar Computing Research Institute, Doha, Qatar, 2011
- Invited Talk, Institute for Language, Cognition and Computation Seminar, University of Edinburgh, 2010
- Keynote, “Morphosyntactic Correspondences and a Progress Report on Bitext Parsing”, InFuture 2009, Zagreb, Croatia, 2009

- Invited Talk, German Society for Computational Linguistics & Language Technology Workshop “Costs and Uses of Machine Translation”, 2009
- Invited Talk, Research Track of Euromatrix Second Machine Translation Marathon, 2008
- Invited Talk, Institute for Computational Linguistics, Heidelberg University, 2008
- Invited Talk, Linguistics Department, University of Potsdam, 2007
- Invited Talk, Language Technology Laboratory, DFKI (German Research Center for Artificial Intelligence), 2006
- Invited Talk, Institute for Natural Language Processing (IMS), University of Stuttgart, 2006
- Panel, “Issues in Arabic to English Machine Translation”, MT Summit IX Workshop on Machine Translation for Semitic Languages, 2003
- Invited Talk, US National Institute of Standards Machine Translation Evaluation Workshop, 2003

## **Professional Activities**

### **Senior PC**

- NAACL 2024: Senior Area Chair in Machine Translation
- EMNLP 2023: Senior Area Chair in Machine Translation
- EACL 2023: Senior Area Chair in Multilinguality
- NAACL 2021: Area Chair in Machine Translation and Multilinguality
- EACL 2021: Area Chair in Machine Translation
- ACL 2020: Area Chair in Natural Language Processing Applications
- ACL 2019: Area Chair in Machine Translation
- EAMT 2017: Program Chair (research track)
- EMNLP 2016: (Senior) Area Co-Chair: Machine Translation
- NAACL 2016: (Senior) Area Co-Chair: Multilinguality and Morphology
- EACL 2014: (Senior) Area Co-Chair: Machine Translation
- ACL 2014: Tutorials Co-Chair

### **Journal Reviewer**

- Computational Linguistics (CL)
- Transactions of the Association for Computational Linguistics (TACL)
- Journal of Artificial Intelligence Research (JAIR)
- Artificial Intelligence (AI)
- Journal of Machine Learning Research (JMLR)

### **Conference/Workshop Reviewer**

- Regular reviewer at ACL, NAACL, EACL, EMNLP
- Regular reviewer at Conference on Machine Translation
- Regular reviewer at ACL/NAACL/EACL Workshops on Machine Translation
- NIPS 2010-2011
- COLING 2008, COLING 2010
- MT Summit 2009
- IJCNLP 2008
- Theoretical and Methodological Issues in Machine Translation (TMI) 2004
- Selected Workshop reviews: ACL Workshop on Building and Using Parallel Texts; ACL Workshop on Syntax, Semantics and Structure in Statistical Translation; EACL/ACL Workshops on Computational Approaches to Semitic Languages; Workshop on Hybrid Approaches to Machine Translation; IWSLT International Workshop on Spoken Language Translation; COLING Workshop on Human Judgements in Computational Linguistics; Analyzing and interpreting neural networks for NLP; MT Summit IX Workshop on Machine Translation for Semitic Languages

### **Other Reviewing**

- European Research Council - Proposal Reviewer
- German Research Foundation (DFG) - Proposal Reviewer
- German Academic Exchange Service (DAAD) - Proposal Reviewer
- U.S. National Science Foundation - Proposal Reviewer
- Neural Information Processing Series, "Learning Machine Translation", MIT Press
- Grace Hopper Celebration of Women in Computing, Anita Borg Institute

### **Other Activities**

- Designed and managed creation of the first commercially available statistical machine translation system (Language Weaver Arabic/English), available almost two years before Google Translate
- Software provided to the research community:
  - Implementation of a Discriminative Lexicon in Moses, with Kristina Toutanova and Ondrej Bojar and team at Open Source track of the Eighth Machine Translation Marathon 2013
  - Integration of Vowpal Wabbit, a high-speed learner of classification models, as a lexicon model into the open source Moses statistical machine translation engine, with team at the Johns Hopkins Summer Workshop on Domain Adaptation for Machine Translation 2012
  - Cube pruning algorithm implemented in open source Moses statistical machine translation engine, with team at Open Source track of Euromatrix Second Machine Translation Marathon 2009
  - Word alignment graphical user interface implemented in Java Swing, used by 7 research labs, 2007
- Data provided to the research community:
  - Standard data sets for domain-adaptation experimentation in machine translation created at the Johns Hopkins Summer Workshop on Domain Adaptation for Machine Translation

- Probabilistic Dictionary (with Jinxi Xu) for use in TREC-2002 Cross-Lingual Information Retrieval Task (using English queries to request Arabic documents)
  - BBN release of the Arabic/English UN corpus (with Jinxi Xu, Ralph Weischedel), distributed by the Linguistic Data Consortium
  - Sentence aligned the Ummah corpus, distributed by the Linguistic Data Consortium
- Organizing of Stuttgart Reading Group on Parsing and Machine Translation
- Stuttgart Coordinator for Pascal2 EU Network of Excellence (Pascal2: Pattern Analysis, Statistical Modelling and Computational Learning)