

# Curriculum Vitae

## BEHRANG QASEMIZADEH

Email: [behrangatoffice@gmail.com](mailto:behrangatoffice@gmail.com)

Web: <http://pars.ie>

**Present Position:** Since January 2016, I have been a postdoctoral researcher at the [at the DFG Collaborative Research Centre 991](#) of Heinrich-Heine-Universität Düsseldorf. Under the supervision of Prof. Laura Kallmeyer, I am involved in the *SFB 991* project. With an emphasis on frame semantics, this project is a multidisciplinary investigation of the structure of representations in language, cognition, and science. I am a member of the B08 project that involves the use of data-driven methods for hierarchical frame induction—that is, for empirical investigation of theoretical discussions in the context of the SFB 991 project.

**Long Term Research Goal:** The study of language, and perhaps other sensory information, as a key for investigating the fundamental research question “how does the human mind work?”. As a technologist, my aim is to facilitate man-machine and man-man communication. From the complementary perspective of linguistic anthropology, I am interested in the study of consequences of technological developments in the area of computational linguistics. Simply put, I am striving for a holistic study of communication systems and preparing for the day that man prefers communicating with machines.

**Current Research Areas:** Computational linguistics; natural language processing; information extraction.

**Current Research Topics of Interest:** Vector-based distributional methods for formalising natural language structures; random projections and embedding techniques; sketch-based methods; (inter-)active learning techniques.

## Personal Information

Gender: Male.

Country of Residence: Germany.

## Education

### *Ph.D. Computer Science:*

Date: 08 July 2015.

Thesis Title: Investigating the Use of Distributional Semantic Models for Co-Hyponym Identification in Special Corpora.

Awarding Body: [National University of Ireland, Galway](#).

Advisor: Prof. Dr. Siegfried Handschuh.

### *Master's Degree in Computer Engineering – Artificial Intelligence and Robotics:*

Date: 19 April 2006.

Master's Thesis Title: A Framework for Extracting Temporal Ontology From Annotated Multimedia/Video Data.

Awarding Body: [Iran University of Science and Technology](#).

Advisor: Prof. Dr. Mohammad Reza Kangavari.

### *Additional Studies:*

Participant of the [ENeL training school 2016: Tools and methods for creating innovative e-dictionaries](#), University of Ljubljana, 17 May 2016 to 20 May 2016.

Research graduate visitor to the [APEX Data & Knowledge Management Lab](#), Shanghai Jiao Tong University, 20 April 2011 to 27 June 2011, collaborated with [Tianqi Chen](#) and [Haofen Wang](#).

Participant of the 17th European Summer School on Logic, Language and Information (ESLLI05), Edinburgh, UK, August 2005.

Participant of the Jadertina Summer School in Empirical and Computational Linguistics (JSSECL), Zadar University, September 2006.

Participant of the 3rd DGFS Computational Linguistics Fall School, Bochum, Germany, September 2005.

## Honours and Awards

Full Ph.D. scholarship from the Science Foundation Ireland, 2009–2014.

Two months travel grant in the context of the Marie Curie action IRSES under Grant No. 24761 (Net2).

Master's thesis research grant from the [Iran Research Institute for Information and Communication Technology \(ITRC\)](#).

Outstanding student at the Computer Engineering Department of Iran University of Science and Technology in the academic year 2003–2004.

Ranked under the top 0.001% among more than 10,000 participants of the Iranian nationwide university entrance exam for graduate studies (known as *Konkour*) in the field of computer science in the academic year 2003–2004.

Ranked under the top 0.001% among more than 500,000 participants of the Iranian nationwide university entrance exam for undergraduate studies in the category of mathematics and physics in the academic year 1998–1999.

## Skills

### *Natural Languages:*

English: Fluent;

Modern Persian (i.e., Farsi): Native;

Arabic: Basic Knowledge of Formal (i.e., Quranic) Arabic;

\* I am open to new experiences in life and thus willing to learn new languages!

### *Computer Programming:*

JAVA: Fluent, almost like a native speaker. In addition to Java core class libraries, I have a solid background in:

- Natural Language Processing: Expert knowledge of the Stanford Core-NLP, MALT Parser, Apache OpenNLP libraries, as well as limited lab experience with the GATE and Apache UIMA frameworks;
- Data Analytics and Machine Learning: Considerable knowledge in using Apache Common Math, Apache Mahout, WEKA, JAVA-ML, DL4J, a variety of Support Vector Machines implementations such as LibSVM, as well as libraries for common linear algebra operations such as JAMA, COLT, and ND4J;
- Markup Languages and Formalisms: Familiar with HTML, XML, RDF (as well as how to use the Apache JENA framework for the ‘Semantic Web’ application development);
- Other: I am experienced with distributed TCP/IP, agent-based architectures. In addition, I have explored Apache Hadoop and have a good theoretical knowledge and understanding of its related projects such as Apache Drill, Pig, and SPARK. I am currently exploring GPU-accelerated computing technologies.

Prolog: Fluent in ISO dialect. Specifically, I used the LPA Win-Prolog including the following toolkits:

- Chimera Agents: An agent development and deployment tool;
- TCP/IP Toolkit: For distributed processing over a network;
- ProWeb Server: Dynamic server based application development;
- ProData: Database interface toolkit;
- Intelligence Server: For embedding Prolog codes in programming languages such as Java and C.

**NOTE:** I am comfortable in reusing and modifying other programming languages such as Python, Scala (i.e., functional programming), C, C++, etc. For me, learning a programming language or technology is not an end in itself. It is, however, an integral aspect of the process of realising vague research ideas to real-world applications, or to investigate research hypotheses. Hence, I am ready to learn new languages and technologies to meet the demands of research projects that I get involved in.

### *Additional Computing Skills:*

Web Servers: Apache Tomcat; Apache HTTP Server.

Database: Relational databases such as MySQL, H2DB, Microsoft SQL Server, as well as NoSQL databases such as Redis, Berkeley DB, Neo4J and Kyoto Cabinet.

On-line Document Management System: Drupal (with a little bit of PHP hacking if necessary), Google Sites (and its relevant Java APIs).

Web Interface: HTML, a small bit of PHP, and enough of [Vaadin](#) (JAVA APIs that exploit Google Web Toolkit).

Office Productivity: Microsoft Office,  $\LaTeX$ /TikZ/PGF.

Project Management: Microsoft Project (i.e., Gantt charts for project structuring, employee scheduling, tracking dependencies and work progress), and Basecamp.

Teamwork: Subversion, JIRA, Confluence, Microsoft Visual SourceSafe, and basic Git.

Operating Systems: Linux, Windows.

### *Soft Skills:*

I am committed to my personal development in key areas such as health, work ethic, positive attitude, communication, time management, leadership, and teamwork. I am interested in the concept of *Mindfulness at Work*—that is, achieving balance.

## Professional Research Experience

### *January 2016 – Present:*

Role: Postdoctoral Researcher.

Institute: [DFG Collaborative Research Centre 991, Computational Linguistics Department](#), University of Düsseldorf.

Principal investigator: Prof. Dr. Laura Kallmeyer.

### *October 2014 – November 2015:*

Role: Wissenschaftlicher Mitarbeiter (Scientific Assistant).

Institute: [Chair of Digital Libraries and Web Information Systems](#), University of Passau.

Principal investigator: Prof. Dr. Siegfried Handschuh.

Contributions:

- Teaching: Text mining practical course (8 ECTS); this includes designing course content based on Birds et al.'s *Natural Language Processing with Python* for 7 weeks of 6 hours theoretical/hands-on teaching, followed by defining project topics, and assisting students to achieve their project goals in the remaining of the semester (performance records are available upon request).
- Occasional contributions to research grant proposals: Successfully acquired a grant from NVIDIA for 4 Tesla cards (within the so-called [CUDA research centre scheme](#));  
PI: Prof. Dr. Siegfried Handschuh.
- Disseminating research findings:
  - Participating and presenting research at the colloquiums co-organised by the chair of digital humanities and digital libraries at the University of Passau;
  - 4 research papers.

*April 2009 – October 2014:*

Role: PhD Candidate and Research Assistant<sup>1</sup>.

Institute: INSIGHT Centre for Data Analytics<sup>2</sup>; National University of Ireland, Galway.

Principal investigators: Prof. Dr. Siegfried Handschuh, Dr. Paul Buitelaar.

Operational Manager: Dr. Brian Wall.

Contributions:

- Pursuing a Ph.D.: Investigating corpus-based, distributional approaches to natural language semantics. This study was carried out through a scholarship that was provided in the context of *Lion* and *Lion 2* projects—with the theme of *enabling networked knowledge*—which are supported by the Science Foundation Ireland under grant numbers SFI/02/CE1/I131 and 12/RC/2289.
- Occasional contributions to research deliverables:
  - Design and development of the *MLTagger* tool: A support vector-based key-phrase extractor.
  - Design and development of a Distant Reading tool from corpora: An information extraction tool that extracts triples and generates a graph visualisation in a PDF format.
  - Design and development of the TextBarCode tool: A visual text analytic tool for exploring natural language documents, particularly for knowledge-workers in the domain of digital humanities.
  - Contributing to knowledge transfer deliverables involving the *CISCO Systems, Inc*: Designing algorithms and developing codes for domain-adaptive extraction of topics (keywords) from chat logs and email conversations between employees in a *Jabber*-based communication system (e.g., for expertise finding, customer support, etc.).
- Disseminating research findings:
  - Several research publications (see list of [publications](#)).
  - Development of the ACL RD-TEC: A language resources for the evaluation of terminology extraction and classification methods consisting of manually annotated terms in the [ACL Anthology Reference Corpus](#). This resource is published by the European Language Resources Association (ELRA) under [catalogue number ELRA-T0375](#).<sup>3</sup>
- Other activities:
  - Assisting new students (2013/2014) with disabilities through an initiative by the [disability support service](#) office at the National University of Ireland, Galway.

*December 2007 – April 2009:*

Role: Research Assistant.

Institute: [Speech and Vision Systems; The Institute of Electronics, Communications and Information Technology](#), Queen's University of Belfast.

Principal investigator: Dr. Ian O'Neill, Dr. David Linton.

Other Collaborators: Dr. Philip Hanna, Dr. Chris Fowler, Dr. Darryl Stewart, and Dr. Paul Miller.

Contributions:

- Prototyping a Natural Language Dialogue System for Surveillance Video Retrieval:
 

The role was defined in the context of the ISIS Project (An Integrated Sensor Information System for Crime Prevention) founded by the Engineering and Physical Science Research Council (EPSRC) under grant number EP/E028640/1. I was in charge of bridging the semantic gap between natural language utterances and surveillance video data in a complex multi-agent system. To achieve the objectives of the project, a Prolog-based distributed environment for information consolidation from heterogeneous resources (and various programming environments such as C++ and Java) was developed. The task encompassed the development of (a) a generic high-level event detection system, (b) translating detected events into and from natural language descriptions, and (c) a basic dialogue management system and domain-specific strategies for information seeking dialogues (which are defined in the context of the project for footage retrieval).

<sup>1</sup>I served as research assistant from June 2012 to August 2013; during the rest of this period, I was a PhD student.

<sup>2</sup>Formerly known as the [Digital Enterprise Research Institute \(DERI\)](#).

<sup>3</sup>ISLRN: 699-305-362-089-6.

- Disseminating research findings:
  - 4 peer-reviewed conference papers (see the list of publications).
  - Active participation in project meetings and reviews.

### *Prior to December 2007:*

#### **Basic Language Resources Kit for Modern Farsi (aka. Persian).**

Date: May 2005 – December 2007 (before immigrating to the UK).

Role: Founder and organiser of a set of self-founded (volunteer-based) projects for developing language resources and processing tools for Farsi.

Main Contributions:

- Integration of Farsi into the [MULTEXT-East framework \(officially since version 4\)](#):<sup>4</sup>
  - Devising morphosyntactic specifications for Farsi in the MULTEXT-East Framework;
  - Manual annotation of a Farsi translation of Orwell's 1984 corpus.<sup>5</sup>
- A lexicon of Farsi stems, available from ELRA ([Catalogue Reference: ELRA-L0087](#)).
- Research publications as well as a set of tools (in Microsoft Visual C++ and Java) for Farsi electronic-texts standardisation.

#### **Master Thesis Research Grant: Video Semantics and Retrieval**

Date: August 2005 – April 2006.

Institute: [Iran Research Institute for Information and Communication Technology \(ITRC\)](#) and Iran University of Science and Technology.

Principal investigator: Prof. Dr. Mohammad Reza Kangavari.

Role: Member of a team of two Master's degree students for the development of a prototype system to extract events from subtitled videos and visualise them in a time line.

Contributions: A working demo system developed in C# as well as technical reports and research publications.

#### **Contributions to several commercial products that involve language technologies for Farsi**

- Farsi text-to-speech;
- Spell checking;
- Post-processing the output of an optical character recognition system.

\* Please see the below [section](#) for details of professional experience during this period of time.

## **Teaching Experience**

Text Mining Practical : Summer Semester 2015, University of Passau, Bavaria, Germany.

Text Mining Practical : Winter Semester 2014, University of Passau, Bavaria, Germany.

Data Storage and Retrieval: Summer Semester 2007, University of Applied Science and Technology, Tehran, Iran.

## **Other Professional Experience**

### *December 2012 – Now:*

Company: [ISIS Forensics Ltd. \(UK\)](#)<sup>6</sup>.

Role: Technical Consultant (occasional).

Duties/Project: Design and development of forensic linguistic solutions (gender and age profiling of online personas).

Supervisors: Dr. James Walkerdine, Dr. Phil Greenwood.

<sup>4</sup>See also <http://hdl.handle.net/11356/1043> and <http://hdl.handle.net/11356/1042>.

<sup>5</sup>The core idea behind this project came from Dr. Damir Cavar during the 3rd DGFS Computational Linguistics Fall School. The idea was then realised by generous advice from Dr. Tomaz Erjavec.

<sup>6</sup>Now known as [Relative Insight](#).

*September 2006 – December 2007:*

Company: Text and Speech Technology (Ltd), Tehran, Iran.

Role: Co-Founder and Consultant.

Duties: The company was founded together with my former colleague, Mr. Majid Namnabat. The focus of the company was on the development of language technologies for Farsi. During this time, in addition to technical support, I was in charge of writing proposals for national research institutes and bidding projects from several companies. Although I still hold a share of the company, I have not been involved in any of the company's activities since December 2007.

Contributions: Successful grant applications for the development of language resources for Farsi-English machine translation (through [ITRC](#)); development of a Farsi text-to-speech system and its integration with Microsoft Windows Speech API to be used together with the [JAWS screen reader](#) (through Artificial Intelligence Rayvarz Co.).

Collaborators: Mr. Majid Namnabat and Dr. Saeed Rahimi (currently a lecturer at Sharif University of Technology).

*2005–2006 (In a Number of Occasions):*

Company: Lexical Computing Ltd. (i.e., the [Sketch Engine](#)), UK.

Role: Consultant.

Duties: Development of language resources for statistical Farsi-English machine translation; helping with various issues when the company developed a Persian Web Corpus.

Supervisor: Dr. Adam Kilgarriff.

*September 2004 – September 2006:*

Company: Digital Clone (Shegarf), Tehran, Iran.

Role: Localisation Team Leader.

Duties: Development of language technologies for Farsi and software localisation in the context of the [CloniZER](#) project. In this role, I managed a team of 7 people for the development of three sub-projects including text-to-speech, voice command, and spell checking for Farsi. In this role, I also actively contributed towards the design of the spell-checking tool as well as the phonemic transliteration of Farsi text (including the detection of stress, duration, and pitch) in the text-to-speech system.

Supervisor: Mr. Heydar Semyari.

*March 2004 – August 2004:*

Company: [Artificial Intelligence Rayvarz](#), Tehran, Iran.

Role: System Analyst and Developer (part-time).

Duties: Research and development for a Farsi text-to-speech system, enhancing the output of the company's optical character recognition engine for Farsi.

Supervisor: Mr. Hossein Mirheydari.

Note: Due to several reasons, such as the involved intellectual property holders in the text-to-speech project, I had a continuous engagement with the company until December 2007.

**Other Professional Services**

- Member of the Association for Computational Linguistics (ACL), and of the Artificial Intelligence Association of Ireland.
- Participation in the WP3 of the IC1207 COST Action: [PARSEME \(PARSing and Multi-word Expressions\)](#).
- Reviewer for a number of conferences such as ACL, AEPC, EACL, EKAW, ESWC, ISWC, LREC, NLDB, WWW.
- Reviewer for journals of Data & Knowledge Engineering, IEEE Transactions on Systems, Man and Cybernetics, IEEE Access, and Natural Language Engineering.

## List of Published Language Resources

- The Farsi Section of PARSEME Shared Task Data (v. 1.0) : multilingual corpora annotated with verbal multiword expressions, Published through LINDAT/CLARIN digital library at Institute of Formal and Applied Linguistics, Charles University in Prague, <http://hdl.handle.net/11372/LRT-2282>.
- ACL RD-TEC 2.0: An annotated corpus for evaluating term and entity recognition methods(together with Schumann, Anne-Kathrin), Published through LINDAT/CLARIN digital library at Institute of Formal and Applied Linguistics, Charles University in Prague, <http://hdl.handle.net/11372/LRT-1661>, ISLRN: 178-746-003-351-0.
- ACL RD-TEC: A Terminological resource in the domain of computational linguistics; ELRA Catalogue Reference: ELRA-T0375, ISLRN: 699-305-362-089-6.
- Persian Lexicon: A lexicon of more than 40,000 non-inflected Farsi words with rich semantic and morphological annotations as well as phonemic transliteration; ELRA Catalogue Reference: ELRA-L0087, ISLRN: 547-614-436-004-7.
- FASpell: A dataset for the evaluation of spell checking algorithms; available through LINDAT/CLARIN digital library at Institute of Formal and Applied Linguistics, Charles University in Prague, <http://hdl.handle.net/11372/LRT-1547>.
- Persian MULTEXT-East Framework Lexicon: A small lexicon of 13,427 inflected words in Farsi with morphosyntactic annotation and phonemic transliteration. ELRA Catalogue Reference: ELRA-L0086, ISLRN: 884-966-712-343-0.
- Persian 1984 Corpus: Farsi translation of Orwell’s 1984 manually annotated with phonemic transliterations and morphosyntactic tags according to specifications proposed in the MULTEXT-East framework. ELRA Catalogue Reference: ELRA-W0054, ISLRN: 851-240-629-673-1.

Apart from the ELRA distributions, Farsi sections of the MULTEXT-East resources are available through CLARIN.SI:

- MULTEXT-East "1984" annotated corpus 4.0: See <http://hdl.handle.net/11356/1043>.  
 MULTEXT-East non-commercial lexicons 4.0: See <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] Kata Gábor, 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>.
- [2] 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>.
- [3] 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>.
- [4] Behzad Naderalvojjoud, 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.
- [5] Laura Kallmeyer Q. Zadeh Behrang and Aurelie Herbelot. "Projection Alatoire Non-Negative pour le Calcul de Word Embedding". In: *Actes de TALN 2017, volume 1 : articles longs*. 2017, pp. 109–122.
- [6] Behrang Qasemizadeh and Laura Kallmeyer. "HHU at SemEval-2017 Task 2: Fast Hash-Based Embeddings for SemanticWord Similarity Assessment". In: *SemEval*. Vancouver, Canada: ACL, 2017.

- [7] 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, 2017, pp. 31–47. URL: <http://www.aclweb.org/anthology/W17-1704>.
- [8] 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, 2016, pp. 179–188. URL: <http://aclweb.org/anthology/S/S16/S16-2023.pdf>.
- [9] 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, 2016, pp. 62–72. URL: <http://aclweb.org/anthology/W16-4708>.
- [10] 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>.
- [11] 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](http://www.lrec-conf.org/proceedings/lrec2016/pdf/681_Paper.pdf).
- [12] 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>.
- [13] 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.
- [14] Behrang Qasemizadeh. “Random Manhattan Indexing: An Incremental Word Space Model”. In: *DGfS-CL Postersession 2015*. Deutschen Gesellschaft für Sprachwissenschaft. Universität Leipzig, Mar. 2015.
- [15] 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.
- [16] 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.
- [17] 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.
- [18] 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>.
- [19] 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 4<sup>th</sup> 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>.



- [20] 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](https://www.insight-centre.org/sites/default/files/basic_pages_file/book_of_abstracts-electronic_version.pdf).
- [21] 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](http://www.lrec-conf.org/proceedings/lrec2014/pdf/920_Paper.pdf).
- [22] Behrang QasemiZadeh and Siegfried Handschuh. “Random Manhattan Indexing”. In: *Proceedings of the 25<sup>th</sup> International Workshop on Database and Expert Systems Applications (DEXA)*. Munich, Germany: IEEE, Sept. 2014, pp. 203–208. DOI: 10.1109/DEXA.2014.51. URL: <http://www.uni-weimar.de/medien/webis/events/tir-14/tir14-papers-final/zadeh14-random-manhattan-indexing.pdf>.
- [23] Behrang QasemiZadeh, Saeed Rahimi, and Mehdi Safae 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>.
- [24] 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](http://www.lrec-conf.org/proceedings/lrec2014/pdf/1092_Paper.pdf).
- [25] 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>.
- [26] 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](http://atmykitchen.info/datasets/acl_rd_tec/annotation_guideline/annotation-guideline.pdf).
- [27] 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](http://www.lrec-conf.org/proceedings/lrec2012/pdf/342_Paper.pdf).
- [28] Behrang QasemiZadeh and Paul Buitelaar. “Why Reading Papers When EEYORE 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.
- [29] 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.
- [30] 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.
- [31] 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](http://www.nuigalway.ie/faculties_departments/engineering/documents/abstracts_der_i.pdf).

- [32] 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).
- [33] Chris Fowler and Behrang QasemiZadeh. “Towards a Common Event Model for an Integrated Sensor Information System”. In: *Proceedings of the 1<sup>st</sup> 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](http://ceur-ws.org/Vol-468/semsensweb2009_submission_3.pdf).
- [34] 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 7<sup>th</sup> 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).
- [35] Behrang QasemiZadeh. “Transcription of the Persian Language in the Electronic Format”. In: *CLUKI 2009 - 12<sup>th</sup> Annual Research Colloquium of the Special Interest Group for Computational Linguistics in the UK and Ireland*. 2009.
- [36] 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).
- [37] 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>.
- [38] 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.
- [39] 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.
- [40] 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>.
- [41] 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 12<sup>th</sup> 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>.
- [42] 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>.
- [43] 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.
- [44] Abbas Koochari, Behrang QasemiZadeh, M. Kasaeiyan, and Majid Namnabat. “Ezafe 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>.
- [45] 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.

- [46] 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).
- [47] Behrang QasemiZadeh, Ali Ilkhani, and Amir Ganjeii. “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).
- [48] 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>.
- [49] 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).
- [50] 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 11<sup>th</sup> International CSI Computer Conference (CSICC'2006)*. Tehran, Iran: Computer Society of Iran and Institute for Research in Fundamental Sciences, Jan. 2006, pp. 323–411.
- [51] 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>.
- [52] 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.
- [53] 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.
- [54] 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.
- [55] Behrang QasemiZadeh and Saeed Rahimi. *A Lexicon for Farsi Words*. Technical Report clonizer-SW-LOC-1027. Published in Farsi. Digital Clone, Localisation Department, 2005.
- [56] 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.

## Talks

- *PPP-Hashing and PeARS: A Hashing Technique for Distributed Semantic Indexing*, Computer Laboratory, University of Cambridge, UK. 10 November 2016.
- *Distributional Semantics for Identifying Co-Hyponym Terms in Specialised Corpora*. Computational Linguistics Department at Heinrich-Heine-Universität Düsseldorf, Germany. 02 October 2015.
- *Identifying co-hyponym terms from special corpora*. Ubiquitous Knowledge Processing (UKP) at TU Darmstadt, Germany. 27 May 2015.
- *Extracting terminological resources and co-hyponym terms: the use case of building technology profiles*. Forschungskolloquium im Sommersemester 2015, Philosophische Fakultät, University of Passau, Germany. 05 May 2015.
- *What a pioneer in digital humanities really needs to know about distributional semantics*. Forschungskolloquium im Wintersemester 2014/15, Philosophische Fakultät, University of Passau, Germany. 16 December 2014.
- *Vector spaces for information extraction*. INSIGHT Workshop on Latent Space Methods, University College Dublin (UCD), Ireland. 17 January 2014.
- *Relational tuple extraction from text*. Research Seminar at the Language and Linguistics Department of Sharif University of Technology, Iran. 08 February 2012.

- *Formalising morphosyntactic specification of Farsi in MULTEXT-East framework*. Research Seminar at the Language and Linguistics Department of Sharif University of Technology, Iran. 8 February 2012.
- *Towards technology structure mining from scientific literatures*. During a Visit to the Chair of Computational Linguistics (Prof. Dr. Anette Frank). 03 March 2011.
- *Technology structure mining*. DERI Friday Talk Series, National University of Ireland, Galway. 05 March 2010.
- *Populating the Semantic Web with Relations from Text*. DERI Reading Group Series, National University of Ireland, Galway. 03 February 2010.
- *Use of Ontologies for Dialogue about Audiovisual Data*. Workshop on Spoken Dialogue Management Systems, Ulster University at Jordanstown, Northern Ireland, September 2008.
- *The Farsi corpus project and morphological analysis of Farsi*. Student Conference on Empirical and Computational Linguistics (JSSECL WS CECL 2006), Zadar, Croatia. 16 September 2006.

## Extracurricular Activity

I try to remember that life is too short! I love long distance walks for reflection, as a tool for being mindful and grounded. I also like thinking about very abstract, scattered, and unrelated things—it is a kind of luxury when you are supposed to remain focused at all time. I like the cinema of Wes Anderson and Abbas Kiarostami. From time to time, I enjoy drawing ordinary things that I use in my daily life. I love yoga (it is not a girly thing to do!).

## References

Available upon request.