

Curriculum Vitae

BEHRANG QASEMIZADEH

Email: qasemizadeh@gmail.com
Web: <https://data-science-expert.de>
LinkedIn: www.linkedin.com/in/qasemizadeh

Professional Summary

Senior Data Scientist and researcher with over 15 years of experience in statistical machine learning, probabilistic modelling, Bayesian methods, natural language processing, and time-series analysis. Expertise in developing robust methods for complex data, with contributions spanning algorithm development, representation learning, unsupervised learning, and evaluation methodologies. Established record of peer-reviewed publications, widely used datasets and language resources, international research collaborations, and contributions to grant-funded projects. Experience spans independent academic research and the development of applied AI systems in industrial R&D, bridging methodological innovation and real-world applications.

Research Areas

- **Statistical Machine Learning:** Statistical inference, probabilistic modelling, Bayesian inference, Bayesian optimisation, and uncertainty-aware learning.
- **Data-Centric AI:** Data-centric machine learning for complex structured, multimodal, temporal, and sensor data, with emphasis on robust evaluation, scientific machine learning, and interdisciplinary applications.
- **Representation Learning and Structured Prediction:** Semantic representation learning, self-supervised and unsupervised learning, structured prediction, and information extraction.

Education

- **Ph.D. in Computer Science**, 2015, [National University of Ireland, Galway](#).
- **M.Sc. in Artificial Intelligence and Robotics**, 2006, [Iran University of Science and Technology \(IUST\)](#).
- **B.Eng. in Computer Engineering (Software)**, 2003, [Tehran Azad University](#).

Skills

Technical Expertise

Python: Machine learning and probabilistic modelling (scikit-learn, XGBoost, GPyTorch, Prophet, PyMC), deep learning (PyTorch, Keras), NLP and LLMs (NLTK, spaCy, Hugging Face Transformers), optimization (BoTorch, Z3), scientific computing & data processing (NumPy, SciPy, Pandas, Dask), web apps (FastAPI, Flask, Dash), databases (PostgreSQL, MySQL, SQLite, DuckDB, MongoDB, Redis, TimescaleDB), and data visualization (Plotly, Matplotlib, Seaborn).

Java: Object-oriented software engineering for machine learning and natural language processing systems (Stanford CoreNLP, GATE, Apache OpenNLP, MALLET).

Prolog: ISO Prolog and LPA Win-Prolog environments (Chimera Agents, TCP/IP Toolkit, Intelligence Server).

Semantic Web & Knowledge Graphs: RDF, OWL, XML, Apache Jena.

Data Engineering, MLOps & DevOps: Databricks, Apache Airflow, Git, GitLab, Azure DevOps, MLflow, Docker.

Languages

Farsi: Native; **English:** Fluent; **German:** A2.

Soft Skills

Research leadership • interdisciplinary collaboration • mentoring and supervision • scientific communication • project coordination.

Experience

February 2020 – Present

Senior Data Scientist, Henkel – Center of Excellence for Data Science

Lead the design, development, and deployment of machine learning solutions across R&D, manufacturing, finance, and supply chain. Collaborate with interdisciplinary stakeholders to identify research opportunities, define AI solution strategies, and deliver production-grade machine learning systems.

Selected Achievements:

- Led the development of *Unravel*, a Bayesian optimization proof-of-concept platform for chemical product formulation and design of experiments. *Unravel* evolved into a multi-year strategic project with substantial investment from the Adhesive Technologies business unit. Continued as a core technical contributor, designing and implementing the project's core Bayesian optimization algorithms.
- Developed probabilistic forecasting models for manufacturing process optimization using production sensor time-series data to predict production line speed and subprocess stoppage risk, enabling proactive production control and optimization of chemical dosing.
- Developed a predictive machine learning web application that integrated heterogeneous R&D datasets into a unified modelling pipeline to address data sparsity and support product formulation. The system enables chemists to leverage historical data from previously developed products across different product families, facilitating knowledge transfer during formulation design. As a result, it reduced formulation development time and effort by improving reuse of historical R&D data.
- Developed rolling forecasting and Last Best Estimate (LBE) models to support financial planning and resource allocation, and built an interactive dashboard enabling controllers to review, compare, and adjust forecasts from multiple statistical and machine learning models.
- Developed machine learning models for customer credit risk assessment on highly imbalanced datasets, enabling proactive risk mitigation and supporting financial decision-making.
- Developed a predictive machine learning web application to estimate event durations in SAP-based enterprise resource planning workflows within the supply chain, enabling identification of process bottlenecks and improving time-to-market KPIs.
- Designed and implemented a Lucene-based enterprise search platform for hazard sheets and technical documentation, improving information retrieval for R&D teams.

January 2016 – February 2020:

Postdoctoral Researcher, [DFG Collaborative Research Centre 991](#), University of Düsseldorf.

Conducted independent research in statistical machine learning and natural language processing, developing probabilistic and unsupervised methods for representation learning and structured prediction within a multidisciplinary DFG Collaborative Research Centre. Contributed to research management through grant writing, project coordination, and scientific reporting.

Selected Research Contributions:

- Developed hierarchical Bayesian models for unsupervised semantic frame induction within a multidisciplinary research project on natural language understanding.
- Published 15 peer-reviewed papers in leading NLP and machine learning venues, and developed open-source software, benchmark annotated datasets, and research infrastructure for semantic representation learning, multiword expression identification, and relation extraction.
- Organized three shared tasks at major international NLP conferences, coordinating datasets, evaluation protocols, and international community participation.
- Delivered graduate-level lectures on Semantic Role Labeling and supervised two master's students, resulting in successful thesis completions and peer-reviewed publications.
- Collaborated with international researchers through the PARSEME COST Action (IC1207) and contributed to the successful research grant proposal for the follow-up project *Frame Induction: Event Type Hierarchies and Complex Event Types (FInd)*.

October 2014 – November 2015:

Wissenschaftlicher Mitarbeiter (Scientific Assistant), University of Passau.

Conducted independent research in statistical machine learning and large-scale information extraction, with a particular focus on representation learning and terminology extraction using randomized projection methods. Contributed to grant writing, project reporting, and graduate teaching.

Selected Research Contributions:

- Published 4 peer-reviewed papers, contributed to research seminars and organized colloquia.
- Secured an NVIDIA CUDA Research Center equipment grant (4 Tesla GPUs) and contributed to EU Horizon 2020 research proposals.
- Designed and delivered an 8 ECTS Text Mining course, including lectures, project supervision, and student mentoring.

April 2009 – October 2014:

PhD Candidate and Research Assistant,¹ INSIGHT Centre for Data Analytics, National University of Ireland, Galway.

Selected Research Contributions:

- Completed a Ph.D. in Computer Science on corpus-based and distributional approaches to natural language semantics, supported by a scholarship within the *Lion* and *Lion 2* projects (Science Foundation Ireland: SFI/02/CE1/I131, 12/RC/2289).
- Published 10+ peer-reviewed papers at leading NLP and computational linguistics venues; developed the *ACL RD-TEC* benchmark dataset, now widely used for evaluating terminology extraction and classification methods.
- Developed scalable statistical learning methods for terminology extraction, semantic modelling, and scientific literature mining, resulting in reusable open-source software supporting research in information extraction and text analytics.
- Collaborated with *Cisco Systems* on knowledge-transfer activities by developing domain-adaptive topic extraction algorithms for enterprise chat logs and email communications.
- Awarded a competitive Marie Curie Actions travel grant to undertake a 10-week research visit at the Apex Data & Knowledge Management Lab, Shanghai Jiao Tong University, fostering international research collaboration.

December 2012 – June 2013:

Remote Technical Consultant (part-time), [Relative Insight](#).

Selected Achievements:

- Developed a production-ready forensic text analysis software package for automated gender and age profiling from online forum data.
- Designed and implemented machine learning models using stylometric and linguistic features to infer demographic characteristics from textual data.

December 2007 – April 2009:

Research Assistant, [Institute of Electronics, Communications and Information Technology](#), Queen's University, Belfast. Conducted research and development on natural language dialogue systems and multimodal information retrieval for surveillance video analysis as part of the EPSRC-funded ISIS project.

Selected Research Contributions:

- Developed a Prolog-based distributed framework integrating heterogeneous software components (including C++ and Java) for natural language-driven surveillance video retrieval.
- Designed semantic models and dialogue management components linking natural language queries to high-level event representations extracted from surveillance video.
- Developed methods for high-level event detection, bidirectional mapping between detected events and natural language descriptions, and domain-specific dialogue strategies for interactive video retrieval.
- Published 4 peer-reviewed papers and contributed to interdisciplinary collaboration across the EPSRC-funded project.

September 2006 – December 2007:

Co-founder and NLP Lead, Text and Speech Technology Ltd., Tehran, Iran.

Led NLP research and development, technical consulting, project acquisition for commercial and publicly funded projects.

Selected Achievements:

- Secured research funding from Iran Telecommunication Research Center (ITRC) for the development of language resources supporting Farsi–English machine translation.
- Developed a Persian text-to-speech system and integrated it with the Microsoft Speech API for compatibility with the JAWS screen reader.
- Led the integration of Persian into the MULTTEXT-East framework by defining morphosyntactic specifications and manually annotating the Persian translation of Orwell's *1984*.
- Developed foundational Persian NLP resources, including a lexicon of Persian stems published through ELRA (Catalogue Reference: ELRA-L0087).
- Published 2 peer-reviewed papers and released software tools for processing Persian electronic text.
- Provided external consultancy on Persian language technologies for industrial and research organizations.

2005–2006 (Multiple Engagements):

Consultant, Lexical Computing Ltd. (Sketch Engine), UK.

Provided consultancy on Persian language resources, supporting the development of statistical Farsi–English machine translation and large-scale Persian web corpora.

Selected Achievements:

- Provided language resources and taggers for statistical Farsi–English machine translation.
- Advised on linguistic and language resource issues during the development of a Persian Web Corpus.

August 2005 – April 2006

Master’s Thesis Grant, [Iran Telecommunication Research Center \(ITRC\)](#) & Iran University of Science and Technology.

R&D in the area of Video Semantics and Retrieval.

Selected Research Contributions:

- Produced two peer-reviewed research publications based on the project outcomes.
- Developed a functional C# prototype for video event extraction and timeline visualization.

September 2004 – September 2006:

Localization Team Leader, Digital Clone (Shegarf), Tehran, Iran.

Led the development of Persian language technologies and software localisation.

Selected Achievements:

- Led a multidisciplinary team of linguists and engineers in the development of Persian text-to-speech, voice command recognition, and spell-checking technologies.
- Designed and developed adaptive spell-checking algorithms.
- Developed phonemic transliteration methods for Persian text-to-speech, including stress, duration, and pitch prediction.
- Designed and developed core linguistic resources, including a Persian morphological analyser and part-of-speech tagger, supporting the company’s language technology products.
- Published 4 peer-reviewed papers.

March 2004 – August 2004:

System Analyst and Developer (part-time), [Artificial Intelligence Rayvarz](#), Tehran, Iran.

Conducted R&D and product development for Persian natural language processing.

Selected Achievements:

- Developed the first commercial Persian text-to-speech system.
- Improved the accuracy of the company’s optical character recognition (OCR) engine.

Research Funding and Awards

- **Full Ph.D. Scholarship**, Science Foundation Ireland, 2009–2014.
- **NVIDIA Academic Grant**, University of Dusseldorf, 2019.
- **NVIDIA CUDA Research Center equipment grant**, University of Passau, 2015.
- **Marie Skłodowska-Curie Research Visit Grant**, IRSES Programme, Grant No. 24761 (Net2).
- **Master’s Thesis Research Grant**, Iran Research Institute for Information and Communication Technology (ITRC).
- **Outstanding Academic Performance Award**, Computer Engineering Department, IUST, 2003–2004.
- **Top 0.1% Ranking**, Iranian Nationwide Graduate Entrance Examination (aka. Konkour) in Computer Science, 2003–2004.

Other Professional Services

- Task Organizer for three international shared tasks and community evaluation campaigns at SemEval and PARSEME, coordinating benchmark datasets, evaluation protocols, and international research participation.
- Contributed to the PARSEME COST Action (IC1207), WP3: PARSing and Multi-word Expressions.
- Reviewer for journals including IEEE Transactions on Systems, Man, and Cybernetics, IEEE Access, Natural Language Engineering, and Data & Knowledge Engineering.
- Reviewer for international conferences, including ACL, EACL, WWW, ISWC, ESWC, EKAW, LREC, NLDB,

Additional Studies

- The ENeL Training School 2016: Tools and Methods for Creating Innovative e-Dictionaries, University of Ljubljana, 17–20 May 2016.
- Research visiting scholar at the **APEX Data & Knowledge Management Lab, Shanghai Jiao Tong University**, 20 April 2011 – 27 June 2011.
- The 17th European Summer School on Logic, Language and Information (ESSLLI 2005), Edinburgh, UK, August 2005.
- Jadertina Summer School in Empirical and Computational Linguistics (JSSECL), University of Zadar, September 2006.
- The 3rd DGfS Computational Linguistics Fall School, Bochum, Germany, September 2005.

Published Research Resources

- **PARSEME Shared Task Data (Farsi, v1.0)**: Multilingual corpus annotated with verbal multiword expressions. Distributed via LINDAT/CLARIN Digital Library, Charles University in Prague. <http://hdl.handle.net/11372/LRT-2282>.
- **ACL RD-TEC 2.0**: Annotated corpus for term and entity recognition evaluation (co-authored with Schumann, Anne-Kathrin). LINDAT/CLARIN. <http://hdl.handle.net/11372/LRT-1661>. ISLRN: 178-746-003-351-0.
- **ACL RD-TEC**: Terminological resource for computational linguistics. ELRA Catalogue ELRA-T0375. ISLRN: 699-305-089-6.
- **Persian Lexicon**: Lexicon of 40,000+ non-inflected Farsi words with semantic, morphological, and phonemic annotations. ELRA ELRA-L0087. ISLRN: 547-614-436-004-7.
- **FASpell**: Dataset for evaluation of Persian spell-checking systems. LINDAT/CLARIN. <http://hdl.handle.net/11372/LRT-1547>.
- **MULTEXT-East Persian Lexicon (v4)**: Inflected lexicon (13,427 entries) with morphosyntactic annotation and phonemic transliteration. ELRA ELRA-L0086. ISLRN: 884-966-712-343-0.
- **Persian “1984” Annotated Corpus**: Manually annotated Persian translation of Orwell’s *1984* with morphosyntactic tags and phonemic transliteration (MULTEXT-East compliant). ELRA ELRA-W0054. ISLRN: 851-240-629-673-1. Additional MULTEXT-East distributions (CLARIN.SI):
 MULTEXT-East “1984” corpus v4: <http://hdl.handle.net/11356/1043>
 MULTEXT-East non-commercial lexicons v4: <http://hdl.handle.net/11356/1042>

Filed Patents

- Alistair Baron, Behrang QasemiZadeh, Paul Rayson, James Walkerdine, Phil Greenwood and Awais Rashid. “Text analysis”. *WO Patent App. PCT/GB2013/052,806*. Date Filed: Oct 28, 2013, URL: <https://patentscope.wipo.int/search/en/WO2014068293>.

List of Publications

- [1] Behrang QasemiZadeh, Miriam R. L. Petruck, Regina Stodden, Laura Kallmeyer, and Marie Candito. “SemEval-2019 Task 2: Unsupervised Lexical Frame Induction”. In: *Proceedings of the 13th International Workshop on Semantic Evaluation (SemEval-2019)*. 2019, pp. 16–30.
- [2] Kata Gabor, Davide Buscaldi, Anne-Kathrin Schumann, Behrang QasemiZadeh, Haifa Zargayouna, and Thierry Charnois. “SemEval-2018 Task 7: Semantic Relation Extraction and Classification in Scientific Papers”. In: *Proceedings of The 12th International Workshop on Semantic Evaluation*. New Orleans, Louisiana: Association for Computational Linguistics, 2018, pp. 679–688. URL: <http://aclweb.org/anthology/S18-1111>.
- [3] Laura Kallmeyer, Behrang QasemiZadeh, and Jackie Chi Kit Cheung. “Coarse Lexical Frame Acquisition at the Syntax–Semantics Interface Using a Latent-Variable PCFG Model”. In: *Proceedings of the Seventh Joint Conference on Lexical and Computational Semantics*. New Orleans, Louisiana: Association for Computational Linguistics, 2018, pp. 130–141. URL: <http://aclweb.org/anthology/S18-2016>.
- [4] Natalia Klyueva, Anna Vernerova, and Behrang Qasemizadeh. “Querying Multi-word Expressions Annotation with CQL”. In: *Proceedings of the 16th International Workshop on Treebanks and Linguistic Theories*. 2018. URL: <http://aclweb.org/anthology/W/W17/W17-7611.pdf>.
- [5] Alfredo Maldonado and Behrang QasemiZadeh. “Analysis and Insights from the PARSEME Shared Task dataset”. In: *Multiword expressions at length and in depth: Extended papers from the MWE 2017 workshop*. Ed. by Agata Savary Stella Markantonatou Carlos Ramisch and Veronika Vincze. Phraseology and Multiword Expressions 2. ISSN: 2625-3127. Berlin: Language Science Press, 2018. Chap. 5, pp. 149–175. ISBN: 978-3-96110-124-5. DOI: [DOI:10.5281/zenodo.1469557](https://doi.org/10.5281/zenodo.1469557).
- [6] Behzad Naderalvjooud, Behrang QasemiZadeh, Laura Kallmeyer, and Ebru Akcapinar Sezer. “A Cross-Lingual Approach for Building Multilingual Sentiment Lexicons”. In: *Text, Speech, and Dialogue - 21st International Conference, TSD 2018, Brno, Czech Republic, September 11-14, 2018, Proceedings*. 2018, pp. 259–266.

- [7] Carlos Ramisch, Silvio Ricardo Cordeiro, Agata Savary, Veronika Vincze, Verginica Barbu Mititelu, Archana Bhatia, Maja Buljan, Marie Candito, Polona Gantar, Voula Giouli, Tunga Güngör, Abdelati Hawwari, Uxoá Iñurrieta, Jolanta Kovalevskaitė, Simon Krek, Timm Lichte, Chaya Liebeskind, Johanna Monti, Carla Parra Escartín, Behrang QasemiZadeh, Renata Ramisch, Nathan Schneider, Ivelina Stoyanova, Ashwini Vaidya, and Abigail Walsh. “Edition 1.1 of the PARSEME Shared Task on Automatic Identification of Verbal Multiword Expressions”. In: *Proceedings of the Joint Workshop on Linguistic Annotation, Multiword Expressions and Constructions (LAW-MWE-CxG-2018)*. Santa Fe, New Mexico, USA: Association for Computational Linguistics, Aug. 2018, pp. 222–240. URL: <http://www.aclweb.org/anthology/W18-4925>.
- [8] Agata Savary, Verginica Barbu Mititelu Marie Candito, Fabienne Cap Eduard Bejček, Silvio Ricardo Cordeiro Slavomír Čéplö, Voula Giouli Gülşen Eryiğit, Yaakov HaCohen-Kerner Maarten van Gompel, Simon Krek Jolanta Kovalevskaitė, Johanna Monti Chaya Liebeskind, Lonneke van der Plas Carla Parra Escartín, Carlos Ramisch Behrang QasemiZadeh, and Ivelina Stoyanova & Veronika Vincze. 2018 Federico Sangati. “PARSEME multilingual corpus of verbal multiword expressions”. In: *Multiword expressions at length and in depth: Extended papers from the MWE 2017 workshop*. Ed. by Agata Savary Stella Markantonatou Carlos Ramisch and Veronika Vincze. Phraseology and Multiword Expressions 2. ISSN: 2625-3127. Berlin: Language Science Press, 2018. Chap. 4, pp. 87–148. ISBN: 978-3-96110-124-5. DOI: [10.5281/zenodo.1469555](https://doi.org/10.5281/zenodo.1469555).
- [9] Regina Stodden, Behrang QasemiZadeh, and Laura Kallmeyer. “TRAPACC and TRAPACCS at PARSEME Shared Task 2018: Neural Transition Tagging of Verbal Multiword Expressions”. In: *Proceedings of the Joint Workshop on Linguistic Annotation, Multiword Expressions and Constructions (LAW-MWE-CxG-2018)*. Santa Fe, New Mexico, USA: Association for Computational Linguistics, Aug. 2018, pp. 268–274. URL: <http://www.aclweb.org/anthology/W18-4930>.
- [10] Behzad Naderalvojud, Behrang Qasemizadeh, and Laura Kallmeyer. “Lexicon-Based Deep Learning for Contextual Sentiment Analysis”. In: *GermEval 2017: Shared Task on Aspect-based Sentiment in Social Media Customer Feedback*. 2017, pp. 18–21.
- [11] Laura Kallmeyer Q. Zadeh Behrang and Aurelie Herbelot. “Projection Aléatoire Non-Négative pour le Calcul de Word Embedding”. In: *Actes de TALN 2017, volume 1 : articles longs*. 2017, pp. 109–122.
- [12] Behrang Qasemizadeh and Laura Kallmeyer. “HHU at SemEval-2017 Task 2: Fast Hash-Based Embeddings for Semantic-Word Similarity Assessment”. In: *SemEval*. Vancouver, Canada: ACL, 2017.
- [13] Agata Savary, Carlos Ramisch, Silvio Cordeiro, Federico Sangati, Veronika Vincze, Behrang QasemiZadeh, Marie Candito, Fabienne Cap, Voula Giouli, Ivelina Stoyanova, and Antoine Doucet. “The PARSEME Shared Task on Automatic Identification of Verbal Multiword Expressions”. In: *MWE 2017*. Valencia, Spain: Association for Computational Linguistics, Apr. 2017, pp. 31–47. URL: <http://www.aclweb.org/anthology/W17-1704>.
- [14] Aurelie Herbelot and Behrang Q. Zadeh. “You and me... in a vector space: modelling individual speakers with distributional semantics”. In: **SEM 2016: The Fifth Joint Conference on Lexical and Computational Semantics*. ACL, Aug. 2016, pp. 179–188. URL: <http://aclweb.org/anthology/S/S16/S16-2023.pdf>.
- [15] Behrang Qasemizadeh. “A Study on the Interplay Between the Corpus Size and Parameters of a Distributional Model for Term Classification”. In: *Proceedings of the 5th International Workshop on Computational Terminology (Computerm2016)*. Patrick Drouin and Natalia Grabar and Thierry Hamon and Kyo Kageura and Koichi Takeuchi. Osaka, Japan: The COLING 2016 Organizing Committee, Dec. 2016, pp. 62–72. URL: <http://aclweb.org/anthology/W16-4708>.
- [16] Behrang QasemiZadeh and Laura Kallmeyer. “Random Positive-Only Projections”. In: **SEM 2016: The Fifth Joint Conference on Lexical and Computational Semantics*. ACL, 2016, pp. 188–198. URL: <http://aclweb.org/anthology/S/S16/S16-2024.pdf>.
- [17] Behrang QasemiZadeh and Anne-Kathrin Schumann. “The ACL RD-TEC 2.0: A Language Resource for Evaluating Term Extraction and Entity Recognition Methods”. In: *Proceedings of the Tenth International Conference on Language Resources and Evaluation (LREC 2016)*. Ed. by Nicoletta Calzolari, Khalid Choukri, Thierry Declerck, Marko Grobelnik, Bente Maegaard, Joseph Mariani, Asuncion Moreno, Jan Odijk, and Stelios Piperidis. Portoroz, Slovenia: European Language Resources Association (ELRA), 2016. ISBN: 978-2-9517408-9-1. URL: http://www.lrec-conf.org/proceedings/lrec2016/pdf/681_Paper.pdf.
- [18] Behrang QasemiZadeh. “Investigating the Use of Distributional Semantic Models for Co-Hyponym Identification in Special Corpora”. PhD Thesis: Computer Science. National University of Ireland, Galway, June 2015. URL: <http://hdl.handle.net/10379/5205>.
- [19] Behrang QasemiZadeh. “Random Indexing Revisited”. In: *Natural Language Processing and Information Systems*. Ed. by Chris Biemann, Siegfried Handschuh, Andre Freitas, Farid Meziane, and Elisabeth Metais. Vol. 9103. Lecture Notes in Computer Science. Springer International Publishing, June 2015, pp. 437–442. ISBN: 978-3-319-19580-3. DOI: [10.1007/978-3-319-19581-0_43](https://doi.org/10.1007/978-3-319-19581-0_43).
- [20] Behrang QasemiZadeh. “Random Manhattan Indexing: An Incremental Word Space Model”. In: *DGfS-CL Postersession 2015*. Deutschen Gesellschaft für Sprachwissenschaft. Universität Leipzig, Mar. 2015.
- [21] Behrang QasemiZadeh and Siegfried Handschuh. “Random Indexing Explained with High Probability”. In: *Text, Speech and Dialogue*. Ed. by Pavel Kral and Vaclav Matousek. Vol. 9302. Lecture Notes in Artificial Intelligence. Pilsen, Czech: Springer International Publishing Switzerland, Sept. 2015, pp. 480–489. DOI: [10.1007/978-3-319-24033-6_47](https://doi.org/10.1007/978-3-319-24033-6_47).

- [22] Anne-Kathrin Schumann and Behrang QasemiZadeh. *The ACL RD-TEC Annotation Guideline: A Reference Dataset for the Evaluation of Automatic Term Recognition and Classification*. Annotation Guidelines Ver. 2.6. Saarland University, Oct. 2015. doi: [10.13140/RG.2.1.1939.1446](https://doi.org/10.13140/RG.2.1.1939.1446).
- [23] Anne-Kathrin Schumann and Behrang QasemiZadeh. “Tracing Research Paradigm Change Using Terminological Methods: A Case Study on “Machine Translation” in the ACL Anthology Reference Corpus”. In: *Proceedings of the conference Terminology and Artificial Intelligence*. Ed. by Pamela Faber and Thierry Poibeau. Granada, Spain, Nov. 2015, pp. 115–122.
- [24] Behrang Q. Zadeh and Siegfried Handschuh. “Random Manhattan Integer Indexing: Incremental L1-Normed Vector Space Construction”. In: *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*. Ed. by Alessandro Moschitti, Bo Pang, and Walter Daelemans. Doha, Qatar: Association for Computational Linguistics, Oct. 2014, pp. 1713–1723. URL: <http://www.aclweb.org/anthology/D14-1178>.
- [25] Behrang Q. Zadeh and Siegfried Handschuh. “The ACL RD-TEC: A Dataset for Benchmarking Terminology Extraction and Classification in Computational Linguistics”. In: *Proceedings of the 4th International Workshop on Computational Terminology (Computerm)*. Ed. by Patrick Drouin, Natalia Grabar, Thierry Hamon, and Kyo Kageura. Dublin, Ireland: Association for Computational Linguistics and Dublin City University, Aug. 2014, pp. 52–63. URL: <http://www.aclweb.org/anthology/W14-4807>.
- [26] Behrang QasemiZadeh. “Random Manhattan Indexing: A Randomized Scalable Method for Semantic Similarity Measurement in L1 Normed Spaces”. In: *Book of Abstracts: Insight Student Conference*. Insight Centre, 2014. URL: https://www.insight-centre.org/sites/default/files/basic_pages_file/book_of_abstracts-electronic_version.pdf.
- [27] Behrang QasemiZadeh and Siegfried Handschuh. “Evaluation of Technology Term Recognition with Random Indexing”. In: *Proceedings of the Ninth International Conference on Language Resources and Evaluation (LREC’14)*. Ed. by Nicoletta Calzolari, Khalid Choukri, Thierry Declerck, Hrafn Loftsson, Bente Maegaard, Joseph Mariani, Asuncion Moreno, Jan Odijk, and Stelios Piperidis. Reykjavik, Iceland: European Language Resources Association (ELRA), May 2014, pp. 4027–32. ISBN: 978-2-9517408-8-4. URL: http://www.lrec-conf.org/proceedings/lrec2014/pdf/920_Paper.pdf.
- [28] Behrang QasemiZadeh and Siegfried Handschuh. “Random Manhattan Indexing”. In: *Proceedings of the 25th International Workshop on Database and Expert Systems Applications (DEXA)*. Munich, Germany: IEEE, Sept. 2014, pp. 203–208. doi: [10.1109/DEXA.2014.51](https://doi.org/10.1109/DEXA.2014.51). URL: <http://www.uni-weimar.de/medien/webis/events/tir-14/tir14-papers-final/zadeh14-random-manhattan-indexing.pdf>.
- [29] Behrang QasemiZadeh, Saeed Rahimi, and Mehdi Safaee Ghalati. “Challenges in Persian Electronic Text Analysis”. In: *CoRR abs/1404.4740* (2014). First published in 2006 at the Second Persian Computational Linguist Workshop. URL: <http://arxiv.org/abs/1404.4740>.
- [30] Simon Scerri, Behrang QasemiZadeh, Maciej Dabrowski, and Ismael Rivera. “Extracting Information for Context-aware Meeting Preparation”. In: *Proceedings of the Ninth International Conference on Language Resources and Evaluation (LREC’14)*. Ed. by Nicoletta Calzolari, Khalid Choukri, Thierry Declerck, Hrafn Loftsson, Bente Maegaard, Joseph Mariani, Asuncion Moreno, Jan Odijk, and Stelios Piperidis. Reykjavik, Iceland: European Language Resources Association (ELRA), May 2014, pp. 120–124. ISBN: 978-2-9517408-8-4. URL: http://www.lrec-conf.org/proceedings/lrec2014/pdf/1092_Paper.pdf.
- [31] Behrang Q. Zadeh and Siegfried Handschuh. “Investigating Context Parameters in Technology Term Recognition”. In: *Proceedings of the COLING Workshop on Synchronic and Diachronic Approaches to Analyzing Technical Language*. Ed. by Adam Meyers, Yifan He, and Ralph Grishman. Dublin, Ireland: Association for Computational Linguistics and Dublin City University, Aug. 2014, pp. 1–10. URL: <http://www.aclweb.org/anthology/W14-6001>.
- [32] Behrang QasemiZadeh. *The ACL RD-TEC: A Reference Dataset for the Evaluation of Automatic Term Recognition and Classification in Computational Linguistics: Annotation Guidelines*. Annotation Guidelines. Galway, Ireland: National University of Ireland, Galway, 2013. URL: http://atmykitchen.info/datasets/acl_rd_tec/annotation_guideline/annotation-guideline.pdf.
- [33] Behrang QasemiZadeh, Paul Buitelaar, T. Q. Chen, and Georgeta Bordea. “Semi-Supervised Technical Term Tagging With Minimal User Feedback”. In: *Proceedings of the Eighth International Conference on Language Resources and Evaluation (LREC-2012)*. Ed. by Nicoletta Calzolari, Khalid Choukri, Thierry Declerck, Mehmet Uğur Doğan, Bente Maegaard, Joseph Mariani, Jan Odijk, and Stelios Piperidis. Istanbul, Turkey: European Language Resources Association (ELRA), May 2012, pp. 617–621. ISBN: 978-2-9517408-7-7. URL: http://www.lrec-conf.org/proceedings/lrec2012/pdf/342_Paper.pdf.
- [34] Behrang QasemiZadeh and Paul Buitelaar. “Why Reading Papers When EYORE Will do That for you!?” In: *Proceedings Of NUI Galway – UL Alliance First Annual Engineering And Informatics Research Day*. Ed. by Michael Madden, Peter Corcoran, Marcus Keane, and Sean McGrath. Galway, Ireland: NUI Galway and University of Limerick, Apr. 2011.
- [35] Behrang QasemiZadeh. “Towards Technology Structure Mining from Scientific Literature”. In: *The Semantic Web – ISWC 2010*. Ed. by Peter F. Patel-Schneider, Yue Pan, Pascal Hitzler, Peter Mika, Lei Zhang, Jeff Z. Pan, Ian Horrocks, and Birte Glimm. Vol. 6497. Lecture Notes in Computer Science. Shanghai, China: Springer Berlin Heidelberg, Nov. 2010, pp. 305–312. ISBN: 978-3-642-17748-4. doi: [10.1007/978-3-642-17749-1_21](https://doi.org/10.1007/978-3-642-17749-1_21).
- [36] Behrang QasemiZadeh. *Towards Technology Structure Mining From Text By Linguistics Analysis*. Tech Report DERI TR 010-02-15. Lower Dangan, Galway: Digital Enterprise Research Institute, Feb. 2010.

- [37] Behrang QasemiZadeh and Paul Buitelaar. “Technology Structure Mining From Text by Linguistic Analysis”. In: *College of Engineering & Informatics: Research Day: DERI Research Session*. Galway, Ireland: NUI Galway, Apr. 2010, p. 15. URL: http://www.nuigalway.ie/faculties_departments/engineering/documents/abstracts_der_i.pdf.
- [38] Behrang QasemiZadeh, Paul Buitelaar, and Fergal Monaghan. “Developing a Dataset for Technology Structure Mining”. In: *Proceedings of 2010 IEEE Fourth International Conference on Semantic Computing (ICSC 2010)*. Pittsburgh, Pennsylvania, USA: IEEE, Sept. 2010, pp. 32–39. doi: [10.1109/ICSC.2010.73](https://doi.org/10.1109/ICSC.2010.73).
- [39] Chris Fowler and Behrang QasemiZadeh. “Towards a Common Event Model for an Integrated Sensor Information System”. In: *Proceedings of the 1st International Workshop on the Semantic Sensor Web (SemSensWeb) Collocated with ESWC 2009*. Ed. by Oscar Corcho, Manfred Hauswirth, and Manolis Koubarakis. Vol. 468. CEUR Workshop Proceedings. Crete, Greece: CEUR, June 2009, pp. 1–13. URL: http://ceur-ws.org/Vol-468/semsensweb2009_submission_3.pdf.
- [40] Philip Hanna, Ian O’Neill, Darryl Stewart, and Behrang QasemiZadeh. “Development of a Java-Based Unified and Flexible Natural Language Discourse System”. In: *Proceedings of the 7th International Conference on Principles and Practice of Programming in Java (PPPJ ’09)*. Calgary, Alberta, Canada: Association for Computing Machinery, Aug. 2009, pp. 21–29. doi: [10.1145/1596655.1596659](https://doi.org/10.1145/1596655.1596659).
- [41] Behrang QasemiZadeh. “Transcription of the Persian Language in the Electronic Format”. In: *CLUKI 2009 - 12th Annual Research Colloquium of the Special Interest Group for Computational Linguistics in the UK and Ireland*. 2009.
- [42] Behrang QasemiZadeh, Ian O’Neill, Philip Hanna, and Darryl Stewart. “A Data Model for Content Modeling of Temporal Media”. In: *Future Multimedia Networking*. Ed. by Andreas Mauthe, Sherali Zeadally, Eduardo Cerqueira, and Marilia Curado. Vol. 5630. Lecture Notes in Computer Science. Coimbra, Portugal: Springer Berlin Heidelberg, June 2009, pp. 194–199. ISBN: 978-3-642-02471-9. doi: [10.1007/978-3-642-02472-6_18](https://doi.org/10.1007/978-3-642-02472-6_18).
- [43] Behrang QasemiZadeh, Jiali Shen, Ian O Neill, Paul Miller, Philip Hanna, Darryl Stewart, and Hongbin Wang. “A Speech Based Approach to Surveillance Video Retrieval”. In: *Proceedings of Sixth IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS 2009)*. Genoa, Italy: IEEE, Sept. 2009, pp. 336–339. doi: [10.1109/AVSS.2009.54](https://doi.org/10.1109/AVSS.2009.54). URL: <http://hdl.handle.net/10379/4384>.
- [44] Behrang QasemiZadeh, Ian O’Neill, Philip Hanna, and Darryl Stewart. *Video Retrieval Using Natural Language: From Parse Tree to Database*. Technical Report 2008-isis. Belfast, UK: Queen’s University Belfast, Oct. 2008.
- [45] Behrang QasemiZadeh. “Farsi e-Orthography: An Example of e-Orthography Concept”. In: *Improving Non-English Web Searching (iNEWS07) SIGIR07 Workshop*. Ed. by Fotis Lazarinis, Jesus Vilares, and John I. Tait. Amsterdam, The Netherlands: ACM, July 2007, pp. 62–64.
- [46] Behrang QasemiZadeh, Saeed Rahimi, and Behrooz Mahmoodi Bakhtiari. “The First Parallel Multilingual Corpus of Persian: Toward a Persian BLARK”. In: *Proceedings of the Second Workshop on Computational Approaches to Arabic Script-based Language (CAASL2)*. Ed. by Ali Farghali and Karine Megerdumian. Also available as arXiv preprint (arXiv:1404.4572). Stanford University, July 2007, pp. 114–119. URL: <http://dadegan.ir/catalogue/QasemiZadeh2007.pdf>.
- [47] Behrang QasemiZadeh, Said Rahimi, Morteza Salariyan, and Ali Bahari Salim. “Report of Constructing the First Parallel Multilingual Corpus of Farsi”. Farsi. In: *Proceedings of the 12th International CSI Computer Conference (CSICC’2007)*. Tehran, Iran: Computer Society of Iran and Shahid Beheshti University, Feb. 2007, pp. 323–411. URL: <http://dadegan.ir/catalogue/QasemiZadeh1385.pdf>.
- [48] Restgou Haghi, Mohammadreza Kangavari, and Behrang QasemiZadeh. “A Novel Approach for Video Temporal Annotation”. In: *Proceedings of the International Conference on Multidisciplinary Information Sciences and Technology (InSciT 2006)*. Vol. 1. Multidisciplinary information sciences and technologies; Current research in information sciences and technologies. Merida, Spain: Open Institute of Knowledge, 2006, pp. 279–282. ISBN: 9788461131044. URL: <http://arxiv.org/ftp/arxiv/papers/1404/1404.4543.pdf>.
- [49] Abbas Koochari, Majid Namnabat, Saeed Rahimi, Behrang QasemiZadeh, and Morteza Salariyan. *Modelling Duration in a Persian Text-to-Speech System Using Neural Networks*. Technical Report clonizer-sw-loc-1385-1. Published in Farsi. Enqelab St.. Tehran, Iran: Digital Clone, 2006.
- [50] Abbas Koochari, Behrang QasemiZadeh, M. Kasaeiyan, and Majid Namnabat. “Ezafé Prediction in Phrases of Farsi Using CART”. In: *Proceedings of the International Conference on Multidisciplinary Information Sciences and Technology (InSciT 2006)*. Vol. 2. Multidisciplinary information sciences and technologies; Current research in information sciences and technologies. Merida, Spain: Open Institute of Knowledge, 2006, pp. 329–332. ISBN: 9788461131044. URL: <http://arxiv.org/ftp/arxiv/papers/1404/1404.4543.pdf>.
- [51] Behrang QasemiZadeh. “A Framework for Extracting Temporal Ontology From Annotated Multimedia/Video Data”. Farsi. Published in Farsi. Master’s thesis: Computer Engineering (Artificial Intelligence and Robotics). Tehran, Iran: Iran University of Science and Technology, Computer Engineering Department, Apr. 2006.
- [52] Behrang QasemiZadeh, H.R. Haghi, and Mohammadreza Kangavari. “A Framework for Temporal Content Modeling of Video Data Using an Ontological Infrastructure”. In: *Proceedings of the Second International Conference on Semantics, Knowledge and Grid (SKG ’06)*. IEEE Computer Society, Nov. 2006, pp. 38–38. doi: [10.1109/SKG.2006.4](https://doi.org/10.1109/SKG.2006.4).
- [53] Behrang QasemiZadeh, Ali Ilkhani, and Amir Ganjei. “Adaptive Language Independent Spell Checking Using Intelligent Traverse on a Tree”. In: *IEEE Conference on Cybernetics and Intelligent Systems (CIS)*. Bangkok: IEEE, June 2006, pp. 1–6. doi: [10.1109/ICCIS.2006.252325](https://doi.org/10.1109/ICCIS.2006.252325).

- [54] Behrang QasemiZadeh, Mohammadreza Kangavaria, and S. Hadi Rastgou Haghi. “Temporal Ontology for Video Data Modelling”. In: *Proceedings of the International Conference on Multidisciplinary Information Sciences and Technology (InSciT 2006)*. Multidisciplinary information sciences and technologies; Current research in information sciences and technologies. Merida, Spain: Open Institute of Knowledge, 2006. ISBN: 9788461131044. URL: <http://arxiv.org/ftp/arxiv/papers/1404/1404.4543.pdf>.
- [55] Behrang QasemiZadeh and Saeed Rahimi. “Persian in MULTEXT-East Framework”. In: *Advances in Natural Language Processing*. Ed. by Tapio Salakoski, Filip Ginter, Sampo Pyysalo, and Tapio Pahikkala. Vol. 4139. Lecture Notes in Computer Science. Springer Berlin Heidelberg, 2006, pp. 541–551. ISBN: 978-3-540-37334-6. DOI: [10.1007/11816508_54](https://doi.org/10.1007/11816508_54).
- [56] Behrang QasemiZadeh, Said Rahimi, Morteza Salariyan, Ali Torkamani, Heydar Semiyari, Abbas Koochari, Majid Namnabat, and Loghman Barari. “A Novel Approach for Morphological Analysis of Persian Words”. Farsi. In: *Proceedings of the 11th International CSI Computer Conference (CSICC'2006)*. Tehran, Iran: Computer Society of Iran and Institute for Research in Fundamental Sciences, Jan. 2006, pp. 323–411.
- [57] Loghman Barari and Behrang QasemiZadeh. “Clonizer Spell Checker Adaptive Language Independent Spell Checker”. In: *ICGST International Conference on Artificial Intelligence and Machine Learning (AIML-05)*. Ed. by Ashraf Aboshosha. Vol. 05. Cairo, Egypt: ICGST, Dec. 2005, pp. 65–71. URL: <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.91.7252&rep=rep1&type=pdf>.
- [58] Behrang QasemiZadeh. *A Study on Semantic Structure and Methods of Extracting Concepts from Video Data*. Seminar Report. Published in Farsi. Iran University of Science and Technology, 2005.
- [59] Behrang QasemiZadeh. *An Introduction to Expert System and Frame-Based Knowledge Representation Systems*. Report. Published in Farsi. Iran University of Science and Technology, Computer Engineering Department, 2005.
- [60] Behrang QasemiZadeh and Majid Namnabat. *An Architecture for a Farsi Text-to-Speech System*. Technical Report clonizer-SW-LOC-966. Published in Farsi. Digital Clone, Localisation Department, 2005.
- [61] Behrang QasemiZadeh and Saeed Rahimi. *A Lexicon for Farsi Words*. Technical Report clonizer-SW-LOC-1027. Published in Farsi. Digital Clone, Localisation Department, 2005.
- [62] Behrang QasemiZadeh and Saeed Rahimi. *A Study on Persian Derivational Morphology, Restriction of the Co-Occurrences, and its Application to Computational Morphology*. Technical Report clonizer-SW-LOC-1012. Published in Farsi. Digital Clone, Localisation Department, 2005.

Invited Talks

- *PPP-Hashing and PeARS: A Hashing Technique for Distributed Semantic Indexing*, Computer Laboratory, University of Cambridge, UK, 10 Nov 2016.
- *Distributional Semantics for Identifying Co-Hyponym Terms in Specialised Corpora*, Computational Linguistics Group, Heinrich-Heine-Universität Düsseldorf, Germany, 02 Oct 2015.
- *Identifying Co-Hyponym Terms from Specialised Corpora*, UKP Lab, TU Darmstadt, Germany, 27 May 2015.
- *Extracting Terminological Resources and Co-Hyponym Terms: Building Technology Profiles*, Research Colloquium, University of Passau, Germany, 05 May 2015.
- *What a Pioneer in Digital Humanities Needs to Know About Distributional Semantics*, Research Colloquium, University of Passau, Germany, 16 Dec 2014.
- *Vector Spaces for Information Extraction*, INSIGHT Workshop on Latent Space Methods, University College Dublin, Ireland, 17 Jan 2014.
- *Relational Tuple Extraction from Text*, Research Seminar, Sharif University of Technology, Iran, 08 Feb 2012.
- *Formalising Morphosyntactic Specification of Farsi in the MULTEXT-East Framework*, Research Seminar, Sharif University of Technology, Iran, 08 Feb 2012.
- *Towards Technology Structure Mining from Scientific Literature*, University of Passau, Germany, 03 Mar 2011.
- *Technology Structure Mining*, DERI Friday Talk Series, National University of Ireland Galway, Ireland, 05 Mar 2010.
- *Populating the Semantic Web with Relations from Text*, DERI Reading Group Series, National University of Ireland Galway, Ireland, 03 Feb 2010.
- *Use of Ontologies for Dialogue about Audiovisual Data*, Workshop on Spoken Dialogue Management Systems, Ulster University, UK, Sep 2008.
- *The Farsi Corpus Project and Morphological Analysis of Farsi*, Student Conference on Empirical and Computational Linguistics (JSSECL), Zadar, Croatia, 16 Sep 2006.