Miguel Ballesteros. PhD – Curriculum Vitae

AI Architect - Oracle (OCI)

Personal Email: miguelballesteros1@gmail.com
Homepage: https://miguelballesteros.github.io/

EMPLOYMENT

AI Architect. Oracle (OCI). New York City, New York, United States.
February 2025 - Present

Principal Applied Scientist. Amazon Web Services. New York City, New York, United States.
October 2021 - February 2025

Senior Applied Scientist. Amazon Web Services. New York City, New York, United States.
July 2019 - October 2021

Research Scientist/Research Staff Member– IBM Research, Yorktown Heights New York, United States.
October 2016 - July 2019

Postdoctoral researcher – Visiting Professor – Department of Information and Communication
Technologies, at Pompeu Fabra University, in Barcelona. Spain.
January 2013 - October 2016

Visiting Postdoctoral researcher and adjunct instructor/lecturer at the Language Technologies Institute.
Carnegie Mellon University. Pittsburgh, U.S.
March 2015 – April 2016

PhD Student. Researcher. NIL research group (nil.fdi.ucm.es). Department of Software Engineering and
Artificial Intelligence. Complutense University of Madrid. Spain. September 2010-December 2012.

Scholarship grant. University Consortium Collaboration Grant: UNED, UCM, UC3M. Spain. May 2010-
September 2010.

Scholarship grant. Department of Informatic Systems and Computation. Complutense University of Madrid,
Spain. Granted by the Spanish Ministry of Education and Science. January 2009 - June 2009.

QUALIFICATIONS

PhD in Computational Linguistics (Computer Science), Complutense University of Madrid, Spain.
December 2012.

Master of Science (60 ECTS). Research in Computer Science, NLP Project, Complutense University of
Madrid, Spain.
June 2010.
This may be considered as the first year of the PhD studies.

Computer Science Engineering. Complutense University of Madrid, Spain.
Thesis title: TALFi: una herramienta para teoría de autómatas y lenguajes formales.
June 2009.
*This is a 5 year degree (350 ECTS). In the American system it would be equivalent to Bachelor and a Master with a
final project.*

LANGUAGES

Spanish (native)
English (fluent)
Italian (basic)

INTERNATIONAL RESEARCH STAYS

2014. Carnegie Mellon University, Pittsburgh, U.S. Supervised by Noah A. Smith and Chris Dyer.
1st October 2014 - 9th November 2014.

2013. Singapore University of Technology and Design, Singapore. Supervised by Yue Zhang.
14th June 2013 - 25th July 2013.

2012. Uppsala University, Sweden. Supervised by Joakim Nivre. Phd Student.
19th March 2012 - 19th April 2012.

2011. Uppsala University, Sweden. Supervised by Joakim Nivre. Phd Student.
3rd October 2011 - 23rd December 2011.

PUBLICATIONS

Please see publications at [Google Scholar](#)

More than 70 peer-reviewed papers and articles at ACL, EMNLP, NAACL, CL, TACL, EACL and related conferences/workshops.

ACADEMIC SUPERVISION

Graduate Supervision

Supervisor of Laura Perez Mayos' Doctoral Thesis (co-supervised with Leo Wanner). Phd Thesis. Indepth exploration of the syntactic capabilities of auto-encoding language modes for downstream applications. Universitat Pompeu Fabra, Barcelona, Spain. 2022

Many interns mentored at IBM and AWS (see publications).

Undergraduate Supervision

Supervisor of 3 final Degree Project of undergraduate students Pompeu Fabra University, in Barcelona, Spain. 2013-2015

TEACHING EXPERIENCE

Co-instructor of the Algorithms for Natural Language Processing Course 11711. Carnegie Mellon University. With Chris Dyer and Robert Frederking. Academic Year: 2015-2016

Instructor of (1) Discrete Mathematics and Graph Theory and main instructor of (2) Software Engineering for Web Applications. Computer Science Degree for the Department of Information and Communication Technologies, at Pompeu Fabra University, in Barcelona, Spain. Academic Year: 2014-2015

Instructor of (1) Discrete Mathematics and Graph Theory, main instructor of (2) Software Engineering for Web Applications, and instructor of (3) Databases. Computer Science Degree for the Department of Information and Communication Technologies, at Pompeu Fabra University, in Barcelona, Spain. Academic Year: 2013-2014

Instructor of (1) Discrete Mathematics and Graph Theory, main instructor of (2) Software Engineering for Web Applications, and instructor of (3) Databases. Computer Science Degree for the Department of Information and Communication Technologies, at Pompeu Fabra University, in Barcelona, Spain. Academic Year: 2012-2013

Teaching assistant in the Software Engineering subject (Ingeniería del Software) of the Computer Science Degree. Faculty of Computer Science. Complutense University of Madrid, Spain. Academic year: 2011-2012.

Online courses tutor for the Department of Software Engineering and Artificial Intelligence. Complutense University of Madrid, Spain. Academic year: 2011-2012.

SENIOR ACADEMIC SERVICES

Senior Area Chair and Area Chair. ACL Rolling Review. 2023 - present

Action Editor of Computational Linguistics Journal. 2020 - present

Action Editor of Transactions for the Association for Computational Linguistics (TACL Journal). 2019 - present

Senior Area Chair North American Chapter of the Association for Computational Linguistics (NAACL 2022)

Tutorial Chair. North American Chapter of the Association for Computational Linguistics. (NAACL 2022)

Area chair for ACL Rolling Review. 2021 - present.

Area chair of Tagging, Chunking, Syntax and Parsing of the Empirical Methods for Natural Language Processing Conference (EMNLP 2019).

Area chair. International Conference of Computational Linguistics (COLING 2018).

External Reviewer of Yiqing Zhang's Master Thesis. University of Melbourne, Australia. November 2018.

Area chair of Tagging, Chunking, Syntax and Parsing of the International Association for Computational Linguistics Conference (ACL 2016).

Member of Kepa Bengoetxea's PhD examination committee. [Dependency Parsing for Basque]. Basque Country University, San Sebastian, Spain. 16th December 2014.

Reviewer for Spanish National Projects. MICINN-ANEP research projects

AWARDS

Selected as one of the Best Reviewers (Best Reviewer Award).
Conference on Empirical Methods in Natural Language Processing. 2018.

Outstanding paper award.
What do Recurrent Neural Network Grammars Learn about Syntax?
European Chapter of the Association for Computational Linguistics. April 2017

Best paper award.
Towards the understanding of gaming audiences by modeling Twitch emotes
Third Workshop on Noisy User-generated Text (W-NUT 2017 at EMNLP 2017). September 2017

Best Student Paper Award
JADAWeb: A CBR System for Cooking Recipes. In Proceedings of the Computer Cooking Contest Workshop at the 18th International Conference on Case-based Reasoning. **Best Student Paper Award**. (CCC-ICCBR 2010).

PARTICIPATION IN PROJECTS (European, MIT-IBM)

MULTISENSOR: Mining and Understanding of multilingual content for intelligent sentiment enriched context and social oriented interpretation. (Program: FP7-ICT-610411. European Union)

KRISTINA: A knowledge based information agent with social competence and human interaction capabilities. (Program: H2020-645012-RIA. European Union)

Co-Pi of Hierarchically structured symbolic architectures with neural control for robust and human-like natural language parsing with Prof. Roger Levy. Exploratory Proposal. Funded.

INVITED TALKS

Research Seminar - Guest Lecture at Columbia University, Invited by Kathleen Mckeown. New York, United States. November 2022.

Invited speaker for the Automatic Summarization for Creative Writing workshop at COLING 2022, Gyeongju, Republic of Korea. October 2022.

Research Seminar at Spain AI, Spain (Virtual). 20th January 2022.

Research Seminar at Uppsala University, Sweden (Virtual). September 22nd 2021.

Research Seminar at University Politecnica de Catalunya, Spain (Virtual). September 5th 2020.

Research Seminar at Xerox Research Centre, France. Host: Xavier Carreras. September 8th 2016.
Topic: Transition-based Natural Language Processing.

UPF-DTIC Research Seminar (Universitat Pompeu Fabra, Barcelona, Spain). Host: Leo Wanner. May 10th 2016. Topic: Deep Learning for Transition-based Natural Language Processing.

Research Seminar at IBM Watson. Host: Yaser Al-onaizan. April 5th 2016.
Topic: Transition-based Natural Language Processing.

Research Seminar at IBM Almaden. Host: Laura Chiticariu.
March 28th 2016. Topic: Transition-based Natural Language Processing.

Research Seminar at Google New York. Host: Slav Petrov. March 23rd 2016.
Topic: Transition-based Natural Language Processing.

Research Seminar at Carnegie Mellon University. Host: Alan Black.
March 10th 2016.
Topic: Transition-based Natural Language Processing.

Research Seminar at Microsoft Research Redmond. Host: Chris Quirk.
March 3rd 2016. Topic: Greedy Transition-based Dependency Parsing with Stack-LSTMs.

Research Seminar at Microsoft Research Redmond. Host: Chris Quirk. October 9th 2015. Topic: Greedy Transition-based Dependency Parsing with Stack-LSTMs.

Seminar at the University of Washington. Host: Noah Smith. October 8th 2015.
Topic: Greedy Transition-based Dependency Parsing with Stack-LSTMs.

CLSP student seminar. Johns Hopkins University. Host: Jason Eisner. September 25th 2015. Topic: Greedy Transition-based Dependency Parsing with Stack-LSTMs.

Seminars of the IXA Research Group. Basque Country University, Spain. Host: Koldo Gojenola. December 19th 2014. Topic: Going to the Roots of Dependency Parsing.

Machine Learning Lunch Seminars. Carnegie Mellon University. Host: Noah A. Smith.
October 6th 2014. Topic: Going to the Roots of Dependency Parsing.

Seminars of the TALP Research Group. Universitat Politècnica de Catalunya, Spain. Host: Xavier Carreras. April 25th 2014. Topic: Going to the Roots of Dependency Parsing.

Seminars at the Singapore University of Technology and Design, Singapore. Host: Yue Zhang. July 12th 2013. Topic: Deep-Syntacting Parsing [Parsing into Deeper Syntactic Dependencies].

Seminars of the Department of Linguistics and Philology, Uppsala University, Sweden. Host: Joakim Nivre. March 30th 2012. Topic: MaltOptimizer: A System for MaltParser Optimization.

Department of Computer Science, Stockholm University, Kista, Sweden. Host: Hercules Dalianis. October 20th 2011. Topic: Enhancing and Applying Dependency Analysis in Different Topics: Negation, Speculation and Text Simplification.

Seminars of the Department of Linguistics and Philology, Uppsala University, Sweden. Host: Joakim Nivre. October 14th 2011. Topic: Enhancing and Applying Dependency Analysis in Different Topics.

UCNLG+Eval - The 4th UCNLG Workshop: Language Generation and Evaluation at EMNLP 2011, Edinburgh, United Kingdom. Host: Anja Belz. July 31st 2011.
Talk Title: A Proposal for a Spanish Surface Realization Shared Task. (UCNLG+Eval - EMNLP 2011)

PRESS RELEASES

MIT News Press Release - Link

Venturebeat - Link

El Periodico. Spanish newspaper - Link

El Economista. Spanish newspaper - Link