Daniel Ortiz Martínez

CONTACT INFORMATION

Webinterpret
Gran Vía Corts Catalanes
630 4th Floor
Barcelona, Spain

Phone: (+34) 93 270 25 86
E-mail: daniel.o@webinterpret.com

EDUCATION

MSc in Bioinformatics

University of Valencia, Valencia, Spain   Graduation date: September 2016

- Thesis title: Systems Biology Strategies to Study Cancer Metabolism
- Advisors: Joaquín Dopazo Blázquez and Vicente Arnau Llombart

PhD in Pattern Recognition and Artificial Intelligence

Technical University of Valencia, Valencia, Spain   Graduation date: October 2011

- Thesis title: Advances in Fully-Automatic and Interactive Phrase-Based Statistical Machine Translation
- Advisors: Francisco Casacuberta Nolla and Ismael García Varea

MSc in Pattern Recognition and Artificial Intelligence

Technical University of Valencia, Valencia, Spain   Graduation date: November 2005

- Thesis topic: Search Algorithms for Phrase-based Statistical Machine Translation

BSc in Computer Science Engineering

University of Castilla La Mancha, Albacete, Spain   Graduation date: September 2003

- Specialization in program of Pattern Recognition and Artificial Intelligence
- Thesis topic: Stack Decoding Algorithms for Statistical Machine Translation

PROFESSIONAL EXPERIENCE

Natural Language Processing Research Scientist   February 2016 to today

Webinterpret

- Research on new natural language processing and statistical machine translation techniques

Research Consultant (Data Scientist) for Webinterpret   April 2015 to January 2016

- Introduction of natural language processing techniques into the Webinterpret's workflow

Assistant Professor   December 2010 to January 2016

Technical University of Valencia, Valencia, Spain

- Statistics (undergrad level, Spanish and English): Courses 2011/12, 2012/13, 2013/14, 2014/15
- Operational Research (undergrad level, Spanish): Courses 2010/11, 2012/13

Post-Doctoral Researcher   February 2012 to December 2014

PRHLT Research Group, Technical University of Valencia, Valencia, Spain

- CASMACAT research project, funded by the 7th Framework Programme of the European Commission (funding amount: €2 500 000)
  - Online and active learning techniques for statistical machine translation
  - Scalable and parallel estimation of HMM models using Map-Reduce

Independent Expert (FP7 research project reviewer)   March 2011 to May 2011

European Commission

- Scientific reviewer of the EuromatrixPlus research project (budget: €5.94M), EuromatrixPlus belongs to the 7th Framework Programme of the European Union

Research Assistant   July 2008 to January 2012

Instituto Tecnol. de Informática, Technical University of Valencia, Valencia, Spain

- MIPRCV research project, part of the CONSOLIDER programme of the Spanish Government which hosted excellent research groups according to Nature (funding amount: €4 500 000)
  - Application of the incremental EM algorithm to the estimation of HMM models
  - Design and implementation of stochastic error correction models for string pairs

Computational Linguistic Researcher   March 2003 to June 2008

Technical University of Valencia, Valencia, Spain

- Participation in several research projects funded by the Spanish Government

RESEARCH STAYS	Three months research visit to the Lershtul für Informatik VI at Aachen University of Technology (RWTH) (Germany), from May 2006 to July 2006
SUPERVISED MSc AND PhD THESES	Jesús González Rubio. Technical University of Valencia. PhD Thesis. May 2014
	Álvaro Peris Abril. Technical University of Valencia. MSc Thesis. September 2014

SELECTED PUBLICATIONS

Daniel Ortiz-Martínez. Online Learning for Statistical Machine Translation. *Computational Linguistics*, 03/2016; DOI: 10.1162/COLI_a_00244

Antonio L. Lagarda, **Daniel Ortiz-Martínez**, Vicent Alabau, Francisco Casacuberta. Translating without In-domain Corpus: Machine Translation Post-Editing with Online Learning Techniques. *Computer Speech & Language Journal*, 11/2014; DOI: 10.1016/j.csl.2014.10.004

Daniel Ortiz-Martínez, Francisco Casacuberta. The New Thot Toolkit for Fully Automatic and Interactive Statistical Machine Translation. *Proceedings of the European Chapter of the Association for Computational Linguistics (EACL) conference*, Gothenburg, Sweden, April 2014

Jesús González-Rubio, **Daniel Ortiz-Martínez**, Francisco Casacuberta. Active learning for interactive machine translation. *Proceedings of the European Chapter of the Association for Computational Linguistics (EACL) conference*, Avignon, Paris, 2012

Daniel Ortiz-Martínez, Ismael García-Varea, Francisco Casacuberta. Online Learning for Interactive Statistical Machine Translation. *Proceedings of the North American Chapter of the Association for Computational Linguistics - Human Language Technologies (NAACL HLT) conference*, Los Angeles, US, 2010

Link to full publication list (currently comprises more than 50 research papers)

SELECTED INVITED TALKS

"Incremental Learning for Statistical Machine Translation", Workshop on Future Directions on Translation Research, Nara, Japan, December 2012

"Thot: New Features to Deal with Larger Corpora and Long Sentences", In TC-STAR OpenLab on Speech Translation Workshop, Trento, Italy, March 2006

SCIENTIFIC AND PROGRAM COMMITTEES

Scientific reviewer of international conferences and journals (IUI, ACL, COLING, Computer Speech and Language Journal, Knowledge Engineering Review Journal)

Member of the program committee of conferences and workshops (CAEPIA, ACL, AMTA)

OPEN SOURCE SOFTWARE

Thot Toolkit for Statistical Machine Translation
- Hosted on github http://daormar.github.io/thot/
- Provides search algorithms and scalable and parallel estimation of statistical models
- Funded by the European Commission and by the Spanish Government
- +50 000 lines of code

AWARDS AND SCHOLARSHIPS

Co-supervisor of the PhD Thesis "On the Effective Deployment of Currrent Machine Translation Technology" written by Jesús González Rubio, which received the LRC best PhD Thesis award of 2014

Honorable mention award in the Information and Communications Technology category at Valencia Idea competition, Spain, 2010

Best paper award for "Phrase-level alignment generation using a smoothed loglinear phrase-based statistical alignment model" given by the European Association for Machine Translation. Germany, 2008

Predoctoral Fellowship (FPI), public competitive call, Spanish Government, (2003-2007)

5th course of Computer Science studied in the Technical University of Valencia under the Seneca scholarship programme of the Spanish Government (2001-2002)

TECHNICAL SKILLS

Programming Languages: C, C++, Python, AWK, UNIX shell scripting
Machine Learning Toolkits: SNNS, SRILM, GIZA, Moses
Computer Applications: Statgraphics, LaTeX, most common productivity packages (Windows and Linux)