

# Diego Marcheggiani

## Curriculum Vitae

### Experience

- Jul 2017 – **Visiting Post-doc Researcher**, *University of Edinburgh - School of Informatics - Institute for Language, Cognition and Computation (ILCC)*, Edinburgh, UK.  
Now Exploiting and inducing linguistic structures with neural networks.
- Sep 2015 – **Post-doc Researcher**, *University of Amsterdam*, Institute for Logic, Language, and Computation (ILLC), Amsterdam, NL.  
Now Deep learning methods for semantic role labeling.
- Sep 2014 – **Visiting Researcher**, *University of Amsterdam*, Institute for Logic, Language, and Computation (ILLC), Amsterdam, NL.  
Nov 2014 Unsupervised relation extraction with variational discrete autoencoders.
- Oct 2012 – **Visiting Researcher**, *Université Pierre et Marie Curie*, Laboratoire d'informatique de Paris 6, Paris, FR.  
Dec 2012 Active learning with partially labeled sequences.
- Apr 2011 **Visiting Researcher**, *Swedish Institute of Computer Science (SICS)*, Stockholm, SE.  
Multi-label conditional random fields for opinion mining.
- Sep 2009 – **Research Fellow**, *ISTI-CNR*, Networked Multimedia Information Systems Laboratory, Human Language Technologies group, Pisa, IT.  
Jul 2015 Machine learning for sequence labeling.

### Education

- 2010–2014 **PhD in Computer Science**, *Università Ca Foscari di Venezia*, Venice, IT.  
Thesis topic: “Beyond Linear-chain: a Journey through Conditional Random Fields for Information Extraction from Text.”  
Supervisors: Fabrizio Sebastiani, Andrea Esuli.
- 2006–2009 **M.Sc. in Computer Science**, *Università di Pisa*, Pisa, IT.  
Thesis topic: “Active Learning for Information Extraction”  
Supervisors: Fabrizio Sebastiani, Andrea Esuli.
- 2002–2006 **B.Sc. in Computer Science**, *Università degli Studi di Perugia*, Perugia, IT.  
Curriculum in Computer Networks

### Service

#### Tutorial

- EMNLP 2017 Tutorial on Neural Methods for Semantic Role Labeling.

#### Program Committee

ACL, NAACL, EACL, EMNLP, COLING, ICLR, NIPS, ICML, WWW, IJCAI.

#### Other

- 2016–2017 University of Amsterdam–ILLC Deep Learning Reading Group.

---

## Teaching

- 2017 **Guest Lecturer**, *University of Amsterdam*, Unsupervised Language Learning. Recurrent Neural Networks and Encoder-Decoder Models.
- 2016 **Guest Lecturer**, *University of Amsterdam*, Unsupervised Language Learning. Recurrent Neural Networks and Encoder-Decoder Models, Memory Networks and Neural Turing Machines.

---

## Publications

Diego Marcheggiani and Laura Perez-Beltrachini. Deep graph convolutional encoders for structured data to text generation. In *Proceedings of the 11th International Conference on Natural Language Generation (INLG)*, pages 1–9, 2018.

Diego Marcheggiani, Joost Bastings, and Ivan Titov. Exploiting semantics in neural machine translation with graph convolutional networks. In *Proceedings of the the 16th Annual Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (NAACL-HLT)*, pages 486–492, 2018.

Diego Marcheggiani and Fabrizio Sebastiani. On the effects of low-quality training data on information extraction from clinical reports. *ACM Journal of Data and Information Quality*, 9(1):1:1–1:25, 2017.

Joost Bastings, Ivan Titov, Wilker Aziz, Diego Marcheggiani, and Khalil Simaan. Graph convolutional encoders for syntax-aware neural machine translation. In *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 1947–1957, 2017.

Diego Marcheggiani and Ivan Titov. Encoding sentences with graph convolutional networks for semantic role labeling. In *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 1507–1516, 2017.

Diego Marcheggiani, Anton Frolov, and Ivan Titov. A simple and accurate syntax-agnostic neural model for dependency-based semantic role labeling. In *Proceedings of the 21st Conference on Computational Natural Language Learning (CoNLL)*, pages 411–420, 2017.

Diego Marcheggiani, Ehsan Khoddam, and Ivan Titov. Unsupervised induction of shallow semantic representations with feature-rich models. In *Book of Abstracts of the 9th International Conference on Construction Grammar (ICCG)*, 2016.

Diego Marcheggiani and Ivan Titov. Discrete-state variational autoencoders for joint discovery and factorization of relations. *Transactions of the ACL (TACL)*, 4:231–244, 2016.

Diego Marcheggiani and Ivan Titov. Autoencoding for joint relation factorization and discovery from text. In *Proceedings of the 4th International Conference on Learning Representations - Workshop track (ICLR)*, 2016.

Salud M. Jiménez Zafra, Giacomo Berardi, Andrea Esuli, Diego Marcheggiani, Maria Teresa Martín-Valdivia, and Alejandro Moreo Fernández. A multi-lingual annotated dataset for aspect-oriented opinion mining. In *Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 2533–2538, 2015.

Giacomo Berardi, Andrea Esuli, and Diego Marcheggiani. Word embeddings go to Italy: A comparison of models and training datasets. In *Proceedings of the 6th Italian Information Retrieval Workshop (IIR)*, 2015.

Giacomo Berardi, Diego Ceccarelli, Andrea Esuli, and Diego Marcheggiani. On the impact of entity linking in microblog real-time filtering. In *Proceedings of the 30th Annual ACM Symposium on Applied Computing (SAC)*, pages 1066–1071, 2015.

Diego Marcheggiani, Oscar Täckström, Andrea Esuli, and Fabrizio Sebastiani. Hierarchical multi-label conditional random fields for aspect-oriented opinion mining. In *Advances in Information Retrieval - 36th European Conference on IR Research (ECIR). Proceedings*, pages 273–285, 2014.

Diego Marcheggiani and Thierry Artières. An experimental comparison of active learning strategies for partially labeled sequences. In *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 898–906, 2014.

Andrea Esuli, Diego Marcheggiani, and Fabrizio Sebastiani. An enhanced crfs-based system for information extraction from radiology reports. *Journal of Biomedical Informatics*, 46(3):425–435, 2013.

Giacomo Berardi, Andrea Esuli, and Diego Marcheggiani. Isti@trec microblog track 2012: Real-time filtering through supervised learning. In *Proceedings of The Twenty-First Text REtrieval Conference (TREC)*, 2012.

Giacomo Berardi, Andrea Esuli, Sergiu Gordea, Diego Marcheggiani, and Fabrizio Sebastiani. Metadata enrichment services for the europeana digital library. In *Proceedings of the Second International Conference on Theory and Practice of Digital Libraries (TPDL)*, pages 508–511, 2012.

Giacomo Berardi, Andrea Esuli, Diego Marcheggiani, and Fabrizio Sebastiani. Isti@trec microblog track 2011 exploring the use of hashtag segmentation and text quality ranking. In *Proceedings of The Twentieth Text REtrieval Conference (TREC)*, 2011.

Andrea Esuli, Diego Marcheggiani, and Fabrizio Sebastiani. Isti@semeval-2 task 8 boosting-based multiway relation classification. In *Proceedings of the 5th International Workshop on Semantic Evaluation (SemEval@ACL)*, pages 218–221, 2010.

Andrea Esuli, Diego Marcheggiani, and Fabrizio Sebastiani. Sentence-based active learning strategies for information extraction. In *Proceedings of the First Italian Information Retrieval Workshop (IIR)*, pages 41–45, 2010.