

Paramita Mirza

Preussenstrasse 32, 66111 Saarbrücken, GERMANY

☎ (+49) 176-2556-1055 | ✉ paramita135@gmail.com | 🏠 paramitamirza.com | 📱 paramitamirza | 🌐 paramitamirza

Research Interests

Natural Language Processing, Computational Semantics, Information Extraction, Knowledge Harvesting

Experience

Max Planck Institute for Informatics

Saarbrücken, Germany

Postdoctoral researcher

Jul. 2016 - PRESENT

- Research focuses include causation and common sense knowledge harvesting, emotional aspect of relation extraction and the semantics of numbers in natural language texts.
- Supervised by Prof. Gerhard Weikum.

Free University of Bozen-Bolzano

Bolzano, Italy

Research Assistant

May. 2016 - Jun. 2016

- Research focus was extracting relation cardinality, i.e., cardinal numbers in texts representing the cardinality of objects for a particular relation, e.g., number of children.
- Supervised by Dr. Simon Razniewski and Prof. Werner Nutt.

National University of Singapore

Singapore

Postgraduate Intern

Jun. 2015 - Sep. 2015

- Research focus was exploring word embeddings and deep learning techniques for temporal ordering of events.
- Supervised by Dr. Min-Yen Kan and Prof. Hwee Tou Ng.

Samsung Electronics Indonesia

Bekasi, Indonesia

Software R&D Engineer

Oct. 2008 - Aug. 2010

- Main responsibilities included developing and debugging firmware of electronic products, including DVD & MP3 players and Home Theater Systems.

Education

University of Trento

Trento, Italy

Ph.D., Doctoral Program in Information and Communication Technology (ICT)

2012 - 2016

- Supported by project specific grants by Fondazione Bruno Kessler, Center for Information and Communication Technology (FBK-ICT IRST).
- Graduated on April 12, 2016.
- Thesis: "Extracting Temporal and Causal Relations between Events" advised by Dr. Sara Tonelli.

Free University of Bozen-Bolzano (FUB)

Bolzano, Italy

M.Sc., European Master's Program in Language and Communication Technologies

(LCT)

2011 - 2012

- Second year of the LCT program, supported by Erasmus Mundus scholarships.
- Graduated on October 12, 2012 (grade: 110/110).
- Thesis: "Exploiting Combinatory Categorical Grammar (CCG) within Distributional Semantic Models" advised by Dr. Raffaella Bernardi.

University of Lorraine

Nancy, France

M.Sc., European Master's Program in Language and Communication Technologies

(LCT)

2011 - 2012

- First year of the LCT program, supported by Erasmus Mundus scholarships.

Bandung Institute of Technology

Bandung, Indonesia

B.Sc., Informatics

2004 - 2008

- Partially supported by college scholarships from Samsung Electronics Indonesia.
- Graduated on July 2, 2008 (grade: 3.43/4.00).
- Thesis: "Support Vector Machine Application for Information Extraction from Job Fair Advertisements" advised by Dr. Dwi Hendratmo Widyantoro.

Project Experience

Newsreader

Trento, Italy

EU-funded research project www.newsreader-project.eu

Nov. 2012 - Nov. 2015

- Research area: event & temporal information extraction, temporal ordering & causation.
- Developed a system for extracting temporal ordering and causation between events from financial news, to be used in automated construction of event timelines.
- Led to PhD dissertation, 3 conference publications [2, 7, 8] and 4 workshop publications [5, 9–11].
- Supervised by Dr. Sara Tonelli and Dr. Anne-Lyse Minard

MAGIK: Managing Completeness of Data

Bolzano, Italy

magik-demo.inf.unibz.it

May. 2012 - Oct. 2012

- Research area: information and knowledge management.
- Developed a system for managing, querying and visualizing database completeness.
- Led to 2 demo publications [12, 14]
- Supervised by Dr. Ognjen Savkovic and Prof. Werner Nutt.

COMPOSES: Compositional Operations in Semantic Space

Rovereto, Italy

EU-funded research project clic.cimec.unitn.it/composes

Feb. 2012 - Jun. 2012

- Research area: distributional semantics.
- Explored the benefits of combinatory categorial grammar for distributional semantic models.
- Led to Master's thesis and a conference publication [13].
- Supervised by Dr. Raffaella Bernardi and Prof. Roberto Zamparelli

Skills

Programming Python, C/C++, Scala, Java, Prolog, PHP, JSP, LaTeX

Databases MySQL, PostgreSQL, Microsoft SQL Server 2005

Others Weka, GATE, Apache Hadoop

Languages Indonesian (native), English (fluent), German (basic knowledge), Italian (basic knowledge)

Achievements & Awards

- | | | |
|------|------------------------------------------------------------------------------------------------------------|--------------------------------|
| 2014 | Travel scholarship , The 52nd Annual Meeting of the Association for Computational Linguistics (ACL) | Baltimore, USA |
| 2014 | Travel scholarship , European Chapter of the Association for Computational Linguistics (EACL) | Gothenburg, Sweden |
| 2013 | Travel scholarship (Google) , 3rd Lisbon Machine Learning Summer School | Lisbon, Portugal |
| 2010 | Erasmus Mundus scholarship , European Master's Program in Language and Communication Technologies | Nancy, France & Bolzano, Italy |
| 2006 | College scholarship , Samsung Electronics Indonesia | Bekasi, Indonesia |

Academic Activities

Reviewer, Journal of Natural Language Engineering

PC member, EMNLP 2015, COLING 2016, EMNLP 2017

References

Prof. Gerhard Weikum, Max Planck Institute for Informatics

weikum@mpi-inf.mpg.de

Dr. Sara Tonelli, Fondazione Bruno Kessler

satonelli@fbk.eu

Prof. Werner Nutt, Free University of Bozen-Bolzano

nutt@inf.unibz.it

Conference Publications

2017

- [1] **Paramita Mirza**, Simon Razniewski, Fariz Darari, and Gerhard Weikum. “Cardinal Virtues: Extracting Relation Cardinalities from Text”. In: *(to appear) Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers)*. Acceptance rate: 19.1%. Vancouver, Canada, July 2017.

2016

- [2] **Paramita Mirza** and Sara Tonelli. “CATENA: CAusal and TEmporal relation extraction from NATural language texts”. In: *Proceedings of COLING 2016, the 26th International Conference on Computational Linguistics: Technical Papers*. Acceptance rate: 32.4%. Osaka, Japan, Dec. 2016.
- [3] **Paramita Mirza** and Sara Tonelli. “On the contribution of word embeddings to temporal relation classification”. In: *Proceedings of COLING 2016, the 26th International Conference on Computational Linguistics: Technical Papers*. Acceptance rate: 32.4%. Osaka, Japan, Dec. 2016.
- [4] **Paramita Mirza**, Simon Razniewski, and Werner Nutt. “Expanding Wikidata’s Parenthood Information by 178%, or How To Mine Relation Cardinality Information”. In: *Proceedings of the ISWC 2016 Posters & Demonstrations Track*. Kobe, Japan, Oct. 2016.

2015

- [5] **Paramita Mirza** and Anne-Lyse Minard. “HLT-FBK: a Complete Temporal Processing System for QA TempEval”. In: *Proceedings of the 9th International Workshop on Semantic Evaluation (SemEval 2015)*. Denver, Colorado, June 2015.
- [6] **Paramita Mirza**. “Recognizing and Normalizing Temporal Expressions in Indonesian Texts”. In: *Proceedings of the 14th International Conference of the Pacific Association for Computational Linguistics*. PACLING '15. Acceptance rate: 40%. Bali, Indonesia, May 2015.

2014

- [7] **Paramita Mirza** and Sara Tonelli. “An Analysis of Causality between Events and its Relation to Temporal Information”. In: *Proceedings of COLING 2014, the 25th International Conference on Computational Linguistics: Technical Papers*. Acceptance rate: 30.8%. Dublin, Ireland, Aug. 2014.
- [8] **Paramita Mirza** and Sara Tonelli. “Classifying Temporal Relations with Simple Features”. In: *Proceedings of the 14th Conference of the European Chapter of the Association for Computational Linguistics*. Acceptance rate: 24.6%. Gothenburg, Sweden, Apr. 2014.
- [9] **Paramita Mirza** and Anne-Lyse Minard. “FBK-HLT-time: a complete Italian Temporal Processing system for EVENTI-Evalita 2014”. In: *Proceedings of the 4th International Workshop EVALITA-2014*. Pisa, Italy, Dec. 2014.
- [10] **Paramita Mirza**. “Extracting Temporal and Causal Relations between Events”. In: *Proceedings of the ACL 2014 Student Research Workshop*. Baltimore, Maryland, USA, June 2014.
- [11] **Paramita Mirza**, Rachele Sprugnoli, Sara Tonelli, and Manuela Speranza. “Annotating Causality in the TempEval-3 Corpus”. In: *Proceedings of the EACL 2014 Workshop on Computational Approaches to Causality in Language (CAtoCL)*. Gothenburg, Sweden, Apr. 2014.

2012 - 2013

- [12] Ognjen Savković, **Paramita Mirza**, Alex Tomasi, and Werner Nutt. “Complete Approximations of Incomplete Queries”. In: *Proc. VLDB Endow.* 6.12 (Aug. 2013).
- [13] **Paramita Mirza** and Raffaella Bernardi. “CCG Categories for Distributional Semantic Models”. In: *Proceedings of the International Conference Recent Advances in Natural Language Processing RANLP 2013*. Hissar, Bulgaria, Sept. 2013.
- [14] Ognjen Savković, **Paramita Mirza**, Sergey Paramonov, and Werner Nutt. “MAGIK: Managing Completeness of Data”. In: *Proceedings of the 21st ACM International Conference on Information and Knowledge Management*. CIKM '12. Maui, Hawaii, USA, 2012.