

MASSIMO NICOSIA

Staff Research Engineer - Google DeepMind

m.nicosia@gmail.com ♦ <http://www.linkedin.com/in/massimonicosia>

WORK EXPERIENCE

Google DeepMind

Staff Research Engineer

Nov 2022 - Present

Zürich, Switzerland

- Multimodal NLU in Gemini, i18n of semantic parsing models for the Assistant.

Google

Senior Software Engineer in Research

Nov 2020 - Nov 2022

Zürich, Switzerland

- i18n of semantic parsing models for the Assistant, conversational semantic parsing.

Google

Software Engineer in Research

July 2018 - Nov 2020

Zürich, Switzerland

- i18n of semantic parsing models for Search and YouTube.
-

Google

PhD Software Engineering Intern

June 2017 - September 2017

Zürich, Switzerland

- Adversarial training for text generation. Tech: Python (TensorFlow).
Host: *Aliaksei Severyn*.

Google

PhD Software Engineering Intern

June 2016 - October 2016

London, United Kingdom

- Dependency parsing with neural networks. Tech: C++ and Python (TensorFlow).
Hosts/Advisors: *Bernd Bohnet, Ryan McDonald, Michael Collins*.

Google

PhD Software Engineering Intern

June 2015 - November 2015

Zürich, Switzerland

- Sentence compression with neural networks (*dist_belief*) and the structured perceptron. Tech: C++.
Hosts: *Katja Filippova, Enrique Alfonseca*.

Qatar Computing Research Institute

Research Associate and Consultant

December 2013 - February 2015

Doha, Qatar

- Main designer and developer of a multilingual UIMA pipeline (English, Italian, Arabic) for Question Answering and other NLP tasks.
- Main designer and developer of the systems for the SemEval-2015 task 3 on Community Question Answering (1st system for Arabic and 3rd for English).
- Research on Question Answering, Automatic Crossword Solving, Machine Translation.

University of Trento

Research Intern

September 2012 - November 2013

Trento, Italy

- Development of a UIMA pipeline for deep natural language analysis and machine learning based re-ranking of candidate answers for Question Answering. Collaboration with the IBM Watson team.
- Participation in the Semantic Textual Similarity shared task at SemEval 2013.
- Collaboration with the University of Siena for improving the WebCrow crossword solver by application of machine learning methods and re-ranking.
- Research on Question Answering, Semantic Textual Similarity, Automatic Crossword Solving.

- Design and development of festivals' websites, newsletters and voting/polling applications.

HONORS AND AWARDS

Assistant Tech Impact Award H1 2023

For contributing to one of the most impactful projects in the Google Assistant.

Winner of the Zero-Shot MMNLU-22 Semantic Parsing Challenge organized by Amazon

Our winning submission was based on Translate & Fill, the synthetic data augmentation technique for semantic parsing models we published at EMNLP 2021.

ICT Days 2018 Best Poster Award

Prize awarded by Seac S.p.A. during the PhD Poster Competition organized by the University of Trento.

CIKM AnalytiCup 2017

Semi-finalist (top 10) at the CIKM 2017 Lazada Product Title Quality Challenge.

Honourable mention and travel award for the quality of the submission and technical report.

Google European Doctoral Fellowship 2015

Fellowship in Statistical Natural Language Processing.

Distinguished Young Paper CLiC-it 2014

A Retrieval Model for Automatic Resolution of Crossword Puzzles in Italian Language.

Merit Scholarship from the University of Trento

Prize for finishing the Master Course on time and with excellent results.

PUBLICATIONS

◇ Sebastian Ruder, Jonathan H. Clark, Alexander Gutkin, Mihir Kale, Min Ma, Massimo Nicosia, Shruti Rijhwani, Parker Riley, Jean-Michel A. Sarr, Xinyi Wang, John Wieting, Nitish Gupta, Anna Katanova, Christo Kirov, Dana L. Dickinson, Brian Roark, Bidisha Samanta, Connie Tao, David I. Adelani, Vera Axelrod, Isaac Caswell, Colin Cherry, Dan Garrette, Reeve Ingle, Melvin Johnson, Dmitry Panteleev, Partha Talukdar. *XTREME-UP: A User-Centric Scarce-Data Benchmark for Under-Represented Languages*. EMNLP Findings 2023.

◇ Jonas Pfeiffer, Francesco Piccinno, Massimo Nicosia, Xinyi Wang, Machel Reid, Sebastian Ruder. *mmT5: Modular Multilingual Pre-Training Solves Source Language Hallucinations*. EMNLP Findings 2023.

◇ Massimo Nicosia, Zhongdi Qu, Yasemin Altun. Translate & Fill: Improving Zero-Shot Multilingual Semantic Parsing with Synthetic Data. EMNLP Findings 2021.

◇ Thomas Müller, Francesco Piccinno, Massimo Nicosia, Peter Shaw, Yasemin Altun. Answering Conversational Questions on Structured Data without Logical Forms. EMNLP 2019.

◇ Massimo Nicosia, Alessandro Moschitti. *Semantic Linking in Convolutional Neural Networks for Answer Sentence Selection*. EMNLP 2018.

◇ Massimo Nicosia, Alessandro Moschitti. Accurate Sentence Matching with Hybrid Siamese Networks. CIKM 2017.

◇ Massimo Nicosia, Alessandro Moschitti. *Lazada Product Title Quality Challenge: Constructing Features for a Diversified Ensemble of Classifiers* (Honorable Mention Winner). CIKM AnalytiCup 2017.

◇ Massimo Nicosia, Alessandro Moschitti. *Learning Contextual Embeddings for Structural Semantic Similarity using Categorical Information*. CoNLL 2017.

◇ Kateryna Tymoshenko, Alessandro Moschitti, Massimo Nicosia and Aliaksei Severyn. *RelTextRank: An Open Source Framework for Building Relational Syntactic-Semantic Text Pair Representations*. ACL 2017.

◇ Alessandro Moschitti, Kateryna Tymoshenko, Panos Alexopoulos, Andrew Walker, Massimo Nicosia,

Guido Vetere, Alessandro Faraotti, Marco Monti, Jeff Z. Pan, Honghan Wu and Yuting Zhao. *Question Answering and Knowledge Graphs*. In *Exploiting Linked Data and Knowledge Graphs in Large Organizations*. Springer, 2017.

◊ Massimo Nicosia, Alessandro Moschitti. *Crossword Puzzle Resolution in Italian using Distributional Models for Clue Similarity*. IIR 2016.

◊ Aliaksei Severyn, Massimo Nicosia, Gianni Barlacchi, Alessandro Moschitti. *Distributional Neural Networks for Automatic Resolution of Crossword Puzzles*. ACL 2015.

◊ Gianni Barlacchi, Massimo Nicosia, Alessandro Moschitti. *SACRY: Syntax-based Automatic Crossword puzzle Resolution sYstem*. ACL 2015.

◊ Massimo Nicosia, Gianni Barlacchi, Alessandro Moschitti. *Learning to Rank Aggregated Answers for Crossword Puzzles*. ECIR 2015.

◊ Massimo Nicosia, Simone Filice, Alberto Barrón-Cedeño, Iman Saleh, Hamdy Mubarak, Wei Gao, Preslav Nakov, Alessandro Moschitti, Giovanni Da San Martino, Lluís Màrquez, Shafiq Joty, Walid Magdy. *QCRI: Answer Selection for Community Question Answering - Experiments for Arabic and English*. SemEval-2015 at NAACL 2015.

◊ Gianni Barlacchi, Massimo Nicosia, Alessandro Moschitti. *A Retrieval Model for Automatic Resolution of Crossword Puzzles in Italian Language*. CLiC-it 2014 (Distinguished Young Paper).

◊ Francisco Guzmán, Shafiq Joty, Lluís Màrquez, Alessandro Moschitti, Preslav Nakov and Massimo Nicosia. *Learning to Differentiate Better from Worse Translations*. EMNLP 2014.

◊ Gianni Barlacchi, Massimo Nicosia, Alessandro Moschitti. *Learning to Rank Answer Candidates for Automatic Resolution of Crossword Puzzles*. CoNLL 2014.

◊ Aliaksei Severyn, Massimo Nicosia, Alessandro Moschitti. *Building Structures from Classifiers for Passage Reranking*. CIKM 2013.

◊ Aliaksei Severyn, Massimo Nicosia, Alessandro Moschitti. *Learning Adaptable Patterns for Passage Reranking*. CoNLL 2013.

◊ Aliaksei Severyn, Massimo Nicosia, Alessandro Moschitti. *Learning Semantic Textual Similarity with Structural Representations*. ACL 2013.

◊ Aliaksei Severyn, Massimo Nicosia, Alessandro Moschitti. *iKernels-Core: Tree Kernel Learning for Textual Similarity*. *SEM 2013.

MENTORING

Deep Learning Indaba 2023 in Accra, Ghana

Presenter of a Google DeepMind RFP (Request For Plot), research program for facilitating collaborations between student and Google researchers; mentor during the Mentorship Breakfast sessions.

REVIEWER FOR INTERNATIONAL CONFERENCES

Acl Rolling Reviews (2021), EMNLP (2021, 2020, 2017), ACL (2021, 2020, 2019, 2018, 2017, 2015), EACL SRW (2021), CoNLL (2021, 2020, 2019, 2018), AACL (2020, 2019, 2018, 2015), ECIR (2020), NAACL HLT (2020, 2018, 2016, 2015), SIGIR (2018, 2017, 2016, 2014), CLiC-it (2018, 2017), IJCoL (2018), WWW (2017), IJCAI (2016), IEEE BigData Congress (2016), BioASQ workshop in CLEF (2015), IEEE SMC (2015).

EDUCATION

University of Trento, Italy

PhD in Computer Science

November 2014 - April 2018

Trento, Italy

Area: Natural Language Processing and Machine Learning

Thesis title: Structural Kernels and Neural Network Models for Question Answering Systems

Advisor: Prof. Alessandro MOSCHITTI

University of Trento, Italy

M.Sc. in Computer Science - Data, Media and Knowledge

Graduated with full marks and honour (110/110 cum laude)

September 2012

Trento, Italy

Award: Merit Scholarship for completing the Master Course on time and with excellent results.
Thesis title: *Adaptable Question Answering Pipeline Based On UIMA*
Advisor: Prof. Alessandro MOSCHITTI

University of Bari, Italy

B.Sc. in Computer Science - Technologies for software production

Graduated with full marks and honour (110/110 cum laude)

Thesis title: *Comfort indexes: a case study for managing knowledge in the environmental domain*

Advisor: Prof. Giovanni SEMERARO

July 2010

Bari, Italy

TEACHING EXPERIENCE

University of Trento, Italy

Advanced Natural Language Processing and Information Retrieval

2015, 2016, 2017, 2018

Trento, Italy

TECHNICAL SKILLS

Python, Java, C++, C, HTML, CSS, JavaScript, PHP, MySQL, XML, L^AT_EX, C#, Unity3D, RDF, OWL, R, Protégé, Git.

COURSES

Machine Learning Summer School (MLSS) 2016

Acceptance rate: 30%.

11-21st May 2016

Cádiz, Spain

Coursera

Machine Learning, The Data Scientist's Toolbox, R Programming, Getting and Cleaning Data, Exploratory Data Analysis, Reproducible Research.

E.N.A.I.P. (Professional training center), Monopoli, Italy

April 2010 - September 2010

Communication, marketing, legal and administrative requirements for starting a web agency.

OTHER INFORMATION

Languages: English (working proficiency) — Italian (mothertongue).

I participated in machine learning competitions at Kaggle.com (highest global rank achieved: 317th).

I play guitar, I have published an instrumental post-rock album and I am interested in music, movies and skateboarding.